



TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

Lauri Wirola

**Studies on Location Technology Standards Evolution in
Wireless Networks**



Julkaisu 874 • Publication 874

Tampere 2010

Tampereen teknillinen yliopisto. Julkaisu 874
Tampere University of Technology. Publication 874

Lauri Wirola

Studies on Location Technology Standards Evolution in Wireless Networks

Thesis for the degree of Doctor of Technology to be presented with due permission for public examination and criticism in Rakennustalo Building, Auditorium RG202, at Tampere University of Technology, on the 12th of March 2010, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2010

ISBN 978-952-15-2328-1 (printed)
ISBN 978-952-15-2336-6 (PDF)
ISSN 1459-2045

Abstract

Positioning capability in mobile terminals (phones, laptops, netbooks) is required for various purposes including location-based commercial services, navigation and local search. Other uses include positioning of emergency calls and cellular networks also have internal use for the UE (User Equipment) location information. Moreover, the law and enforcement authorities have needs for positioning terminals and persons carrying them.

A variety of techniques to position a mobile UE is available in the cellular networks. The first category of methods includes utilizing the cellular base station coverage area information and combining that with distance estimates to the base stations. The distance estimates may be based on time delay, time difference or received signal strength measurements. The methods are discussed especially in the context of information they require from the cellular network.

The second category consists of using wireless networks to assist GNSS-based (Global Navigation Satellite System) methods. In the work the physics and prerequisites of the GNSS-based positioning methods are discussed and the significance of the assistance obtainable from the telecommunications networks is highlighted. The assistance is shown to significantly improve time-to-first-fix and sensitivity of the Assisted GNSS receiver.

The study introduces the positioning architectures and protocols in the 3GPP GERAN, UTRAN and E-UTRAN networks known as GSM, 3G and 4G in everyday language, respectively. It is shown that within each network there are dedicated logical and/or physical entities for positioning purposes and their significance to the introduced positioning methods is discussed.

In the discussion emphasis is given on the GERAN control plane positioning protocol called Radio Resource LCS Protocol (LCS for Location Services) also known as RRLP. Especially the GNSS-branch of RRLP is detailed, because its structure has also been copied to UTRAN and also to some extent to the E-UTRAN positioning protocols. It is shown that the generic GNSS structure in RRLP has made it straightforward to add the support for new GNSSs as they emerge. RRLP currently supports GPS in its entirety, GLONASS, QZSS, Galileo and various SBASs. In addition, the location protocol defined for the user plane, OMA SUPL (Open Mobile Alliance Secure User Plane Location protocol), is described and its relation to the 3GPP-defined positioning protocols is discussed.

A plethora of positioning techniques is shown not to be covered in the current wireless positioning standards. Examples include advanced GNSS-based methods, Precise Point Positioning and carrier phase -based relative positioning, which are envisioned to be introduced into the wireless positioning standards in the coming years. It is shown that in order to realize advanced GNSS-based methods in the cellular standards both new assistance data types and protocol features are needed.

Radiomap-based and fingerprint-based methods are also shown not to be covered adequately by the current positioning standards. Introducing the support for these methods is seen as an important milestone towards increasing the availability of the positioning services and solving the indoor positioning challenge. The thesis outlines protocol requirements for these methods as well.

Finally, costs and benefits of the proposed new positioning methods are critically analyzed. Considerations include accuracy, availability and bandwidth aspects. It is found that, in general, accuracy requirements also increase bandwidth consumption. This implies that the development of the new features should be concentrated in the user plane. Hence, OMA SUPL Release 3 together with the extended E-UTRAN positioning protocol are seen as the most promising platforms for introducing new positioning technologies into the wireless positioning standards. Moreover, the outlined standardization roadmap shows that the proposed features are realizable in the near future.

Preface

We choose to go the moon. We choose to go to the moon in this decade and do the other things, not because they are easy, but because they are hard, because that goal will serve to organize and measure the best of our energies and skills, because that challenge is one that we are willing to accept, one we are unwilling to postpone, and one which we intend to win, and the others too.

- President John F. Kennedy in Houston, 12th September, 1962

As seems to always happen I got into the positioning business by accident. In the spring of 2005 I was finishing my master's thesis on acoustics in the Nokia Research Center and figuring out what to do next. At the time the positioning group in Nokia had a temporary position open for a research engineer. I applied, was offered the position and decided to take my chances. I was confident that in a year or so I would be able to show my enthusiasm as well as skills and to become a permanent employee. I succeeded.

My history with GPS goes back to 1996, when I was 14 and I bought my very first GPS device. The Selective Availability degradation was still on and I had fun looking at the coordinates change and velocity fluctuate while sitting in the backyard. Then in 2000 everything changed as the Selective Availability degradation was dismantled. Over the years in university I did various assignments on GPS as well as Galileo and got familiar with the very basics of the satellite-based positioning. However, I never guessed that in a few years time my daily work would have to do with positioning.

Since 2005 I have had a pleasure to diversify my competences from satellite-based methods to cellular-based positioning as well as to location standards and various other technologies in the area. I feel privileged to work in a team that takes the positioning technologies from the university research to consumer products and industry standards. This, I believe, truly distinguishes our group from others. Therein I especially need to acknowledge my manager PhD Jari Syrjärinne, who I think outperforms the majority in both technological and managerial skills. He has been the main driver of this thesis work asking every now and then about the progress. Without Jari's commitment to publish and communicate our research and results this work would not have been possible. Moreover, Jari's comments to the thesis drafts are highly appreciated and have helped me to gradually improve the work. I also wish to express my deepest gratitude to Nokia Inc. for supporting the work in the form of paying the conference bills and being able to publish parts of my work.

My first touch with standardization took place in 2006 and my first 3GPP standardization meeting was in March 2007 in Montreal, Canada. I attended the meeting with my colleague M.Sc. Ismo Halivaara, who has ever since directed me in my daily work with standards. Many of the ideas presented in the thesis are results of the discussions that have most often taken place in Ismo's office hidden from the sunlight. I owe Ismo a big thank for commenting and helping me with the work.

In addition to Jari and Ismo, I would like to express my thanks to everyone I have had a pleasure to interact and work over the recent years in Nokia. Especially I would like to mention my colleagues Kimmo Alanen, Ilkka Kontola, Seppo Turunen, Tommi Laine, Mikko Blomqvist, Jani Käppi, Ville Eerola, Tony Höijer, Samuli Pietilä, Altti Jokinen, Harri Valio, Paula Syrjärinne, Tuomo Honkanen and Jari Mannermaa.

It is also not an accident that this thesis looks at the positioning domain from the perspective of physics. Physics became my favorite subject during the high school years in Tampereen lyseon lukio. Therein I had a pleasure to be taught physics by a truly inspiring teacher Heikki Juslen. This enthusiasm continued in Nokia

from 2001 onwards, when I worked under the supervision of Adjunct Professor Leo Kärkkäinen, who showed me how solid knowledge of physics can be applied to a variety of real-world problems.

The same spirit has also been reinforced by Professor Lauri Kettunen, who has always emphasized me how the practical problems we need to solve in everyday work actually require solid modeling from the first principles and knowledge of natural sciences. Lauri's door at the university has always been open, whenever I have wanted to discuss issues concerning my career or the progress of the thesis. Lauri's comments to the drafts have especially helped me to bring modeling perspective into the work.

I would also like to thank Professor Terry Moore and PhD Ville Ruutu for their constructive feedback on the manuscript. Moreover, Professor Christian Tiberius and Professor Sandra Verhagen from Delft University of Technology, The Netherlands, are acknowledged for providing me support during the work and co-authoring one of the publications that is included in the thesis.

In addition to scientific inspiration by Heikki, Leo and Lauri, I have also been inspired by the example and guidance by my uncle Matti Kähkönen. He has both supported me in my scientific ambitions, but also encouraged me to move away from my comfort zone and to expand my knowledge base to economics and business. It is a completely different perspective to the subject and I want to thank Matti for making me consider issues also from that angle.

The next set of thanks belongs to the people from various companies that I have had the possibility to interact in both 3GPP and OMA standardization fora. The non-comprehensive list of people I would like to acknowledge includes Andreas Wachter, Stephen Edge, Sven Fischer, Khiem Tran, James Winterbottom, Richard Wu, Haeyoung Jun, Norman Shaw, Javier De Salas, Kevin Judge, Giorgio Ghinamo and Michel Monnerat. Sometimes atmosphere is quite heated in the standardization meetings, but most of the time we work together to make location technologies better and afterwards have a drink. As I final note on standardization, I have found that standardization is excellent exercise for anyone interested in technology, processes and politics.

In addition to colleagues I have also had a chance to interact and share my worries as well as joys with my friends including Aki Virola, Tuomas Turto, Eero Niemelä, Jaakko Kuivanen, Ville Myllylä, Jaakko Vuento, Toni Ristimäki and Miika Huikkola. I owe all of them a big thank for joining me in fishing, hunting, hiking, dining, jogging, skiing, having sauna...

The biggest gratitude I have for my family that includes mother Tuula, father Hannu and sisters Katri and Laura as well as Katri's three daughters. My family has always been highly supportive of my studies and my parents have done wonderful job in always making sure that I have had all the prerequisites to study and to success. Finally, I wish to thank my beloved Laura for supporting me and taking good care of me in the final stages of this work.

I want to dedicate this work to my three wonderful nieces Mathilda, Beata and Adele.

Tampere, 28th January 2010,

Lauri Virola

Contents

Abstract	i
Preface	iii
Contents	vii
List of Publications	xi
List of Supplementary Publications	xiii
List of Abbreviations and Symbols	xv
Abbreviations	xv
Symbols, Greek alphabet	xxii
Symbols, other	xxiii
List of Figures	xxvii
List of Tables	xxix
1 Introduction	1
1.1 Regulatory drivers	3
1.2 Commercial drivers	4
1.3 GNSS evolution drivers	5
1.4 RAN evolution drivers	6
1.5 Brief history of positioning standards	6
1.6 Scope	8

2	Overview of relevant positioning methods	9
2.1	Satellite-based methods	9
2.1.1	Physical models involved	9
2.1.2	GNSS descriptions	11
2.1.2.1	GPS	11
2.1.2.2	GLONASS	12
2.1.2.3	Galileo	13
2.1.2.4	QZSS	14
2.1.2.5	Space-Based Augmentation Systems	14
2.1.2.6	Compass	15
2.1.2.7	IRNSS	16
2.1.3	Point positioning	16
2.1.4	Precise Point Positioning	20
2.1.5	Relative methods	22
2.1.5.1	Measurement model	22
2.1.5.2	Observables	24
2.1.5.3	Baseline solution	25
2.2	Assisted GNSS	27
2.2.1	Case for assistance	27
2.2.2	AGNSS infrastructure	29
2.2.3	Effect of obtaining navigation model	30
2.2.4	Effect of reference position, time and frequency	31
2.2.5	Effect of data bit assistance	34
2.2.6	Effect of integrity information	35
2.2.7	Effect of correction data	35
2.3	Methods for Radio Access Networks	35
2.3.1	Coverage area -based methods	36
2.3.2	Ranging-based methods	38
2.3.3	Time Difference of Arrival	42
3	Standards	47
3.1	Third Generation Partnership Project	47
3.2	3GPP Location Services	48
3.2.1	Stage 1 description of LCS	48

3.2.2	Stage 2 description of LCS	49
3.2.2.1	GERAN	50
3.2.2.2	UTRAN	52
3.2.2.3	E-UTRAN	55
3.2.2.4	GERAN - UTRAN - E-UTRAN interworking . . .	57
3.2.3	Stage 3 description of LCS	57
3.2.3.1	GERAN RRLP	57
3.2.3.2	GERAN LCS broadcast	64
3.2.3.3	UTRAN RRC	65
3.2.3.4	E-UTRAN LPP	66
3.2.4	Minimum performance requirements	68
3.3	Open Mobile Alliance	69
3.3.1	Secure User Plane Location protocol	69
3.3.1.1	Releases 1 and 2	69
3.3.1.2	Upcoming Release 3	72
4	Way forward	75
4.1	Harmonization of AGNSS information	75
4.1.1	SV clock models	77
4.1.2	GST - UTC models	78
4.1.3	Inter-GNSS time models	80
4.1.4	Fine time assistance	81
4.1.5	Performance aspects	83
4.2	Missing information elements in positioning standards	85
4.2.1	Navigation models	85
4.2.2	Navigation model request	87
4.2.3	Ionosphere models	88
4.2.4	Troposphere models	91
4.2.5	Radiomaps and fingerprints	96
4.2.6	Support for RTK and PPP	101
4.2.7	Assistance stream and IP broadcast	104
4.2.8	Assistance data push	106
4.2.9	Other issues	109
4.3	Costs and Benefits	110

4.4 Roadmap	113
5 Contribution	115
5.1 Author's contribution to the publications	115
5.2 Author's contribution to the positioning standards	120
6 Conclusions	121
Bibliography	124
Publications	139

List of Publications

This thesis is based on the following papers published in open literature.

- [P1] L. Wirola and K. Alanen and J. Käppi and J. Syrjärinne. Bringing RTK to Cellular Terminals Using a Low-Cost Single-Frequency AGPS Receiver and Inertial Sensors. In *Proceedings of IEEE ION PLANS 2006*, San Diego, USA, April 25th-27th, pp. 645–652, (2006)
- [P2] L. Wirola and S. Verhagen and I. Halivaara and C. Tiberius. On the feasibility of adding carrier phase assistance to cellular GNSS assistance standards. *Journal of Global Positioning Systems*, Vol. 6, No. 1, pp. 1–12, (2007)
- [P3] L. Wirola and I. Kontola and J. Syrjärinne. The Effect of the Antenna Phase Response on the Ambiguity Resolution. In *Proceedings of IEEE ION PLANS 2008*, Monterey, USA, May 6th-8th, pp. 606–615, (2008)
- [P4] J. Syrjärinne and L. Wirola. Setting a New Standard: Assisting GNSS Receivers That Use Wireless Networks. *InsideGNSS*, October issue, Vol. 1, No. 7, pp. 26–31, (2006)
- [P5] L. Wirola and J. Syrjärinne. Bringing the GNSSs on the Same Line in the GNSS Assistance Standards. In *Proceedings of ION Annual Meeting*, Boston, USA, April 23rd-25th, pp. 242–252, (2007)
- [P6] L. Wirola and J. Syrjärinne. GLONASS Orbits in GPS/Galileo-style Ephemerides for Assisted GNSS. In *Proceedings of ION National Technical Meeting*, San Diego, USA, January 28th-30th, pp. 1032–1039, (2008)
- [P7] L. Wirola and I. Halivaara and J. Syrjärinne. Requirements for the next generation standardized location technology protocol for location-based services. *Journal of Global Positioning Systems*, Vol. 7, No. 2, pp. 91–103, (2008)

List of Supplementary Publications

Other publications related to this work, but not included in this thesis:

- [S1] K. Alanen and L. Wirola and J. Käppi and J. Syrjärinne. Inertial Sensor Enhanced Mobile RTK Solution Using Low-Cost Assisted GPS Receivers and Internet-Enabled Cellular Phones. In *Proceedings of IEEE ION PLANS 2006*, San Diego, USA, April 25th-27th, pp. 920–926, (2006)
- [S2] K. Alanen and L. Wirola and J. Käppi and J. Syrjärinne. Using Low-Cost GPS and Internet-Enabled Wireless Phones. *InsideGNSS*, May/June issue 2006, pp. 32–39, (2006)
- [S3] L. Wirola and J. Syrjärinne. Introduction to Multi-Mode Navigation Model. *Discussion paper presented in 3GPP TSG GERAN2#29bis*, Sophia-Antipolis, France, May 22nd-24th, (2006)
- [S4] L. Wirola and J. Syrjärinne. Bringing All GNSS into Line. *GPS World*, Vol. 18, No. 9, pp. 40–47, (2007)

List of Abbreviations and Symbols

Abbreviations

3GPP	Third Generation Partnership Project
ADR	Accumulated Delta Range
AGANSS	Assisted GANSS
AGNSS	Assisted GNSS
AGPS	Assisted GPS
AMPS	Advanced Mobile Phone Service
AOA	Angle-Of-Arrival
ASN	Abstract Syntax Notation
ATIS	Alliance for Telecommunications Industry Solutions
AWGN	Additive White Gaussian Noise
BCCH	Broadcast Control Channel
BS	Base Station
BSC	Base Station Controller
BSIC	Base Station Identity Code
BSS	Base Station Subsystem
BTS	Base Transceiver Station
C/A	Coarse/Acquisition
C-RNC	Controlling RNC
CBC	Cell Broadcast Center
CDMA	Code Division Multiple Access
CID	Cell Identity

CNSS	Chinese Navigation Satellite System
CODE	Center for Orbit Determination in Europe
DGANSS	Differential GANSS
DGNSS	Differential GNSS
DGPS	Differential GPS
DLL	Delay-Locked Loop
DOY	Day-Of-Year
E911	Enhanced 911
E-OTD	Enhanced OTD
E-SLP	Emergency SLP
E-SMLC	Evolved SMLC
E-UTRAN	Evolved UTRAN
ECEF	Earth-Centered Earth-Fixed coordinate system
ECI	Earth-Centered Inertial coordinate system
ECID	Enhanced CID
EDGE	Enhanced Data rates for Global Evolution
EGNOS	European Geostationary Navigation Overlay Service
EKF	Extended Kalman Filter
eNode B	Base Station in E-UTRAN
ENU	East-North-Up coordinate system
EOP	Earth-Orientation Parameters
EUI	Extended Unique Identifier
FCC	Federal Communication Commission
FDD	Frequency-Division Duplex
FDMA	Frequency Division Multiple Access
GAGAN	GPS-Aided Geo-Augmented Navigation
GANSS	Galileo and Additional Navigation Satellite Systems
GEO	Geostationary Earth Orbit
GERAN	GSM/EDGE RAN

GLONASS	Globalnaya Navigatsionnaya Sputnikovaya Sistema
GMLC	Gateway Mobile Location Center
GNSS	Global Navigation Satellite System
GPRS	General Packet Radio Service
GPS	Global Positioning System
GSM	Global System for Mobile communications
GSO	Geosynchronous Orbit
GST	GNSS System Time
H-SLP	Home SLP
HDOP	Horizontal Dilution Of Precision
HELD	HTTP-Enabled Location Delivery
HRPD	High-Rate Packet Data
HTTP	Hypertext Transfer Protocol
HW	Hardware
I-WLAN	Interworking WLAN
ICD	Interface Control Document
IETF	Internet Engineering Task Force
IF	Ionosphere-Free
IGS	International GNSS Service
IGSO	Inclined GSO
IMES	Indoor Messaging System
IP	Internet Protocol
IRNSS	Indian Regional Navigation Satellite System
Jaxa	Japanese Aerospace Exploration Agency
LAC	Local Area Code
LAMBDA	Least-squares Ambiguity Decorrelation Adjustment
LBS	Location Based Services
LCS	Location Services
LCS-AP	LCS Application Protocol

LEX	L-band Experimental Signal
LMU	Location Measurement Unit
LPP	LTE Positioning Protocol
LPPa	LPP Annex
LR	Location Request
LTE	Long-Term Evolution
LTE-A	LTE Advanced
MAC	Media Access Control
MCC	Mobile Country Code
MEMS	Micro-Electromechanical System
MEO	Medium Earth Orbit
MME	Mobility Management Entity
MNC	Mobile Network Code
MO-LR	Mobile-Originated Location Request
mRTK	mobile RTK
MS	Mobile Station (UE in GERAN vocabulary)
MSAS	Multi-Functional Satellite Augmentation System
MSB	Most Significant Bit
MT-LR	Mobile-Terminated Location Request
NI-LR	Network-Induced Location Request
NICT	National Institute of Information and Communications Technology
NMR	Network Measurement Result
Node B	Base Station in UTRAN
NTCM	Neustrelitz TEC Model
OMA	Open Mobile Alliance
OTD	Observed Time Difference
OTDOA	Observed TDOA
OTDOA-IPDL	OTDOA Idle Period Downlink

P-CPICH	Primary Control Pilot Channel
PDCP	Packet Data Convergence Protocol
PLL	Phase-Locked Loop
PPP	Precise Point Positioning
PRN	Pseudorandom Number
PSAP	Public Safety Answering Point
PVT	Position-Velocity-Time
QOS	Quality of Service
QZSS	Quazi-Zenith Satellite System
RAN	Radio Access Network
RF	Radio Frequency
RMS	Root Mean Squared
RNC	Radio Network Controller
RRC	Radio Resource Control protocol
RRLP	Radio Resource LCS Protocol
RSS	Received Signal Strength
RTD	Real-Time Difference
RTI	Real-Time Integrity
RTK	Real-Time Kinematic
RTT	Round-Trip Time
S1-AP	S1 Application Protocol
S-RNC	Serving RNC
SAIF	Submeter-class Augmentation with Integrity Function
SAS	Standalone SMLC
SBAS	Space-Based Augmentation System
SCTP	Stream Control Transmission Protocol
SDCM	System of Differential Correction and Monitoring
SET	SUPL-Enabled Terminal

SFN	System Frame Number
SIM	Subscriber Identification Module
SLC	SUPL Location Center
SLP	SUPL Location Platform
SMLC	Serving Mobile Location Center
SMS	Short Messaging Service
SMSC	SMS Center
SNR	Signal-to-Noise Ratio
SPC	SUPL Positioning Center
SS	Supplementary Service
STD	Slant Total Delay, troposphere
SUPL	Secure User Plane Location
SV	Space Vehicle
TA	Timing Advance
TAI	International Atomic Time (Temps Atomique International)
TCP	Transport Control Protocol
TD-SCDMA	Time-Division Synchronous CDMA
TDD	Time-Division Duplex
TDMA	Time-Division Multiple Access
TDOA	Time Difference Of Arrival
TEC	Total Electron Content
TIA	Telecommunications Industry Association
TLS	Transport Layer Security
TOA	Time-Of-Arrival
TOD	Time-Of-Day
TOT	Time-Of-Transmission
TOW	Time-Of-Week
TP	Technical Plenary

TS	Technical Specification
TTFF	Time-To-First-Fix
U-TDOA	Uplink TDOA
UARFCN-DL	UTRAN Absolute Radio Frequency Channel Downlink
UE	User Equipment
ULP	User plane Location Protocol
UMB	Ultra Mobile Broadband
UMTS	Universal Mobile Telecommunications System
URA	User Range Accuracy
USNO	US Naval Observatory
UTC	Universal Time Coordinated
UTRAN	UMTS Terrestrial RAN
V-SLP	Visited SLP
VMF	Vienna Mapping Function
VRR	Virtual Reference Receiver
WAAS	Wide Area Augmentation System
WAP	Wireless Application Protocol
WAP PPG	WAP Push Proxy Gateway
WARN	Wide Area Reference Network
WCDMA	Wideband CDMA
WiMAX	Worldwide Interoperability for Microwave Access
WL	Wide-Lane
WLAN	Wireless Local Area Network
WTSC	Wireless Technologies and Systems Committee
ZHD	Zenith Hydrostatic Delay
ZWD	Zenith Wet Delay

Symbols, Greek alphabet

α	Azimuth of travel
β	Elevation (angle) of travel
Δ_j^i	OTDOA of the i^{th} BS with respect to j^{th} BS
ϵ	Measurement noise
ϵ_0	Permittivity of free space
ϵ_ϕ	ADR measurement noise
ϵ_ρ	Pseudorange noise
$\epsilon_{\dot{\rho}}$	Range-rate noise
ϵ_{SF}	Slow fading
ϕ	ADR measurement
ϕ^{IF}	ADR ionosphere-free linear combination
$\phi^{\lambda_{WL}}$	ADR wide-lane observable
γ	Dispersion factor
λ	Wavelength
λ_{L1}	Wavelength at L1 frequency
λ_{L2}	Wavelength at L2 frequency
λ_{WL}	Wide-lane wavelength
ψ	Longitude
ν	SV elevation
ρ	Pseudorange
ρ_{L1}	Pseudorange at L1
ρ_{L2}	Pseudorange at L2
ρ_{IF}	Pseudorange ionosphere-free
$\dot{\rho}$	Range-rate
θ_i	Auxiliary information of the i^{th} fingerprint
φ	Latitude

Symbols, other

AG_{BS}	Base station antenna gain
A_{BS}	Base station antenna gain pattern
\underline{x}_i	Base station position (i^{th})
t_i	Base station time (i^{th})
\underline{b}	Baseline vector
Q	Covariance
H_i	Diurnal variation
r	Double difference of geometric ranges
\underline{r}	Double difference of geometric ranges (vector)
f_k	GLONASS frequency channel
$G_{\underline{n}}, G_{\underline{e}}$	Gradients along north and east directions
t_{gd}	Group delay
c_g	Group velocity
e^-	Electron
q_e	Electron charge
N_e	Electron density
m_e	Electron mass
F_{10}	F10 10.7-cm Solar flux
F_i	Fingerprint (i^{th})
F	Fingerprint database
\underline{p}_i	Fingerprint position (i^{th})
f	Frequency
t_{GST}	GNSS System Time
$t_{GST_1}(t_{GST_0})$	GST_1 at the time GST_0
h	Height
n	Index of refraction

n_g	Index of refraction (group)
n_ϕ	Index of refraction (phase)
\mathbb{F}	Index set
N	Integer ambiguity
$\underline{N}^{\lambda_{WL}}$	Integer ambiguity, double difference, wide lane
$\underline{\tilde{N}}^{\lambda_{WL}}$	Integer-ambiguity, double difference, wide lane, integer-valued estimate
$\underline{\hat{N}}^{\lambda_{WL}}$	Integer-ambiguity, double difference, wide lane, real-valued estimate
$\Delta t_I, \Delta t'_I$	Integer second time difference
I	Ionosphere delay
I_{L1}	Ionosphere delay at L1
I_{L2}	Ionosphere delay at L2
N_I	Ionospheric refractivity
$p(\underline{y} \underline{x}^{UE})$	Likelihood
m_h	Mapping function, hydrostatic
m_w	Mapping function, wet
t_{GST}^{ref}	Model reference time
$p(\underline{y})$	Normalization factor
n_{SV}	Number of SVs observed
\underline{y}	Observation
PL	Path loss
c_ϕ	Phase velocity
f_N	Plasma frequency
a_i	Polynomial coefficient (i^{th})
\underline{x}	Position
P_{Losses}	Power losses
$p(\underline{x}^{UE})$	Prior distribution for the UE position

$p(\underline{x}^{UE} \underline{y})$	Probability distribution for the UE position conditioned on the observation \underline{y}
δt_j^i	RTD of the i^{th} BS vs j^{th} BS
Y_j	Seasonal variation
N_i	Set of communication nodes heard at the i^{th} fingerprint
\underline{a}_{ij}	Signal measurement record at the i^{th} fingerprint
s	Signal path
S_l	Solar activity dependency
L_k	Solar zenith angle dependency
c	Speed of light in vacuum
δt_{SV}^{GST}	SV clock bias with respect to GST
$\delta \dot{t}_{SV}^{GST}$	SV clock drift with respect to GST
p, q	SV indices
\underline{x}_{SV}	SV position
$\dot{\underline{x}}_{SV}$	SV velocity
δt_i^{UE}	Time bias between the UE and the i^{th} base station clocks
k	Time uncertainty index
P_0	Transmit power
T	Troposphere delay
h_T	Troposphere effective height
a_h, b_h, c_h	Troposphere mapping function parameters, hydrostatic
a, b, c	Troposphere mapping function parameters, hydrostatic (Niell's)
a_w, b_w, c_w	Troposphere mapping function parameters, wet
h_0	Troposphere reference height
N_h	Tropospheric refractivity, hydrostatic
N_w	Tropospheric refractivity, wet
AG^{UE}	UE antenna gain

A^{UE}	UE antenna gain pattern
δt_{GST}^{UE}	UE clock bias with respect to GST
$\delta \dot{t}_{GST}^{UE}$	UE clock drift with respect to GST
k, m	UE indices
O^{UE}	UE orientation
\underline{x}^{UE}	UE position
t^{UE}	UE time
$\underline{\dot{x}}^{UE}$	UE velocity
$t_{uncertainty}$	Uncertainty in time assistance
\hat{r}	Unit vector
\underline{e}	Unit vector in ENU system, east
\underline{n}	Unit vector in ENU system, north
\underline{u}	Unit vector in ENU system, up
t_{UTC}	UTC time
w_i	Weight (i^{th})

List of Figures

1.1	Future GNSS landscape.	7
2.1	Loss of data payload.	28
2.2	High-level AGNSS infrastructure.	30
2.3	UE in the intersection of three wireless communication nodes.	37
2.4	Timing Advance.	40
2.5	Principle of hyperbolic positioning.	43
2.6	Hyperboloids in TDOA positioning.	46
3.1	GERAN LCS architecture.	51
3.2	GERAN LCS protocol stack.	52
3.3	UTRAN LCS architecture.	54
3.4	E-UTRAN LCS architecture.	56
3.5	E-UTRAN LCS protocol stack: LPP.	56
3.6	E-UTRAN LCS protocol stack: LPPa.	57
3.7	RRLP assistance data structure.	60
3.8	OMA SUPL protocol stack.	70
3.9	OMA LCS architecture.	71
4.1	Bandwidth - Accuracy - Availability comparison.	112
4.2	Location technology standards roadmap.	114

List of Tables

1.1	FCC E911 requirements.	3
2.1	Data broadcast by SBAS SVs.	15
2.2	GNSS signal attenuation.	29
3.1	Items in the Generic Assistance element in RRLP.	62
3.2	Sensitivity test with fine time assistance.	69
3.3	Pass criteria for the sensitivity test with fine time assistance.	69
3.4	SUPL Release 3 Work Item contents.	73
4.1	GNSS System Times.	76
4.2	Native SV clock models and mappings to the harmonized model.	77
4.3	Harmonized SV clock model.	78
4.4	UTC standards.	79
4.5	Native UTC-GST models and mappings to the harmonized model.	79
4.6	Harmonized UTC model.	80
4.7	Native inter-GNSS time models and mappings to the harmonized model.	81
4.8	Harmonized GNSS-GNSS time model.	81
4.9	Harmonized fine time assistance.	83
4.10	Penalty due to model harmonization.	84
4.11	Contents of troposphere assistance.	94
4.12	RTK measurement message.	103
4.13	Requirements for the assistance data push.	108

Chapter 1

Introduction

Developing positioning and location standards has substantial market demand and various regulatory drivers in several countries. Already now AGPS-enabled (Assisted Global Positioning System) mobile terminals constitute a significant share of the global navigation device market. In the second quarter of 2009 80% of the smart phones sold worldwide had an AGPS functionality corresponding to approximately 30 million units [1]. Moreover, in Europe and Middle-East region up to 90% of the smart phones sold are already equipped with AGPS [2] [3]. Moreover, modern smart phones are location-aware at least through the cellular network base station information and some units can also utilize WLAN (Wireless Local Area Network) access points in positioning.

When discussing the positioning and location features in various devices it is commonplace to mix positioning technology with applications and value-added services. A clear distinction must be made here. The term positioning technologies refers to the set of positioning infrastructure, radio signals and algorithms that provide position, velocity and time information to the UE (User Equipment). In consumer appliances providing only the plain coordinates has little use and, hence, on top of the positioning technologies one needs location applications that utilize the position information. Such applications include geo-tagging of photos and turn-by-turn guidance by a navigation equipment. This work deals with the specific aspects of obtaining the plain position data.

One method of providing position information is through utilizing the positioning methods standardized in wireless telecommunication networks. Such methods utilize both the features of the telecommunication networks (radio signals and protocol signaling) as well as provide data that assist in utilizing positioning systems external to telecommunications networks. These external technologies include GNSS-based (Global Navigation Satellite System) methods that are assisted by the information from the telecommunication network. The assistance includes, among other things, navigation models (orbit and clock parameters) for the SVs (Space Vehicle), reference location and reference time. In an assisted situation the UE needs not to receive the navigation models from the SVs, but the UE retrieves the models over the cellular network considerably reducing TTFF (time-to-first-fix). Moreover, reference location and time data improve sensitivity significantly. The positioning is thus enabled in adverse signal conditions such as urban canyons. The improvement in user experience is significant compared to the performance of the autonomous GPS or simple CID-based (Cell Identity) positioning.

Positioning protocol standardization is nowadays concentrated in 3GPP (The Third Generation Partnership Project, [4]) and 3GPP2, which define positioning protocols for the control planes of GERAN (GSM/EDGE Radio Access Network), UTRAN (UMTS Terrestrial RAN), E-UTRAN (Evolved UTRAN) and CDMA/CDMA2000 (Code Division Multiple Access) radio access networks. GERAN, where EDGE stands for Enhanced Data rates for Global Evolution, is better known as GSM (Global System for Mobile communications). UTRAN, where UMTS stands for Universal Mobile Telecommunications System, is commonly referred to as 3G or WCDMA (Wide-band CDMA). To be precise, UTRAN supports a variety of radio access technologies including FDD (Frequency-Division Duplex) and TDD (Time-Division Duplex) with the UTRA-FDD being the WCDMA access technology. Finally, E-UTRAN is better known as LTE (Long-Term Evolution) or 4G.

Solutions for IP-networks (Internet Protocol) include OMA (Open Mobile Alliance, [5]) SUPL (Secure User Plane Location protocol) Release 1 [6] and draft Release 2 [7] that encapsulate control plane positioning protocols defined

Table 1.1: FCC E911 requirements.

Percentage of calls	UE-involved	Network-based
67%	50 m	150 m
95%	150 m	300 m

by 3GPP/3GPP2 as sub-protocols to ULP (User plane Location Protocol). In addition to the capabilities of 3GPP and 3GPP2 positioning protocols the (draft) SUPL Release 2 adds items for, for instance, WLAN-based and WiMAX-based (Worldwide Interoperability for Microwave Access) positioning.

The drivers for the positioning technology standardization in wireless telecommunication networks include legislative and commercial reasons as well as GNSS and RAN (Radio Access Network) evolution. These drivers are now considered in turn.

1.1 Regulatory drivers

The regulatory drivers for standardizing positioning technologies for telecommunication networks include emergency and lawful intercept LCS (Location Services) [8]. The Lawful Interception LCS supports authorities to locate and track desired persons and targets.

The emergency LCS refers to providing the position of the UE to the emergency authorities in case an emergency call is made. The FCC (Federal Communications Commission, [9]) Wireless E911 (Enhanced 911) ruling mandates the emergency call positioning in the United States. The resolution applies to both fixed and wireless connections. It also includes that the mobile operators in the United States may only sell their customers location-capable UEs and that UEs as well as networks must comply with the accuracy and availability requirements [10] summarized in Table 1.1. The table shows that the accuracy requirements are different depending upon, whether the UE is directly involved in the position determination or if positioning is performed by the network.

The requirements are currently enforced in the US so that 100% coverage is required from the operators relying on the network-based methods. In contrast, in the UE-involved case 95% penetration of location capable UEs is required. However, the initial requirements did not specify the method of statistics calculation and currently the accuracy figures are averaged over the area covered by the operator. Therefore, the FCC has drafted new requirements that the mobile operators need to comply with the accuracy requirements in each PSAP (Public Safety Answering Point) service area individually by September 11th 2012. Several companies have, however, stated that the current positioning technologies may be insufficient to fulfill the requirements [11].

The role of standardized positioning technologies in these LCS applications is to provide harmonized performance irrespective of the UE and the network. In the lack of standards each equipment manufacturer would be forced to define their own technologies and implementations leading to high cost as well as problems in interoperability and performance. The key factor in legislative use cases is that both the end user and the authorities can trust on the location performance in terms of guaranteed level of availability and accuracy.

1.2 Commercial drivers

The specification [8] identifies two categories of commercial LCS in telecommunications networks. The first category consists of the value-added services that may include providing the customers services based on location. An example of such LBS (Location-Based Service) are commercials that are triggered, when the UE enters a predefined geographical area. The second category is internal to the network to which the UE is connected to. These internal use cases may include positioning of UEs for traffic and coverage area measurements for the purposes of, say, network capacity planning.

The Canalys mobile navigation and location market trends outlook [12] identifies the following six segments in the LBS market:

- **Social** - Friend finder
- **Productivity** - Fleet and cargo management
- **Information** - City guides, local service searches
- **Navigation** - Turn-by-turn navigation
- **Commerce** - Location-aware advertising
- **Security** - Family and pet trackers as well location-based authentication

While the segments differ in application domain they all rely on reliable and sufficiently accurate positioning technologies in the background. To the LBS providers the standardized positioning technologies provide assurance that the perceived quality-of-service is independent of the platform used and only subject to normal external conditions such as satellite geometry or radio propagation conditions. For example, the LBS application provider can be sure that, on average, the performance of an UE-integrated GPS is independent of the type of UE the end users are using. This is advantageous to the service provider, because typically the positioning is provided by the UE vendor, but the applications may originate from various third parties. However, because the customers do not make a distinction between the positioning technologies and applications, the application provider brand might suffer from unequal performance over geographical areas, mobile operators and UEs. Standardized solutions address these issues.

1.3 GNSS evolution drivers

The GPS ICD (Interface Control Document) [13] was released to the public in 1995. At that time a single signal called L1 C/A (see Chapter 2.1.2.1) was released for non-military use. Until recently this single GPS signal has been the primary positioning technology.

However, over the recent years the GNSS landscape has evolved rapidly and is in continuous change. The GPS modernization program has brought and will bring additional signals to the civilian users for better performance. The European Galileo [14] satellite system is in development and the Russians have updated, and will update, their GLONASS (Globalnaya Navigatsionnaya Sputnikovaya Sistema,

[15]). The Japanese are working on their regional GPS enhancement called QZSS (Quazi-Zenith Satellite System, [16]) and both Indians as well as Chinese have expressed their intent to develop their own satellite-based positioning capabilities. The respective schedules for these initiatives are summarized in Figure 1.1 and the characteristics of the different systems are discussed in Chapter 2.1.2.

This GNSS evolution has led to the need to modernize the assistance data standards. Previously the AGNSS (Assisted GNSS, see Chapter 2.2) standards only supported GPS L1 C/A signal and the corresponding data elements. However, the near-term availability of new GNSSs has led to new work items being opened in 3GPP working groups for extending GNSS support in the standardized positioning protocols. These standards and their evolution are discussed in Chapter 3.2.3.

1.4 RAN evolution drivers

In addition to the GNSS evolution, introductions of new cellular networks require work in positioning technologies. GERAN and UTRAN networks have long had positioning infrastructure, but the recent LTE work has also resulted in the need for positioning technology work in E-UTRAN. Although much of the elements defined previously can be re-used, they still need to be adapted to the new network architectures. The new E-UTRAN positioning infrastructure and protocol are discussed in Chapters 3.2.2.3 and 3.2.3.4, respectively.

1.5 Brief history of positioning standards

Location technology standardization started in the GSM domain in the late 90's in the T1P1 committee within the ATIS (Alliance for Telecommunications Industry Solutions, [17]) forum. The T1P1 is nowadays known as WTSC (Wireless Technologies and Systems Committee). Simultaneously with the T1P1 work LCS standardization was also commenced in the TR45 group under TIA (Telecommunications Industry Association, [18]) for AMPS (Advanced Mobile Phone Service), TDMA (Time Division Multiple Access) and CDMA networks.

The TIA standards were solely for the US due to it being the first country to mandate wireless E911. A few years later the work was also commenced in 3GPP RAN for UTRAN.

Regarding the GSM domain the T1P1 committee defined the whole LCS solution for the GSM Releases 98 and 99 including the requirements, architecture and the various required protocols. The work included the GERAN positioning protocol called RRLP (Radio Resource LCS Protocol). In 2000 the work was transferred from ATIS to 3GPP GERAN and the RRLP Release 98 and Release 99 were numbered as 3GPP TS (Technical Specification) 04.31 [19]. For Release 4 RRLP was re-numbered to the present 3GPP TS 44.031 [20].

The RRLP Release 98 defined the support for Assisted GPS for GPS L1 C/A signal and a GERAN-specific positioning method called E-OTD (Enhanced Observed Time Difference). This was the original content of RRLP that was virtually untouched for several years until in August 2004 a work item for GNSS was opened in 3GPP GERAN. The resulting RRLP Release 7 introduced the support for A-Galileo and multi-frequency measurements, but also a generic structure for easy inclusion of additional GNSSs later. It should be noted that reaching an agreement on the Release 7 contents took about one year, because of two competing proposals.

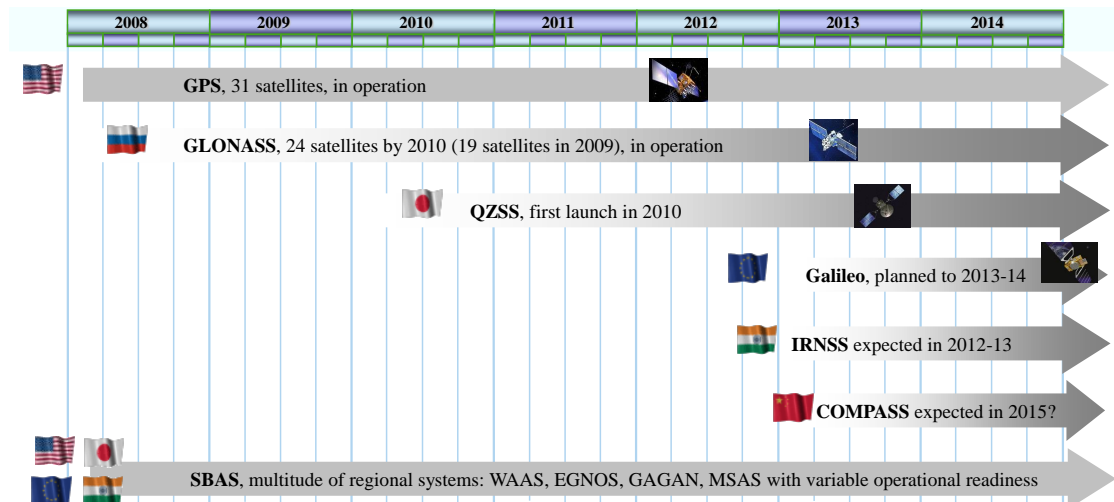


Figure 1.1: Future GNSS landscape. Image credit Jari Syrj arinne.

The two proposals that competed in 3GPP differed in how the new systems should be brought into the standards. One proposal was that the systems would be added one-by-one, while the other suggested a more generalized approach to enable as smooth and bit-efficient addition of known and future GNSSs as possible. The proposition also included the possibility for performance improvements, such as carrier-phase measurements, orbit extensions and the use of non-native navigation models. This generalized approach was chosen in autumn 2006 [21].

Although the Release 7 only included Galileo as GANSS (Galileo and Additional Navigation Satellite Systems) the generic format opened the path for adding other GNSSs in later releases. The RRLP Release 8 will include the possibility to provide UEs with assistance data for all the existing and some future GNSSs. The Release 8 adds the support for Russian GLONASS, QZSS, Modernized GPS [22] [23] [24] as well as various SBASs (Space-Based Augmentation System), such as WAAS (Wide-Area Augmentation System, [25]) and EGNOS (European Geostationary Navigation Overlay Service). Note that these developments initially concerned RRLP only, but the same solution was later approved into UTRAN [26] as well as SUPL draft Release 2 [27]. Currently 3GPP is working towards the E-UTRAN positioning solution, which will re-use major parts of RRLP.

1.6 Scope

As indicated, all the 3GPP-defined RANs have their own RAN-specific control plane positioning protocols. This work concentrates on positioning architectures, technologies and protocols in the 3GPP networks that cover GERAN, UTRAN and E-UTRAN radio access networks. Moreover, the scope includes the OMA-defined SUPL protocol that re-utilizes the 3GPP-defined positioning protocols.

On the other hand, other positioning architectures re-using and interacting with the 3GPP-/OMA-defined protocols and architectures are not in the scope. Such include WiMAX as well as IETF-defined (Internet Engineering Task Force, [28]) technologies including HELD (HTTP-Enabled Location Delivery, [29]). Moreover, the 3GPP I-WLAN (Interworking WLAN) is not considered.

Chapter 2

Overview of relevant positioning methods

This chapter will cover the positioning technologies that are in the scope of the present 3GPP-/OMA-defined positioning protocols. Moreover, emphasis is also given on the technologies that are expected to be in the scope of either 3GPP or OMA protocols in the near future. An additional restriction to the current chapter is that the positioning technologies are introduced from the UE perspective. This is to say that the methods not requiring UE interaction are not in the scope. An example of such technology is the AOA-method (Angle-Of-Arrival) that is currently being defined for E-UTRAN [30]. In the AOA the cellular network base stations measure the signal angle-of-arrival for triangulation and the UE is, hence, not involved in the positioning except for acting as a source of normal communication-related radio signals.

2.1 Satellite-based methods

2.1.1 Physical models involved

GNSS-based methods rely on the UE making an accurate TOA (Time-Of-Arrival) measurement on the time-stamped SV signals. Now, each GNSS has its own GST

(GNSS System Time, see Table 4.1) and assume first that both the UE and SV clocks are perfectly synchronized to the GST. In such a case recording the TOA of the signal, of which TOT (Time-Of-Transmission) is known, allows for determining the time-of-flight and the distance between the UE and the SV. The first modeling prerequisite becomes evident now. The UE must have a model of the SV orbit in order to estimate the SV position. Having these models and measuring the distance to at least three SVs allows for determining the three-dimensional UE position.

The second modeling prerequisite can be observed by noting that the SV clocks are not synchronized to the GST, but both the GST and SV clocks are based on freely running atomic clocks. Hence, the time difference between GST and each SV clock must be modeled in order to be able to relate the SV transmissions, which are based on the SV clocks, to the common reference time, GST. The modeling challenge becomes evident by considering the speed of light. Each 10-ns error in the model contributes to 3-meter error in range.

Regarding clocks it should also be noted that the UE clock is not synchronized to the GST. In fact, the time difference between the UE clock and GST is estimated as part of the PVT (Position-Velocity-Time) solution in the UE. Hence, the three-dimensional PVT in reality requires measurements from at least four SVs. Fixing the altitude reduces the requirement to three. In the end, therefore, the UE clock, GST and SV clocks are synchronized through the models. It can, therefore, be stated that GNSS-based positioning is a clock synchronization exercise.

The third modeling prerequisite deals with converting time to distance. In the case of free space the conversion is straightforward, because of the known and constant speed of light. However, the SV signals propagate through an atmosphere consisting of the ionosphere and the troposphere. The ionosphere is dispersive and delays ranging signal due to the presence of charged particles. The troposphere, on the other hand, is non-dispersive, but delays the ranging signal likewise. The atmospheric effects need to be modeled and compensated for.

The discussion above shows that GNSS-based positioning requires a variety of physical models (SV orbits, clocks, troposphere, ionosphere) that belong to the domains of mechanics, quantum mechanics, thermodynamics, plasma physics

and electromagnetics. The various effects and models utilized become evident in Chapter 2.1.3, in which the underlying (pseudo)range measurement is formalized. However, before that the currently available as well as future GNSSs are discussed.

2.1.2 GNSS descriptions

2.1.2.1 GPS

GPS consists of three components. The Control Segment is responsible for maintaining SVs and their orbits, keeping the data on board SVs up-to-date and valid as well as maintaining timing with respect to UTC(USNO) (Universal Time Coordinated maintained by US Naval Observatory) [31]. Secondly, the User Segment consists of the users of the GPS service. Thirdly, the Space Segment consists of SVs in orbit.

The nominal GPS constellation is 24 SVs positioned in six orbital planes having 60° separation in longitude. The orbit altitude is approximately 26600 km from the center of the Earth resulting in the orbital period of 11h 58min. Hence, each day the same SV geometry can be observed from a given location 4 minutes earlier than on the previous day. The orbit inclination is only 55° occasionally leading to compromised SV availability in high latitudes.

The legacy GPS includes only a single signal for the civilian use called L1 C/A (Coarse/Acquisition) [13]. In addition to L1 C/A the GPS SVs broadcast L1P(Y) and L2P(Y) encrypted military signals. Each SV broadcasts in the same central frequency with L1 and L2 having center frequencies at 1575.42 MHz and 1227.60 MHz, respectively. The SVs are discriminated using CDMA with PRN (Pseudorandom Number) Gold codes. Currently the bulk of civilian GPS UEs operate using L1 C/A only. However, dual-frequency receivers using various proprietary techniques to track the unknown P(Y) signal are widely used in the professional community.

The GPS modernization brings along three new signals for civilian use. These are referred to as L1C [24], L2C [22] and L5 [23]. The L1C and L2C have the same center frequencies as the current L1 and L2 signals. However, the L5 signal

will allocate a new band with a center frequency of 1176.45 MHz. The new signals bring along new signal modulations, new navigation data payload called CNAV (as opposed to NAV payload in legacy GPS) and increased resistibility against multipath and interference [31].

While the L1C signal will become available with GPS Block III SVs with the first launch scheduled for 2014 [32], the first satellite carrying experimental L5 payload was launched 24th March 2009 [33]. Moreover, the currently launched eight IIR-M block SVs broadcast the L2C ranging signal, but the CNAV payload has so far only been activated for single SVs for testing purposes [34] [35].

2.1.2.2 GLONASS

GLONASS is the GNSS maintained by the Russian Department of Defense and Russian Space Agency. The nominal GLONASS constellation consists of 24 SVs that are positioned in three orbital planes with eight SVs in each plane. In each plane there are seven active SVs and one spare. The orbit inclination is 64.8° and the orbit altitude is approximately 19100 km with respect to the surface of the Earth. Due to the higher inclination GLONASS has better coverage in high latitudes than GPS.

Whereas all the GPS SVs transmit at the same frequency and discrimination is based on PRN codes, GLONASS utilizes FDMA (Frequency Division Multiple Access). The carrier is modulated by a PRN code, but the same spread spectrum code is utilized by all the SVs and the discrimination is based on the frequency diversity. The GLONASS-M series SVs broadcast civil signals on both L1 and L2 [36] and the channel center frequencies are given by

$$\begin{aligned} f_{k1} &= (1602.0000 + f_k \cdot 0.5625)\text{MHz} \\ f_{k2} &= (1246.0000 + f_k \cdot 0.4375)\text{MHz}, \end{aligned} \tag{2.1}$$

where $f_k = \{-7, -6, \dots, 5, 6\}$ denotes the channel number allocated for the SV [15].

As of October 2009 the GLONASS constellation has 19 SVs (all GLONASS-M) and the Russians are committed to re-introducing the full constellation by the end

of 2010 [36]. Moreover, with the changes in the GLONASS program management the future constellation may be overpopulated to up to 30 SVs [37]. In 2010 the introduction of GLONASS-K series SVs will introduce the third open GLONASS signal in the L3 band (1197.648 - 1212.255 MHz) below the GPS L2. The L3 signal will be the first GLONASS signal to use CDMA. The CDMA signals for L1, L2 and L5 are under discussion for GLONASS modernization. [38] [39]

Russia has also expressed intentions to implement a GLONASS SDCM (System of Differential Correction and Monitoring) that is a Russian WAAS/EGNOS counterpart for the Russian territory. The two GEO (Geostationary Earth Orbit) SVs to support SDCM are expected to be launched in 2010 and 2011. [40]

2.1.2.3 Galileo

Galileo is the European GNSS designed to have a constellation of 30 SVs in Walker constellation having a 56° orbit inclination close to that of GPS. The SVs will be placed in three orbital planes with nine active SVs and one spare SV in each plane. The orbit altitude is 23 222 km from the surface of the Earth and the orbit repeats itself after 17 orbits, or 10 days.

Currently the Galileo constellation consists of only experimental SVs, GIOVE-A and GIOVE-B. Four In-Orbit-Validation SVs are expected to be launched by the mid-2011 and the initial constellation of 16 SVs is expected in 2013 [34] [41] [42].

Galileo will broadcast in three bands E1, E5 and E6. The E1 signal overlaps with GPS L1 band. The E5 band is sub-divided in E5a and E5b bands allocating $\pm 25 \cdot 1.023$ -MHz band having a center frequency at 1191.795 MHz. Both E1 and E5 signals provide Open Service to the general public. The E6 allocating $\pm 20 \cdot 1.023$ -MHz band around 1278.75 MHz, on the other hand, carries only restricted Commercial Service. The other services provided by Galileo are Public Regulated Service for authorities and Safety-Of-Life for safety-critical use cases including aviation.

The Interface Specification for Galileo is still in the draft phase [14] and is only available for R&D use. The lack of public non-license ICD for commercial purposes has delayed the development of Galileo receivers for mass market.

2.1.2.4 QZSS

QZSS (Quazi-Zenith Satellite System) is the Japanese GPS augmentation system operated by Jaxa (Japanese Aerospace Exploration Agency). QZSS will provide GPS augmentation in the area of Japan and Australia. The QZSS constellation will consist of three satellites in a highly elliptical orbit resulting in a figure-of-eight ground track with a lobe over Japan. The orbit design is such that the SVs spend maximum time over Japan so that at any time there is one high-elevation QZSS SV maximizing QZSS availability in urban environment. The first QZSS satellite is expected to be launched in 2010 [43].

QZSS SVs will broadcast L1C/A, L1C, L2C and L5 signals that are compatible with GPS and improve the availability of positioning services in the area of Japan. The QZSS SVs will also provide corrections to ranging enhancing the positioning performance. The L1-SAIF (Submeter-class Augmentation with Integrity Function) broadcast provides correction data to the ranging measurements including ionosphere and troposphere corrections. Finally, the LEX (L-band Experimental Signal) includes GPS orbit and clock model information for faster GPS SV acquisition. [16]

2.1.2.5 Space-Based Augmentation Systems

Space-Based Augmentation Systems provide ranging and correction data to augment GNSSs. SBAS SVs are in geostationary orbits meaning that the SVs are maintained in the orbit slots that are from $\mp 0.05^\circ$ to $\mp 0.1^\circ$ wide [44] making them practically stationary as seen from the Earth.

The various SBASs deployed include WAAS in North-America [25], European EGNOS, the Japanese MSAS (Multi-functional Satellite Augmentation System, [45]) and the Indian GAGAN (GPS-Aided GEO-Augmented Navigation, [46]).

The SBAS systems are interoperable [47] and have been allocated the PRN code space 120-138. The navigation and correction message format follows the WAAS specification [25]. The correction data provided to SBAS users is summarized in Table 2.1.

Table 2.1: Data broadcast by SBAS SVs.

Message type	Explanation
Fast corrections	Corrections to the pseudorange measurements, their rate of change and degradation
Long-term corrections	Corrections to the orbit and clock models broadcast by the GNSS SVs
RMS error	Estimated range error after the fast and long-term corrections
Integrity	Information on the health of the SV signal broadcasts
Ionosphere delay	Ionosphere delay correction in a grid
Clock-ephemeris covariance	Relative covariance matrix for clock and orbit errors
GEO navigation data	Orbit and clock models for the SBAS SV

2.1.2.6 Compass

Compass, or Beidou-2, is the Chinese Navigation Satellite System (CNSS). It currently lacks a public ICD and, hence, little is known about the Compass system. Compass will provide the users with two open and three restricted signals at three carrier frequencies. The open service signals are centered at 1575.42 (overlapping GPS L1 and Galileo E1) and 1191.795 MHz (overlapping Galileo E5b). [48] [49]

The Beidou-2 constellation will consist of 24 MEO (Medium-Earth Orbit), three GEO and three IGSO (Inclined Geosynchronous Orbit) SVs [50]. The Beidou-2 constellation currently includes two SVs - one in the GEO and the other in the MEO orbit [51]. However, before introducing a global system, China plans to set up a regional version with 12 SVs (five GEO, three IGSO, four MEO) by the end of 2012 [50]. Moreover, the Beidou-1 constellation includes four SVs in the GEO orbit [52].

CNSS uses CDMA spread spectrum technology similar to GPS and Galileo, but the details of the signal structure and PRN codes are unknown, although attempts

have been made to extract the information from the Beidou broadcasts [53]. Moreover, the broadcast data format and its contents are unknown. However, China has indicated that a draft ICD may be available in 2010 [50].

2.1.2.7 IRNSS

IRNSS is the to-be Indian Regional Navigation Satellite System with four GSO and three GEO SVs. IRNSS will include an open standard positioning service at L5 band at 1176.45 MHz and a restricted precision service in the same band. The signal structures are expected to be close to those of modernized GPS signals. However, the final specification is still due. India says that the entire constellation will be in place by 2012. [41] [54]

2.1.3 Point positioning

The most important GNSS receiver functions include signal acquisition, tracking, data demodulation and PVT calculation. Acquisition and tracking refer to finding the SV signal and maintaining a lock at least on the ranging code, respectively. On the other hand, data demodulation refers to decoding the necessary data elements from the SV broadcast and, finally, PVT calculation refers to estimating the UE location. While the details of the GNSS receiver functions are out-of-scope of this thesis and the technology choices differ from GNSS to GNSS, all the GNSSs provide in their broadcasts time, ranging signal and navigation data. Therefore, the following presentation on the satellite-based positioning methods can be applied to any GNSS.

Point positioning refers to utilizing GNSS pseudorange ρ and Doppler $\dot{\rho}$ measurements to solve the UE position, velocity and time. This section summarizes the mathematics behind calculating the PVT solution from the signal measurements. Assume that the pseudorange measurement ρ has appropriately been reconstructed from the code phase measurements [55] obtained from the GNSS chipset. Then ρ fulfills

$$\rho = \|\underline{x}_{SV} - \underline{x}^{UE}\| + c\delta t_{GST}^{UE} - c\delta t_{SV}^{GST} + I(s) + T(s) + \epsilon_{\rho}, \quad (2.2)$$

where \underline{x}_{SV} and \underline{x}^{UE} are the SV and UE positions, respectively. Terms δt_{GST}^{UE} and δt_{SV}^{GST} are the advances of the UE and SV clocks with respect to GST, respectively, and c is the speed of light in vacuum. Note that the SV position and clock advance are evaluated at the time of transmission. $I(s)$ and $T(s)$ are the ionosphere and troposphere delays, respectively, which are functions of the signal path s . Finally, ϵ_ρ is the term for measurement errors that include receiver noise, interference, multipath and receiver hardware offsets [31]. While there are methods for mitigating multipath and estimating hardware biases, they are not in the scope of the current work and are, hence, considered as noise.

Considering then the unknown terms in the measurement model, one needs to estimate \underline{x}^{UE} as well as the unknown GST t_{GST} through estimating the time offset δt_{GST}^{UE} between the user and GNSS system clocks. The SV position \underline{x}_{SV} is calculated from the orbit model included in the SV signal broadcast. Two widely used orbit parameterizations are Quasi-Keplerian parameterization and representing position-velocity-acceleration of the SV at given epoch for orbit integration.

The SV clock offset δt_{SV}^{GST} at t_{GST} is obtained from the SV clock model that the SVs also broadcast. The model is typically a polynomial. For example, in the case of GPS the adopted clock model is a second-order polynomial. Hence,

$$\delta t_{SV}^{GST}(t_{GST}) = a_0 + a_1(t_{GST} - t_{GST}^{ref}) + a_2(t_{GST} - t_{GST}^{ref})^2, \quad (2.3)$$

where t_{GST}^{ref} is the reference time of the model and a_0 , a_1 as well as a_2 the model coefficients. The total SV clock correction may also include terms for relativistic correction and group delay. For example, in the case of GPS the UE needs to estimate the relativistic effect due to the orbit eccentricity and correct the clock offset by an appropriate amount [13]. However, in the case of GLONASS the relativistic correction is included in the clock model by the ground segment. Moreover, the clock model may also need to be corrected for the group delay between the HW (hardware) circuits in the SV. For instance, in GPS the clock model refers to the L2 P(Y) broadcast. Hence, the L1 single frequency user needs to correct the clock offset by the group delay t_{gd} between the L1 and L2 circuits. The t_{gd} -term is broadcast by the SVs.

Atmospheric terms can be handled in various ways. Tropospheric delay T is in the order of few meters and can be modeled in the UE using, for example, the Saastamoinen model. The model requires information on the atmospheric conditions (pressure, temperature, partial water pressure), but the information can also be derived from the model for the standard atmosphere. In such a case only the orthometric height of the UE is needed. Troposphere models are considered in more detail in Chapter 4.2.4. [56]

Ionospheric delay I has a profound effect on the positioning performance. Its effect may be several tens of meters [56]. The delay may be calculated based on the model or it can be estimated in the UE. For example, GPS broadcast includes a global Klobuchar model [13] that accounts for approximately 50% of the variation in the ionospheric delay. Other models are considered in Chapter 4.2.3. In multi-frequency receivers ionospheric delay may be estimated by noting that the ionospheric medium is dispersive. Hence, the delay may be estimated or, for example, the measurements may be combined to form a linear combination free of ionosphere effects.

As an example consider GPS measurements at L1 and L2 frequencies. With simplifying assumptions it is obtained (see Chapter 4.2.3) that the ionospheric delays at L1 and L2 frequencies are related by

$$I_{L2} = \left(\frac{\lambda_{L2}}{\lambda_{L1}} \right)^2 I_{L1} = \gamma I_{L1}, \quad (2.4)$$

where λ_{L1} and λ_{L2} refer to the wavelengths at the corresponding frequencies. The term γ is called the dispersion factor. Therefore, the linear combination

$$\rho_{IF} = \frac{\rho_{L2} - \gamma \rho_{L1}}{1 - \gamma} = 2.542 \rho_{L1} - 1.546 \rho_{L2} \quad (2.5)$$

removes the ionospheric component and is called an IF (ionosphere-free) observable. The drawback of the combination is increased noise in the measurement. Assuming equal measurement variances for both ρ_{L1} and ρ_{L2} the resulting ρ_{IF} has a variance that is $(2.542^2 + 1.546^2) \approx 9$ times greater than the variance of individual measurements.

Finally, the UE velocity solution $\dot{\underline{x}}^{UE}$ is obtained by differentiating Equation 2.2 with respect to t_{GST} . It is then obtained for $\dot{\rho}$

$$\dot{\rho} = \frac{(\underline{x}_{SV} - \underline{x}^{UE}) \cdot (\dot{\underline{x}}_{SV} - \dot{\underline{x}}^{UE})}{\|\underline{x}_{SV} - \underline{x}^{UE}\|} + c\delta t_{GST}^{UE} - c\delta t_{SV}^{GST} + \epsilon_{\dot{\rho}}, \quad (2.6)$$

where the SV velocity $\dot{\underline{x}}_{SV}$ and the SV clock drift with respect to GST \dot{t}_{SV}^{GST} are obtained by differentiating the orbit and the SV clock models with respect to t_{GST} . The term δt_{GST}^{UE} is the drift of the UE clock with respect to GST and is solved simultaneously with the UE velocity.

Whereas the pseudorange measurement ρ is reconstructed from the code phase measurement [55], the rate component $\dot{\rho}$ can be obtained in two ways. The first method is the direct Doppler frequency measurement made by the GNSS receiver from the SV signal. The other method is to use the delta range measurement from the PLL (Phase-Locked Loop) tracking the carrier phase and counting full and partial cycles of the carrier signal over a known interval. This measurement is called the ADR (Accumulated Delta Range) measurement ϕ . The latter method is highly accurate down to millimeter-level in delta range in good signal conditions [31]. ADR ϕ needs to be divided by the measurement interval in order to arrive at $\dot{\rho}$.

There is also a subtle difference in the method of calculation depending on which of the two measurements is utilized. Whereas the Doppler measurement is instantaneous, ADR is measured over an interval. In the former case the SV position, SV velocity and SV clock drift are calculated for the measurement instant and the UE velocity solution represents that particular instant. In the latter case the SV position, SV velocity and SV clock drift are taken in the middle of the measurement interval. The UE velocity and clock drift solution then represent the average value over the interval.

It should be noted that all the required information for positioning (especially orbit, clock and ionosphere models) are available in the GNSS broadcasts. However, the information may likewise be distributed over the telecommunication networks to the Assisted GNSS UEs.

2.1.4 Precise Point Positioning

Whereas in the point positioning various measurement error sources as well as model residuals were bundled in a single term ϵ_ρ and modeled as noise, the principle of PPP (Precise Point Positioning) includes rigorously identifying and estimating these terms in the solution. Alternatively, appropriate corrections and models from various sources may be utilized. With this in mind, PPP is a potential direction for the future Assisted GNSS standards.

The measurement model associated with PPP may be expressed by

$$\begin{aligned} \rho &= \|\underline{x}_{SV} - \underline{x}^{UE}\| + c\delta t_{GST}^{UE} - c\delta t_{SV}^{GST} + I(s) + T(s) + \epsilon_\rho \\ \phi &= \|\underline{x}_{SV} - \underline{x}^{UE}\| + c\delta t_{GST}^{UE} - c\delta t_{SV}^{GST} - I(s) + T(s) + \lambda N + \epsilon_\phi \end{aligned}, \quad (2.7)$$

where $N \in \mathbb{N}$ is the unknown cycle bias (integer ambiguity). ADR measurement is similar to the pseudorange measurement apart from the unknown bias. The bias is due to the PLL starting accumulating cycles from an arbitrary count (typically zero) in the beginning. Note that this property does not affect determining $\dot{\rho}$ from the ADR measurement, because the difference of two ADR measurements is unaffected by the cycle bias. Although the ADR measurement is highly accurate, its drawback is that if the PLL on the signal is lost, cycle slips occur and the bias N changes. The measurement must then be corrected for the number of cycles slipped or the bias must be re-evaluated in the filter. An example of a typical filter used in solving a PPP solution is an EKF (Extended Kalman Filter) that allows for estimating the state of a dynamic system with non-linear differentiable state transition and observation equations. In the case of PPP the observation equation is not a linear mapping.

PPP can achieve decimeter-level accuracies and, therefore, the error sources and modeling needs have to be identified in detail. Especially, in PPP one considers [56] [57]

- Satellite orbit and clock model errors - broadcast satellite orbit and clock models have errors due to their predictive nature. The broadcast orbit models are typically accurate down to few meters RMS [58]. IGS (International

GNSS Service, [59]) publishes post-processed orbit and clock data that can be utilized in PPP. The best quality products achieve orbit and clock RMS accuracies of 2.5 cm and 75 ps, respectively. The 48-h orbit and clock predictions achieve RMS accuracies of 5 cm and 3 ns, respectively.

- Ionosphere delay - can be estimated in the process, if using a multi-frequency receiver. In post-processing it is also possible to use post-processed ionospheric maps available from, for instance, IGS. In the case of real-time PPP and a single-frequency receiver it is possible to use ionosphere models based on space weather forecasts as discussed in Chapter 4.2.3. The models also provide a reasonable initial value for the delay estimation accelerating the convergence of the position solution. Alternatively, due to code-carrier divergence (see Chapter 4.2.3) it is possible to form an ionosphere-free linear combination even from single-frequency measurements to mitigate ionosphere.
- Troposphere delay - can be estimated in the process, but in the post-processing one can also use post-processed troposphere delay maps available from, for instance, IGS. Similarly to ionosphere it is also possible to use troposphere delay forecasts as discussed in Chapter 4.2.4. The models based on forecasts also provide a reasonable initial value for the delay estimation accelerating the convergence of the position solution.
- Phase wind-up correction - GPS signals are right-hand circularly polarized and, hence, the relative rotation of the SV and UE antennas affects the phase measurement. The rotation can be compensated by noting that the SV solar panels point to the sun allowing for modeling the SV rotation.
- Satellite antenna offsets - orbit models are referenced to the SV center-of-mass. However, the measurements are referenced to the satellite antenna phase center, the broadcast point. Offset vectors have been defined for the SV generations.

- Receiver antenna offsets - high-quality UEs typically have antennas with highly isotropic antenna responses and well-defined antenna reference points. However, low-cost UEs may have significant anisotropy in the response as shown in [P3].
- Geophysical phenomena - although the UE were static its coordinates vary in time due to the tectonic plate motion, ocean loading (due to sea-level fluctuations due to tides), solid-Earth tides (time-dependent deformation of Earth due to Sun and Moon) and atmospheric loading. These can be modeled.

2.1.5 Relative methods

Whereas in PPP the UE obtains corrections or estimates the error terms, relative positioning builds on the idea of algebraically canceling the error-prone terms. Technique known as RTK (Real-Time Kinematic) achieves centimeter-level accuracy and is widely used commercially [P1] [P2]. Such a method assumes the existence of a reference station (a physical or a virtual computational reference point [P7]) and a mobile rover that is positioned with respect to the reference station. The theory presented next is shown for a single-baseline case (one rover and one reference station). However, the presentation can be generalized to the multiple-baseline case.

2.1.5.1 Measurement model

The cancellation of common mode errors can be seen by assuming two ADR measurements ϕ_k^p and ϕ_m^p made from the satellite p by the two receivers k and m . The measurement models taking UE and SV clock biases as well as atmospheric delays into account are

$$\begin{aligned} \phi_k^p &= \|\underline{x}^p - \underline{x}_k\| + c\delta t_k^{GST} - c\delta t_{GST}^p - I_k^p + T_k^p + \lambda N_k^p + \epsilon_k^p \\ \phi_m^p &= \|\underline{x}^p - \underline{x}_m\| + c\delta t_m^{GST} - c\delta t_{GST}^p - I_m^p + T_m^p + \lambda N_m^p + \epsilon_m^p \end{aligned} \quad (2.8)$$

where \underline{x}_k and \underline{x}_m are the UE k and m positions, respectively and the terms δt_k^{GST} and δt_m^{GST} refer to the UE clock biases. Vector \underline{x}^p and δt_{GST}^p are the SV p position

and clock bias, respectively. Furthermore, I_k^p , I_m^p , T_k^p and T_m^p refer to the ionospheric advance and tropospheric delay, when the signal propagates from the SV p to the UEs k and m .

Now subtracting the two measurements yields the single difference observable ϕ_{km}^p given by

$$\begin{aligned} \phi_{km}^p = & \|\underline{x}^p - \underline{x}_k\| - \|\underline{x}^p - \underline{x}_m\| + c(\delta t_k^{GST} - \delta t_m^{GST}) - (I_k^p - I_m^p) + (T_k^p - T_m^p) \\ & + \lambda(N_k^p - N_m^p) + \epsilon_{km}^p. \end{aligned} \quad (2.9)$$

In single differencing the SV clock bias with respect to the GST vanishes. Assume then that the receivers k and m both make measurements from the two satellites p and q allowing for forming two single differences ϕ_{km}^p and ϕ_{km}^q . Subtracting these two single differences yields the double difference observable ϕ_{km}^{pq} defined by

$$\begin{aligned} \phi_{km}^{pq} = & \|\underline{x}^p - \underline{x}_k\| - \|\underline{x}^p - \underline{x}_m\| - (\|\underline{x}^q - \underline{x}_k\| - \|\underline{x}^q - \underline{x}_m\|) \\ & - (I_k^p - I_m^p - (I_k^q - I_m^q)) + (T_k^p - T_m^p - (T_k^q - T_m^q)) \\ & + \lambda(N_k^p - N_m^p - (N_k^q - N_m^q)) + \epsilon_{km}^{pq}. \end{aligned} \quad (2.10)$$

In double differencing the UE clock biases with respect to the GST vanish. Hence one is left with the double difference of geometric ranges between the UEs and SVs, double differences of ionosphere advances and troposphere delays and finally with the double difference of integer ambiguities. Note that the integer nature of the ambiguity is preserved in the differencing.

The baseline vector \underline{b} between the two receivers is defined as $\underline{b} = \underline{x}_k - \underline{x}_m$ and, hence, the equation can further be simplified by writing

$$\begin{aligned} \phi_{km}^{pq} = & \|\underline{x}^p - (\underline{x}_m + \underline{b})\| - \|\underline{x}^p - \underline{x}_m\| - (\|\underline{x}^q - (\underline{x}_m + \underline{b})\| - \|\underline{x}^q - \underline{x}_m\|) \\ & - I_{km}^{pq} + T_{km}^{pq} + \lambda N_{km}^{pq} + \epsilon_{km}^{pq}, \end{aligned} \quad (2.11)$$

where double differences for ionosphere advance, troposphere delay and integer ambiguity are expressed in short notation. Note that since one assumes that the

location of the reference receiver \underline{x}_m is known, the double difference of geometric ranges can be written simply as

$$r_{km}^{pq}(\underline{b}) = \|\underline{x}^p - (\underline{x}_m + \underline{b})\| - \|\underline{x}^p - \underline{x}_m\| - (\|\underline{x}^q - (\underline{x}_m + \underline{b})\| - \|\underline{x}^q - \underline{x}_m\|) \quad (2.12)$$

and

$$\phi_{km}^{pq} = r_{km}^{pq}(\underline{b}) - I_{km}^{pq} + T_{km}^{pq} + \lambda N_{km}^{pq} + \epsilon_{km}^{pq}. \quad (2.13)$$

2.1.5.2 Observables

The basic observable of relative positioning is the double difference. Note that the double difference can be formed from both code phase and carrier phase (ADR) measurements. Assume then that there are double difference observables at two distinct frequencies having wavelengths of λ_1 and λ_2 :

$$\begin{aligned} \phi_{km}^{pq,\lambda_1} &= r_{km}^{pq}(\underline{b}) - I_{km}^{pq,\lambda_1} + T_{km}^{pq} + \lambda_1 N_{km}^{pq,\lambda_1} + \epsilon_{km}^{pq,\lambda_1} \\ \phi_{km}^{pq,\lambda_2} &= r_{km}^{pq}(\underline{b}) - I_{km}^{pq,\lambda_2} + T_{km}^{pq} + \lambda_2 N_{km}^{pq,\lambda_2} + \epsilon_{km}^{pq,\lambda_2}. \end{aligned} \quad (2.14)$$

Assuming now that λ_1 and λ_2 correspond to GPS L1 and L2 frequencies the observables presented are referred to as L1 and L2 observables. The observables may be combined to form a WL (wide-lane) observable by the linear combination

$$\begin{aligned} \phi_{km}^{pq,\lambda_{WL}} &= \frac{\lambda_2}{\lambda_2 - \lambda_1} \phi_{km}^{pq,\lambda_1} - \frac{\lambda_1}{\lambda_2 - \lambda_1} \phi_{km}^{pq,\lambda_2} \\ &= r_{km}^{pq}(\underline{b}) - \frac{\lambda_2 - \lambda_1 \left(\frac{\lambda_2}{\lambda_1}\right)^2}{\lambda_2 - \lambda_1} I_{km}^{pq,\lambda_1} + T_{km}^{pq} + \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} N_{km}^{pq,\lambda_{WL}} + \epsilon_{km}^{pq,\lambda_{WL}}. \end{aligned} \quad (2.15)$$

The resulting combination has a considerably longer effective wavelength ($\lambda_{WL} = \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} \approx 86.2 \text{ cm}$) than the individual observables. This makes solving the double difference ambiguity $N^{pq,\lambda_{WL}} = N^{pq,\lambda_1} - N^{pq,\lambda_2}$ easier. However, this comes with the cost of increased noise in the observable. Assuming the same variances for the

L1 and L2 observables the noise in the wide-lane observable is 5.7-fold. It should be noted that it is also possible to form an IF linear combination from the L1 and L2 observables by setting

$$\phi_{km}^{pq,IF} = 77\phi_{km}^{pq,\lambda_1} - 60\phi_{km}^{pq,\lambda_2}. \quad (2.16)$$

This observable has a wavelength of

$$\phi_{km}^{pq,IF} = \frac{\lambda_1}{77} \frac{\lambda_2^2}{\lambda_2^2 - \lambda_1^2} \approx 6.3 \text{ mm} \quad (2.17)$$

making the integer ambiguity resolution challenging even though the integer nature of the ambiguity is preserved in the combination.

Finally note that given that the baseline is short (in the order of few kilometers maximum) the flight paths of the signal from a given satellite to the two receivers can be assumed to be equivalent. In such a case $I_k^p = I_m^p$, $I_k^q = I_m^q$, $T_k^p = T_m^p$ and $T_k^q = T_m^q$ and the WL observable simplifies to

$$\phi_{km}^{pq,\lambda_{WL}} = r_{km}^{pq}(\underline{b}) + \lambda_{WL} N_{km}^{pq,\lambda_{WL}} + \epsilon_{km}^{pq,\lambda_{WL}}. \quad (2.18)$$

The measurements must also be nearly synchronous in order to guarantee the atmosphere coherence and, hence, the cancellation of atmosphere effects as well as the cancellation of SV clock offsets in double differencing. Moreover, if the PLL is not maintained cycle biases and, hence, double difference integer ambiguities change. Should this happen the positioning engine needs to identify the slip and potentially correct the measurements by so many cycles as the PLL slipped. For this purpose the measurement reports exchanged between the two receivers include an identification of the state of the PLL over the previous measurement interval. If the engine cannot compensate for the slip the ambiguities need to be re-evaluated.

2.1.5.3 Baseline solution

Solving the baseline is a special problem due to the integer nature of $N_{km}^{pq,\lambda_{WL}}$. Formally, one is seeking for a solution to the problem

$$\min_{\underline{b}, \underline{N}^{\lambda_{WL}}} \|\underline{\phi}^{\lambda_{WL}} - \underline{r}(\underline{b}) - \lambda_{WL} \underline{N}^{\lambda_{WL}}\|_{Q_{\underline{\phi}^{\lambda_{WL}}}^{-1}}, \quad (2.19)$$

where $Q_{\underline{\phi}^{\lambda_{WL}}}$ is the observable covariance matrix. Unknowns $\underline{b} \in \mathbb{R}^{3x1}$ and $\underline{N}^{\lambda_{WL}} \in \mathbb{Z}^{(n_{SV}-1)x1}$, when there are observations from n_{SV} satellites.

The widely-used solution to the problem is to first seek for the solution in the real-space ignoring the integer constraint. The problem needs to be linearized with respect to \underline{b} and $\underline{N}^{\lambda_{WL}}$ allowing then the use of standard techniques including EKF. Having then obtained the float solution the integer ambiguities can be fixed, for example, using the well-known LAMBDA-method (Least-Squares Ambiguity Decorrelation Adjustment) [60].

The LAMBDA-method seeks for the solution to the problem

$$\min_{\underline{\tilde{N}}^{\lambda_{WL}} \in \mathbb{Z}^{(n_{SV}-1)x1}} \|\underline{\tilde{N}}^{\lambda_{WL}} - \underline{\hat{N}}^{\lambda_{WL}}\|_{Q_{\underline{\tilde{N}}^{\lambda_{WL}}}^{-1}}, \quad (2.20)$$

where $\underline{\hat{N}}^{\lambda_{WL}}$ and $Q_{\underline{\hat{N}}^{\lambda_{WL}}}$ are the real-valued double difference integer ambiguity estimate and its covariance matrix, respectively. Unknown $\underline{\tilde{N}}^{\lambda_{WL}}$ is the integer-valued double difference ambiguity estimate.

The LAMBDA-method seeks for a transformation that preserves the integer nature of the ambiguities, but decorrelates the ambiguities as far as possible. The nearest neighbor search is then performed in the transformed space having an ideal geometric form of $(n_{SV} - 1)$ -sphere. It should be noted that the greater the wavelength of the observable the greater the distance between the grid points in the search space is. Hence, the wide-lane observable is preferred over L1/L2 observables. The solution candidates can then be transformed back to the original space. The best candidate (shortest distance from the real-valued solution in the $Q_{\underline{\hat{N}}^{\lambda_{WL}}}$ -metric) is selected and validated using statistical means.

Having fixed the ambiguities to their integer values the baseline \underline{b} can be estimated and filtered according to Equation 2.19 considering $\underline{N}^{\lambda_{WL}}$ now as a known vector. In a typical arrangement having fixed the wide-lane ambiguities to their integer values, the L1 and L2 ambiguities are solved next. Having done this the L1 double

difference observable is used for filtering the baseline \underline{b} due to the lower noise in the L1 observable than in the wide-lane observable.

2.2 Assisted GNSS

The previous chapters explained the significance of decoding the data payload in addition to the ranging payload from the GNSS SV broadcasts. The data payload contains essential items for PVT including orbit and clock models. However, in compromised signal conditions the capability to decode the data payload is often limited. AGNSS addresses this problem by providing an alternative route to carry the data payload to the UE. Moreover, AGNSS assistance helps the UE also in various other ways to achieve better user experience in terms of availability, speed, accuracy and integrity.

2.2.1 Case for assistance

Assisted GNSS is a concept, in which a GNSS receiver integrated in a UE with communication capability is aided in order to achieve higher availability, better accuracy, higher integrity and improved TTFF (Time-To-First-Fix). Aiding is based on relying both on the data transfer capabilities as well as on the inherent properties of the network including precise timing of radio transmissions.

Higher availability can be understood by considering the combination of the environment, in which the AGNSS-enabled devices are used, the attenuation of the SV broadcasts in such environment and the PLL/DLL (Delay-Locked Loop) tracking thresholds. PLL is needed to demodulate the navigation data from the SV broadcast (carrier phase tracking), whereas DLL is required to track the ranging signal (code phase tracking). As an example, consider the urban environment shown in Figure 2.1. Figure shows how the UEs in open environment or in light canopy are able to receive both the ranging data and the data payload. However, the UE receiving signal from the SV blocked by a building cannot demodulate the data payload. The data payload includes, for example, the SV orbit and clock models and, therefore, the UE cannot calculate the PVT solution. Note, however,

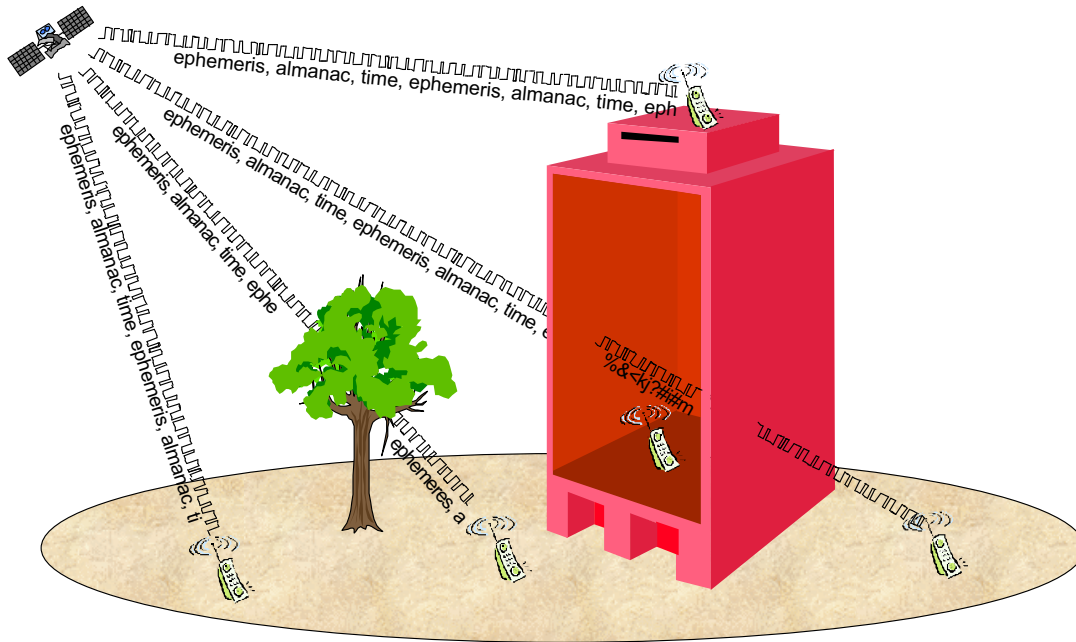


Figure 2.1: Loss of data payload, when the signal attenuates. Reprinted with permission from [55].

that the UE does receive the ranging code. Hence, if the UE obtained the data via an alternative channel the UE would be able to solve the PVT problem.

The loss of navigation data is due to the signal attenuation in blocked environment. Publication [61] shows that the SV signal can attenuate approximately 15 dB from its nominal level, before the error probability grows too high for the data demodulation to be possible. Such 15-dB attenuation levels are found in forest canopy and suburban areas as indicated in Table 2.2.

In urban canyons and indoors the attenuation is typically greater than 15 dB. Therefore, an unassisted UE is incapable of producing a PVT solution due to the UE lacking the data payload. However, in optimal conditions (stationary UE, data wipe-off) the DLL can track the ranging code even down to -160 dB ref 1 mW [62]. This corresponds to approximately 30-dB attenuation, which is the level of attenuation found in urban canyons and mild indoors as shown in Table 2.2. Therefore, in this 15-30 -dB attenuation range AGNSS can tangibly improve performance, because the required data is transferred to the UE via the

Table 2.2: GNSS signal attenuation, when signal passes through different materials and obstacles. Data from [31].

Material	Attenuation (dB)
Light to moderate forest	10
Moderate suburban	6.5
Steel canyons	15
In-building commercial	24
In-building high-rise	30

communication channel in the UE and the UE is, hence, able to provide PVT despite challenging signal conditions. This consideration shows the strength of Assisted GNSS.

2.2.2 AGNSS infrastructure

Figure 2.2 shows the high-level view of the AGNSS architecture. The core of the architecture is the AGNSS server that may lie in the cellular network or in public Internet. These servers provide AGNSS service to the subscribers in each geographical area - for instance in the cellular network the assistance server serves subscribers associated to a subset of base stations. In the cellular domain the assistance data is requested and delivered to the UE in the control messaging of the network (control plane). Alternatively, the assistance data may be delivered over the user plane (IP-network). In both circumstances, however, a point-to-point connection is established between the UE and the AGNSS server and assistance data is transferred from the server to the UE. Depending upon the situation the server may push assistance data to the UE or the UE may request for specific assistance data items from the server.

The AGNSS server may obtain its data from various sources. Sources may include WARNs (Wide Area Reference Network) that are networks of physical GNSS receivers distributed geographically (left hand side in Figure 2.2). These receivers provide integrity information as well as broadcast ephemerides to the AGNSS servers for distribution. On the other hand, orbit and clock models (as

well as other data) can originate from an external service providing, for instance, orbit and clock predictions (right hand side in Figure 2.2). Moreover, additional GNSS services may include meteorological institutes providing troposphere delay forecasts (see Chapter 4.2.4) to the AGNSS subscribers.

In addition to the navigation data the assistance services provide various other information to the UEs including reference position, reference time, a full copy of the data bit sequence broadcast by the SVs and differential GNSS corrections. These various aspects are now considered in turn.

2.2.3 Effect of obtaining navigation model

The navigation model refers to the absolute minimum amount of information on the SVs the UE needs to have in order to calculate the PVT fix. This includes the orbit and clock models for the SVs included in the PVT solution. For example, in the case of GPS receiving this information from the SVs takes in minimum 18 seconds, because the data is distributed over three sub-frames [13] each lasting six seconds. However, in an assisted case the information can be delivered quickly

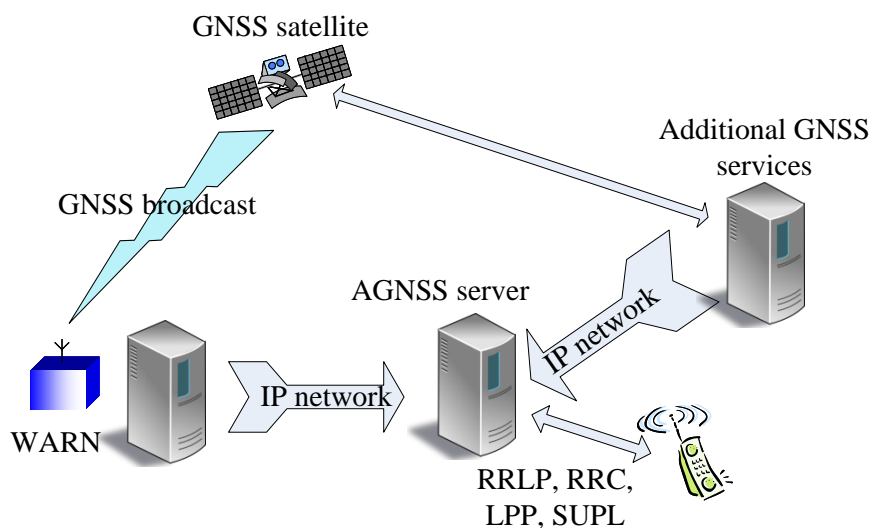


Figure 2.2: High-level AGNSS infrastructure. RRLP, RRC, LPP and SUPL refer to the assistance data standards discussed in Chapter 3.

to the UE over a data link, because the amount of data is fairly small (about 500 bits/SV [20]). This improves the user experience due to the reduced TTFF. Moreover, the UE may retrieve the assistance data from the server, whenever the data in the UE expires. In such a case the UE always has a valid copy of the navigation data and the TTFF is further reduced.

Finally, in case orbit and clock predictions are available, AGNSS-enabled UEs can be provided with navigation models extending days or even weeks ahead. In such a case the UE does not need to connect to the assistance server in the beginning of each positioning session. This improves user experience due to the time saved in not having to download the assistance [63] (user plane).

2.2.4 Effect of reference position, time and frequency

Having reference position, time and frequency has a profound effect on the GNSS receiver performance. Assume that the UE has a copy of the navigation data. This allows the UE to predict the satellite positions with respect to the UE as well as the SV clock offsets given that the GST and UE position are known at some accuracy (reference time and position). Ultimately, the UE can, thus, predict the visible SVs and SV signal code phases with some uncertainty. Likewise Doppler frequencies may also be predicted.

The SV signal acquisition is a three-dimensional search problem [55]. The dimensions are the SV itself, the code phase and the Doppler frequency of the SV carrier signal. In order to obtain an order-of-magnitude estimate of the significance of the search space reduction due to the AGNSS assistance, consider the following example. Assume a GPS receiver that searches SVs in half-chip bins in the code dimension and in 50-Hz bins in the frequency dimension. The number of required search bins to cover the whole search space (not taking into account the overlap of frequency bins) is

$$32 \text{ SV} \frac{\frac{2.6700 \text{ Hz}}{50 \text{ Hz/bin}} \frac{1023 \text{ chip}}{0.5 \text{ chip/bin}}}{\text{SV}} \approx 2 \cdot 10^7 \text{ search bins.} \quad (2.21)$$

The Doppler range of approximately ∓ 6700 Hz is obtained by considering the maximum radial speed (∓ 930 m/s or ∓ 4900 Hz in L1 frequency) of GPS SVs [P2] as well as assuming a 1-ppm (1575 Hz) accurate reference oscillator in the UE. Moreover, assuming car speeds the UE motion can contribute at maximum 200 Hz (137 km/h). Note that the Doppler frequency uncertainties due to the reference position and time are negligible here given that the reference data is reasonable. Reasonable means a few kilometers in the spatial domain and few seconds in the temporal domain. Moreover, the full GPS constellation is 32 SVs and the GPS L1 C/A CDMA ranging code sequence is $2^{10} - 1 = 1023$ chips long.

The first dimension is the actual SV that, for example, in the case of GPS means the PRN code broadcast by the SV. The coarse knowledge of time and position allows the UE to predict the SVs above the horizon and, hence, the search can be limited to a subset of SVs. Assume that the UE can now rule out two-thirds of the SVs that are not visible to it in any case. The search space is then

$$10 \text{ SV} \frac{\frac{2 \cdot 6700 \text{ Hz}}{50 \text{ Hz/bin}} \frac{1023 \text{ chip}}{0.5 \text{ chip/bin}}}{\text{SV}} \approx 5 \cdot 10^6 \text{ search bins.} \quad (2.22)$$

The second search dimension is the code phase of the SV CDMA broadcast. As noted, having the SV navigation model and knowing the coarse UE position and GST allows for predicting the code phase with some uncertainty. The UE can, therefore, limit the code phase search in a small volume within the search space. To obtain an order-of-magnitude estimate for the performance improvement assume that the uncertainty in the time assistance is $\mp 10 \mu s$ (fine time assistance in [64]). Furthermore, assume that the reference location horizontal uncertainty is 3 km. Note that this accurate assistance allows for removing the millisecond ambiguity also.

The GPS L1 C/A CDMA sequence is 1-ms long with 1023 chips. Hence, each chip is approximately $0.98 \mu s$ in temporal or 293 meters in the spatial domain. Therefore, $\mp 10 \mu s$ and 3 km correspond to ∓ 10 -chip and ∓ 10 -chip uncertainties, respectively. Note that the 10-chip uncertainty due to spatial uncertainty is the maximum. In reality the uncertainty maps as $\cos(\nu)$, where ν is the SV elevation from the horizon. Summing these results in $\mp \sqrt{10^2 + 10^2} \approx 14$ -chip uncertainty.

The search space, therefore, reduces to

$$10 \text{ SV} \frac{\frac{2.6700 \text{ Hz}}{50 \text{ Hz/bin}} \frac{2.14 \text{ chip}}{0.5 \text{ chip/bin}}}{\text{SV}} \approx 2 \cdot 10^5 \text{ search bins.} \quad (2.23)$$

Finally, the SV Doppler frequency can be predicted from the radial speed of the SV with respect to UE. Given the small uncertainties in reference time and position, the dominating component in acquiring the Doppler frequency is the performance of the UE oscillator. Assuming the UE has a typical 1-ppm oscillator the frequency uncertainty due to the oscillator is approximately ∓ 1575 Hz (L1 frequency). The maximum Doppler prediction errors due to uncertainties in position and time are 1 Hz/km and 0.8 Hz/s [62]. Moreover, the uncertainty due to the user motion contributes $\mp 146 \cdot \cos(\nu)$ Hz per 100 km/h assuming that the car moves horizontally. The Doppler search space, thus, reduces from ∓ 6700 Hz to, say, ∓ 1700 Hz (∓ 125 Hz due to the UE motion). The new search space volume is now

$$10 \text{ SV} \frac{\frac{2.1700 \text{ Hz}}{50 \text{ Hz/bin}} \frac{2.14 \text{ chip}}{0.5 \text{ chip/bin}}}{\text{SV}} \approx 4 \cdot 10^4 \text{ search bins.} \quad (2.24)$$

However, an AGNSS receiver integrated in a mobile UE may also obtain frequency assistance by locking to the cellular base station carrier frequency. In such circumstances the oscillator frequency uncertainty may drop to ∓ 200 ppb [62], which translates to ∓ 316 Hz in frequency at L1. Note that ∓ 200 ppb already allows for ∓ 100 ppb (158 Hz) uncertainty due to the relative motion between the UE and the base station. The new search space volume is (∓ 125 Hz due to the relative motion between the UE and the SV as previously) thus

$$10 \text{ SV} \frac{\frac{2.441 \text{ Hz}}{50 \text{ Hz/bin}} \frac{2.14 \text{ chip}}{0.5 \text{ chip/bin}}}{\text{SV}} \approx 10^4 \text{ search bins.} \quad (2.25)$$

The simple calculation shows that the search space is reduced by three orders-of-magnitude due to assistance. Moreover, more reduction can be made as time and frequency are decoded from the first acquired satellite. The modern GPS receiver can well have such number of parallel search bins that all the visible SVs can be acquired simultaneously across the whole (reduced) search space.

2.2.5 Effect of data bit assistance

The data bit assistance refers to obtaining a full copy of the bit sequence modulated within the GNSS broadcast. Obtaining the copy as well as assuming appropriate location and time assistance allows the UE to wipe the data payload away from the SV signal. Without data wipe-off coherent integration can only be performed over the period of one bit (20 ms in the case of GPS NAV broadcast). However, data wipe-off allows the UE to remove the $\pi/2$ phase shifts due to the data payload modulation in the SV signal and, hence, coherent integration can be continued over the bit boundaries. Theoretically each doubling of the coherent integration period increases the sensitivity by approximately 1.5 dBs, when keeping the total integration time (coherent and non-coherent) intact [62]. In practice the length of coherent integration is limited by the UE motion and UE frequency stability.

Data bit assistance is especially usable in the case of Galileo due to its signals having high data rates. For instance, in E1 the data rate is 250 symbols/s as compared to 50 Hz in GPS L1 C/A, L1C and L5. The 250 symbols/s data rate limits coherent integration time to 4 ms without the knowledge of data bits. However, with higher data rates also the amount of data to be transferred as assistance to the UE increases and, hence, in certain cases (low bandwidth) the feasibility of data bit assistance may be compromised. This is, however, compensated by the inclusion of dataless pilot signals in both modernized GPS signals (L1C, L2C and L5) as well as Galileo signals (E1, E5a/b and E6). These signals will contain both data and pilot channels multiplexed in the same carrier. The pilot channels do not have data payload and, hence, no data wipe-off is needed for extended coherent integration. In fact, in [65] it is found that with the precondition of position, time and frequency assistance these pilot channels provide significant sensitivity improvement (up to 9 dB) over the signals with data payload. However, the practical performance improvement will be somewhat less because of necessary compromises in the receiver design due to the cost, size and power consumption restrictions. These considerations are especially relevant for Galileo pilot signals that have additional long (temporally) secondary codes multiplexed on primary codes, which arrangement enlarges the acquisition search space significantly.

2.2.6 Effect of integrity information

Integrity refers to the certainty on the authenticity and correctness of the signal. The integrity may be comprised, for instance, if the positioning system is jammed by a hostile party (spoofing). In the civilian use the most likely reason for the compromised integrity is, however, a faulty signal source (SV).

The integrity of the signals is constantly monitored by the GNSS ground segment and the SVs broadcast their health data. However, the GPS ICD does not provide an indication of the time-to-alarm from the ground segment detecting a problem to the SV health information to be updated. In addition to the ground segment, the private and public GNSS networks monitor the SV broadcasts and provide RTI (Real-Time Integrity) information. RTI information can be carried within the assistance data payload to the UEs. The UE can then discard the measurements from the faulty SVs or signals.

2.2.7 Effect of correction data

The correction data refers to, for instance, the DGNSS (Differential GNSS) data that is specified in the assistance standards. The benefits of the DGNSS are well-known and can significantly improve the accuracy down to sub-meter level [31]. Other correction data available in the assistance standards are the ionosphere delay models including the Klobuchar model defined in [13].

2.3 Methods for Radio Access Networks

The various radio access networks provide a plethora of opportunities for the UE positioning. The following sections provide an overview of the methods relevant for positioning the UE based on the information the UE has, can measure or can obtain from the network.

2.3.1 Coverage area -based methods

By coverage area one refers to the maximum area, in which the UE can at least decode the identity of the wireless communication node. In the case of GERAN this can be either the locally unique combination of BSIC (Base Station Identity Code) and BCCH (Broadcast Control Channel) or the globally unique Global CID defined by the MCC (Mobile Country Code), MNC (Mobile Network Code), LAC (Local Area Code) and CID. Similarly, in UTRA-FDD the identification is either the locally unique combination of UARFCN-DL (UTRAN Absolute Radio Frequency Channel Downlink) and P-CPICH (Primary Control Pilot Channel) or globally unique MCC+MNC+CID. Alternatively, in IEEE 802.11 WLAN the identification is MAC (Media Access Control) address or EUI (Extended Unique Identifier) as defined by IEEE (Institute of Electrical and Electronics Engineers, [66]).

The simplest form of the coverage area -based method is to assign the UE the location of the wireless communication node. In the cellular standards this is referred to as the CID-method. This is necessary, if there is no further information on, for example, the sectorization, range, azimuth or beam width of the node antenna. In case this information is available the UE can be assigned a location, say, in the center of mass of the node. However, typically more than one communication node is heard and in such a case the UE can be positioned within the intersection of the node coverage areas. The size of the intersection determines the uncertainty of the location solution. Naturally, the true uncertainty also includes the uncertainty in the coverage area models that are challenging to estimate. In the cellular standards the method, in which supplementary information in addition to the serving cell is utilized in positioning, is called the ECID-method (Enhanced CID).

Figure 2.3 shows an example of the situation, in which the UE can hear three nodes. Note that the nodes may belong to different RANs including GERAN, UTRAN, E-UTRAN and the method only requires knowledge of the coverage areas. Although the cellular standards include elements to carry base station coordinates to the UE, no standard includes elements to carry coverage areas to the terminals. Hence, CID and ECID methods are network-based methods, in which the UE only assists the positioning process by reporting measurements, for instance the list of

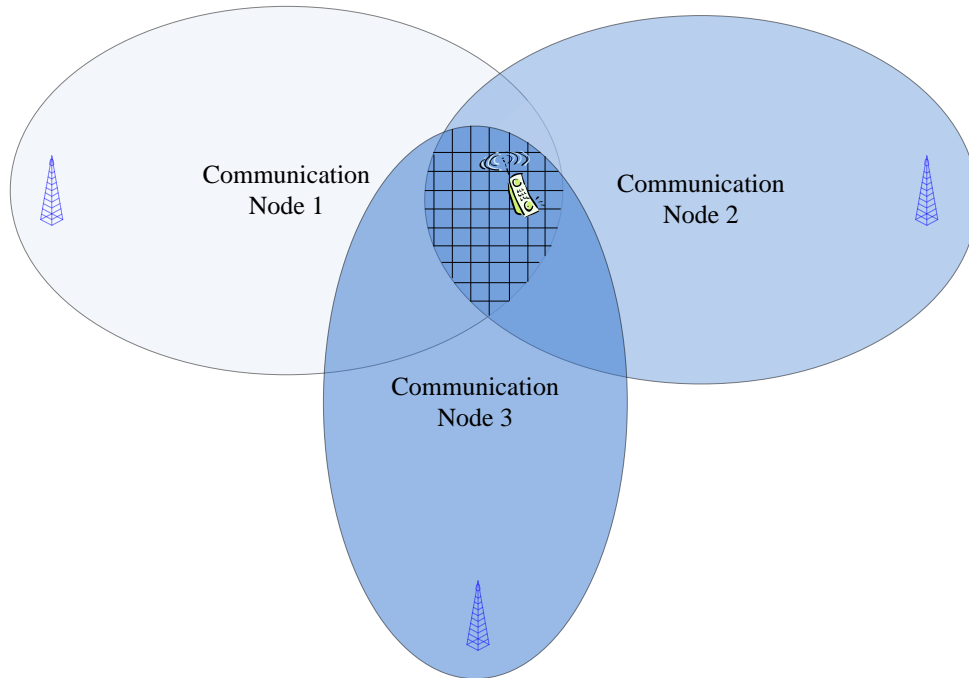


Figure 2.3: UE in the intersection of three wireless communication nodes.

heard BSs, to the network element. The network typically has information on the BS location, antenna direction, 3-dB beam width and the maximum range. However, these parameters model the coverage area with limited accuracy due to the real coverage being a function of, for instance, landscape. The true coverage area may be patchy and non-convex.

The performance of the coverage area -based methods is highly dependent on the air interface technology, network topology, frequency as well as environment (urban, sub-urban, rural). In addition, for instance in UTRAN an additional challenge is cell breathing, which refers to the cell area contracting and expanding as a function of traffic. Publication [67] cites 1400-meter 95% accuracy for suburban and significantly better 600-m 95% accuracy for urban areas using the CID-method. While the reference only assesses the experimental results from a single GSM network the results can be considered representable in a sense that the higher base station density (smaller cells) in urban areas leads to higher accuracy.

2.3.2 Ranging-based methods

Ranging-based methods refer to the techniques, in which the UE has access to the UE-node range information. The information may originate either from the time delay or path loss measurement. When ranges to a sufficient number of nodes can be determined the UE location can be solved through trilateration using standard techniques. Alternatively, the range estimate can be used in conjunction with the coverage area information.

In principle obtaining the time delay is straightforward. Assume that the node and the UE clocks are synchronized and the node transmits a time stamped transmission to the UE. The UE registers the TOA of the transmission and deduces the time delay. However, this assumes that all the nodes in the network as well as the UE clock are synchronized to the reference clock. In practice only the CDMA/CDMA2000 networks in the US are synchronized. Note that given the prerequisites that the node clocks are synchronized, transmissions are time stamped and the node locations are known, the UE clock bias with respect to the RAN reference clock can be estimated. This corresponds to the TOA-based positioning in GNSS.

GERAN is not synchronized, but in certain states the UE has an estimate of the time delay between the serving base station and the UE. This is due to GERAN being a TDMA-based network, in which each UE connects to the base station at the same frequency. Each UE is then assigned a time slot, in which the UE shall transmit. Hence, the UE needs to know the time-of-flight to the base station in order not to interfere with the transmissions from other UEs. This delay is called TA (Timing Advance) and its adjustment process is described in [68]. The time delay, however, is a fairly rough estimate of range with the resolution being one GSM bit. The bit length in GSM is $3.69 \mu s$ or approximately 550 meters in range due to the TA measurement being a round-trip measurement. Hence, the TA measurement can be thought as a doughnut around the communication node. The UE position is uniformly distributed within the doughnut-shaped area defined by the TA measurement. Using the TA measurement for positioning purposes has a further complicating property that the measurement error is positively biased due

to multipath and non-line-of-sight propagation.

The TA is measured only with respect to the serving base station, but can nonetheless assist in the UE positioning. Figure 2.4 shows a situation with the knowledge of the node position, coverage area and the TA doughnut. Without the knowledge on TA the UE position would be uniformly distributed within the coverage area. Now, however, the TA measurement restricts the UE position to the intersection of the coverage area and the TA doughnut. Again, the UE can be assigned a position in the center-of-mass of the intersection with the size of the intersection determining the uncertainty. Note that because the TA is known only with respect to serving BS, it is utilized in conjunction with the coverage area information to limit the location distribution and not as a range estimate in trilateration. Also, because the cell coverage areas are only known to the network, the TA-based methods are only feasible in UE-assisted cases, in which the UE assists in positioning by reporting measurements to the network element performing the calculation.

In [67] it is found that in GERAN the 95% accuracies for urban and suburban environments are 550 and 800 meters, respectively, when using TA in conjunction with the knowledge on the BS service area. Again, the smaller cell size in the areas with higher population density contributes to better positioning performance in urban areas. In UTRAN and E-UTRAN one can expect somewhat better performance, because in UTRA-FDD the resolution of RTT (round-trip time) between the UE and the serving node B (BS in the UTRAN vocabulary) is significantly better than that of TA in GERAN. In UTRA-FDD the RTT measurement resolution is half of the UTRAN chip [69]. The 3.84 Mchip/s rate [70], hence, corresponds to 40-m RTT accuracy. Finally, in E-UTRAN the TA resolution is $\frac{16}{15000 \cdot 2048}$ seconds corresponding to approximately 78-meters in range [71] [72].

The RTT measurement includes in addition to the geometric delay also the internal handling time, the time between reception and transmission, in the UE. This UE Rx-Tx time difference [73] is measured by the UE and reported to the network for subtraction from the RTT measurement. Moreover, in soft hand-over conditions in the Node B coverage area boundaries it is also possible to have RTT

measurements to multiple Node Bs. In such conditions it is potentially possible to trilaterate the UE using only RTT measurements.

The RSS-based (Received Signal Strength) positioning methods estimate the distance between the communication node and the UE based on the signal path loss. Transmission from an omni-directional antenna attenuates as $1/r^2$ in free homogeneous space due to power being constantly distributed over a larger sphere. However, in reality this simplified model is inadequate for practical needs as explained in [74], which gives for the RSS

$$RSS = P_0 - P_{Losses} + AG_{BS}(\hat{r}_{BS}^{UE}, A_{BS}) - PL + AG^{UE}(-\hat{r}_{BS}^{UE}, A^{UE}, O^{UE}) + \epsilon_{SF}, \quad (2.26)$$

where P_0 refers to the BS antenna transmit power, P_{Losses} to the cable losses in the BS system and AG_{BS} to the BS antenna gain towards the UE. The antenna gain is a function of the direction of the UE with respect to the base station \hat{r}_{BS}^{UE} as well as the antenna properties A_{BS} . The term PL refers to the BS-UE path loss and AG^{UE}

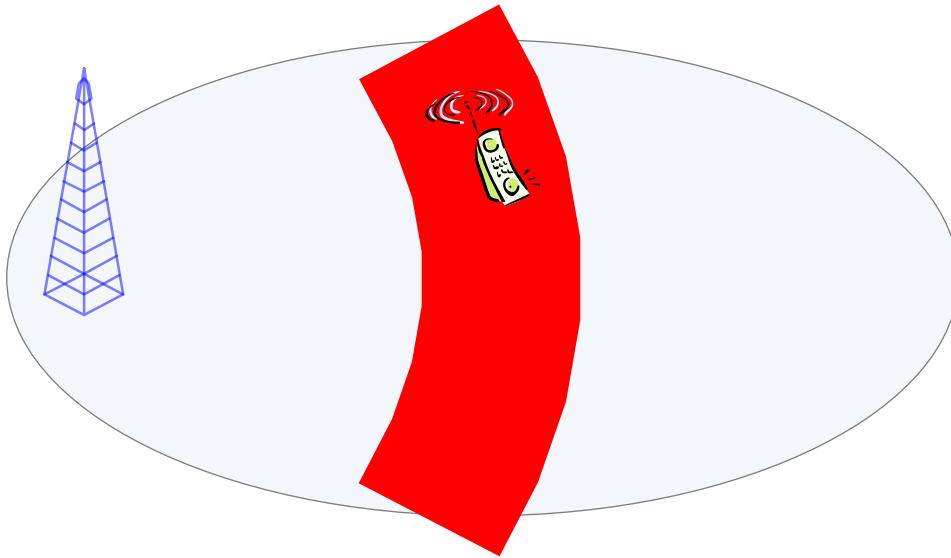


Figure 2.4: Timing Advance.

to the UE antenna gain that is a function of the BS direction $-\hat{r}_{BS}^{UE}$ as seen from the UE, the UE antenna characteristics A^{UE} and the UE orientation O^{UE} . Finally, ϵ_{SF} refers to the slow fading effects due to shadowing. Slow fading is modeled as a random process, for instance, as a log-normally, Rayleigh (no dominant line-of-sight component) or Rician (dominant line-of-sight) distributed variable. The amplitude of fading is typically 5-10 decibels [74].

Equation 2.26 can be paralleled with the measurement equation for GNSS pseudorange in Equation 2.2 that revealed a variety of modeling needs in order to be able to extract the UE position from the measurements. These modeling needs included, amongst other things, atmosphere as well as SV orbits and clocks. In contrast, Equation 2.26 reveals needs for electromagnetic modeling of antennas and radio propagation. Again, the modeling needs have to be adequately addressed before being able to extract position information from the measurements.

Utilizing the path loss component, which depends on the environment and UE-BS distance, in estimating the range is not trivial, because the loss exponent varies heavily depending on the propagation environment. While in free space the loss exponent is two, it can vary up to six in obstructed in-building conditions. One commonly used loss model is the Okumura-Hata model, which is based on experimental data. The loss model uses frequency, effective BS antenna height, UE antenna height and city type (large/medium/small) as inputs [74] [75].

In contrast to the TA/RTT measurements, which are available with respect to the serving BS only, the RSS measurements are typically available with respect to the neighboring BSs as well. This characteristics makes the RSS-based positioning attractive in the cellular networks, because in principle having a set of path loss estimates to a set of BSs allows for trilateration. However, even in such a case the coverage area information is typically used to limit the minimization domain as in [74]. Likewise, it is possible to utilize relative RSSs in positioning. Moreover, the set of RSS measurements is also suitable to be used in the fingerprint-based methods (see Chapter 4.2.5).

Note that the concept of ECID-method also covers utilizing TA/RTT and RSS measurements in order to improve the UE location estimate. In [67] the authors

utilize the RSS+TA-assisted ECID and find 95% accuracies of the method to be 950 and 430 meters in sub-urban and urban environments, respectively. Interestingly, in the experiment taking the RSS into account improves the performance in urban environment, but degrades accuracy somewhat in the sub-urban environment. This may simply follow from the selected path loss model or the model parameters.

The UE has limited information on the factors affecting the received signal strength. Firstly, typically the UE has no information on either the transmit power or the node antenna properties. Obviously, if a network element performs positioning based on the measurements made by the UE, the network can utilize the information it has on the node. This information may include transmit power, antenna azimuth, beam width and maximum range. Typically also the UE antenna pattern is unknown and even if knowledge on the pattern was available, the orientation of the antenna with respect to the base station would most likely be unknown. Hence, the UE antenna is typically considered as noise [74].

Secondly, the parameters of the loss model can only be known by the network. Obviously those parameters together with the node antenna properties could be carried to the UE over the air interface. However, this is superfluous, because the UE measurements can be carried to the network element easier than the model input information to the UE. Therefore, it can be deduced that the TA/RSS-assisted ECID is feasible only in the UE-assisted fashion.

2.3.3 Time Difference of Arrival

TDOA-based (Time Difference of Arrival) methods are utilized, when the base stations are not synchronized to a common time reference and the transmissions are not time stamped. In such a case the exact time of transmission is not known and methods alternative to TOA must be utilized. One possibility is to algebraically remove the UE-BS time offset from the measurement equations. This set of methods can be characterized as hyperbolic positioning.

Figure 2.5 illustrates the principle of hyperbolic positioning from the UE perspective. Assume that the UE at an unknown location \underline{x}^{UE} receives signals from three asynchronous BSs at the time t^{UE} . Let the signals from the BSs at known locations $\underline{x}_1, \underline{x}_2, \underline{x}_3$ be sent at times t_1, t_2 and t_3 according to the individual BS clocks. Now, each of the time delay estimates is a function of the geometric UE-BS range and the unknown time bias $\delta t_1^{UE}, \delta t_2^{UE}$ or δt_3^{UE} between the BS and UE clocks resulting in measurement equations

$$\begin{aligned} t^{UE} - t_1 &= \frac{1}{c} \|\underline{x}^{UE} - \underline{x}_1\| + \delta t_1^{UE} \\ t^{UE} - t_2 &= \frac{1}{c} \|\underline{x}^{UE} - \underline{x}_2\| + \delta t_2^{UE}. \\ t^{UE} - t_3 &= \frac{1}{c} \|\underline{x}^{UE} - \underline{x}_3\| + \delta t_3^{UE} \end{aligned} \quad (2.27)$$

By choosing the base station 1 as the reference base station and subtracting the first measurement equation in Equation 2.27 from the other two results in

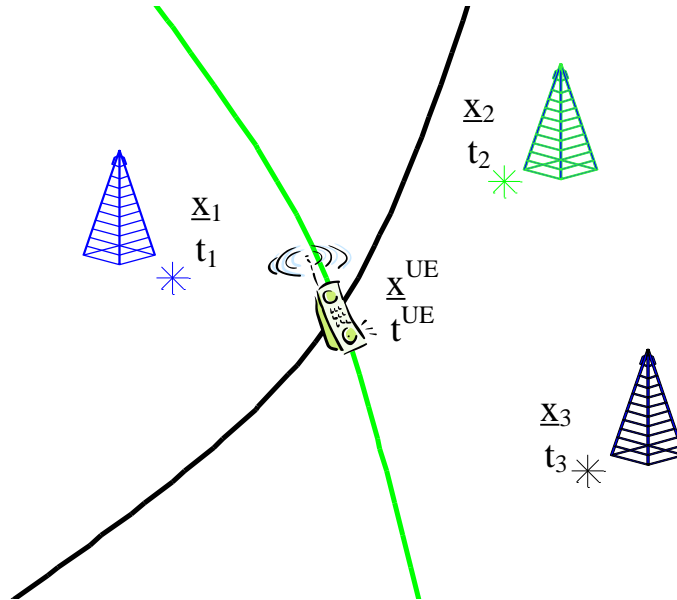


Figure 2.5: Principle of hyperbolic positioning.

$$\begin{aligned}
 (t^{UE} - t_2) - (t^{UE} - t_1) &= \frac{1}{c} (\|\underline{x}^{UE} - \underline{x}_2\| - \|\underline{x}^{UE} - \underline{x}_1\|) + \delta t_2^{UE} - \delta t_1^{UE} \\
 (t^{UE} - t_3) - (t^{UE} - t_1) &= \frac{1}{c} (\|\underline{x}^{UE} - \underline{x}_3\| - \|\underline{x}^{UE} - \underline{x}_1\|) + \delta t_3^{UE} - \delta t_1^{UE}
 \end{aligned} \tag{2.28}$$

Now, the terms

$$\begin{aligned}
 (t^{UE} - t_2) - (t^{UE} - t_1) &= \Delta_1^2 \\
 (t^{UE} - t_3) - (t^{UE} - t_1) &= \Delta_1^3
 \end{aligned} \tag{2.29}$$

are called OTDOA-observables (Observed Time Difference of Arrival) of base stations 2 and 3 with respect to the base station 1, respectively. Furthermore,

$$\begin{aligned}
 \delta t_2^{UE} - \delta t_1^{UE} &= \delta t_1^2 \\
 \delta t_3^{UE} - \delta t_1^{UE} &= \delta t_1^3
 \end{aligned} \tag{2.30}$$

are called RTDs (Real-Time Difference) of base stations 2 and 3 with respect to the base station 1.

The UE can readily measure the timing differences between the hearable base stations. Moreover, in the cellular networks LMUs (Location Measurement Units) measure RTDs between the base stations. Under such conditions the system of equations reduces to a system of one vector unknown \underline{x}^{UE} :

$$\begin{aligned}
 \Delta_1^2 &= \frac{1}{c} (\|\underline{x}^{UE} - \underline{x}_2\| - \|\underline{x}^{UE} - \underline{x}_1\|) + \delta t_1^2 \\
 \Delta_1^3 &= \frac{1}{c} (\|\underline{x}^{UE} - \underline{x}_3\| - \|\underline{x}^{UE} - \underline{x}_1\|) + \delta t_1^3
 \end{aligned} \tag{2.31}$$

The locus of each measurement equation in Equation 2.31 is a hyperbola as shown in Figure 2.5. Hence the name hyperbolic positioning. The UE can be positioned in the intersection of the loci using standard techniques for solving a system of non-linear equations. Two time difference observables are needed in 2D. In 3D three time difference observables are required and the locus of each measurement equation is a 3D-hyperboloid as illustrated in Figure 2.6. In Figure 2.6 the reference

base station is marked in magenta and the first base station as well as the resulting hyperboloid in blue. The figure in the middle shows the second base station and the resulting second hyperboloid in red in addition to the first hyperboloid. Finally, the lower figure shows the third base station and the resulting hyperboloid in green in addition to the first and second hyperboloids.

Note that equally well the measurements could be made by the base station from the burst sent by the UE. The base stations would then tag the received burst with their local clocks. Again assuming that the network has knowledge on the RTDs between the base station clocks the network can estimate the UE location in the manner presented above.

In GERAN hyperbolic positioning is known as E-OTD and in UTRAN the method is known as OTDOA-IPDL (OTDOA Idle Period Downlink). E-OTD and OTDOA-IPDL can both function in the UE-based (UE calculates the location) and UE-assisted (network element calculates the location using the measurements from the UE) modes. For E-UTRAN at least a UE-assisted TDOA-method will be defined. Finally, the uplink method, in which LMUs measure the UE signal, is known as U-TDOA (Uplink TDOA).

The TDOA performance is again highly dependent upon the environment. The simulations in [76] show that OTDOA-IPDL can achieve accuracy in the order of tens of meters in rural areas. However, in urban multipath conditions the 95% accuracies may degrade down to several hundreds meters.

Finally, the LMUs need not be physically in the network, but it is also possible to estimate RTDs using virtual LMUs. The concept of virtual LMUs is implemented, for example, in the product called Matrix [77] by Cambridge Positioning Systems (now part of Cambridge Silicon Radios). Matrix assumes the BS locations known, uses OTD measurements from a set of UEs as an input and solves the UE locations, RTDs and UE clock offsets. For example, in case there are two terminals it is required that they observe the same set of five BSs. Moreover, the concept also works in case of a single moving UE that observes the same set of BSs at distinct locations.

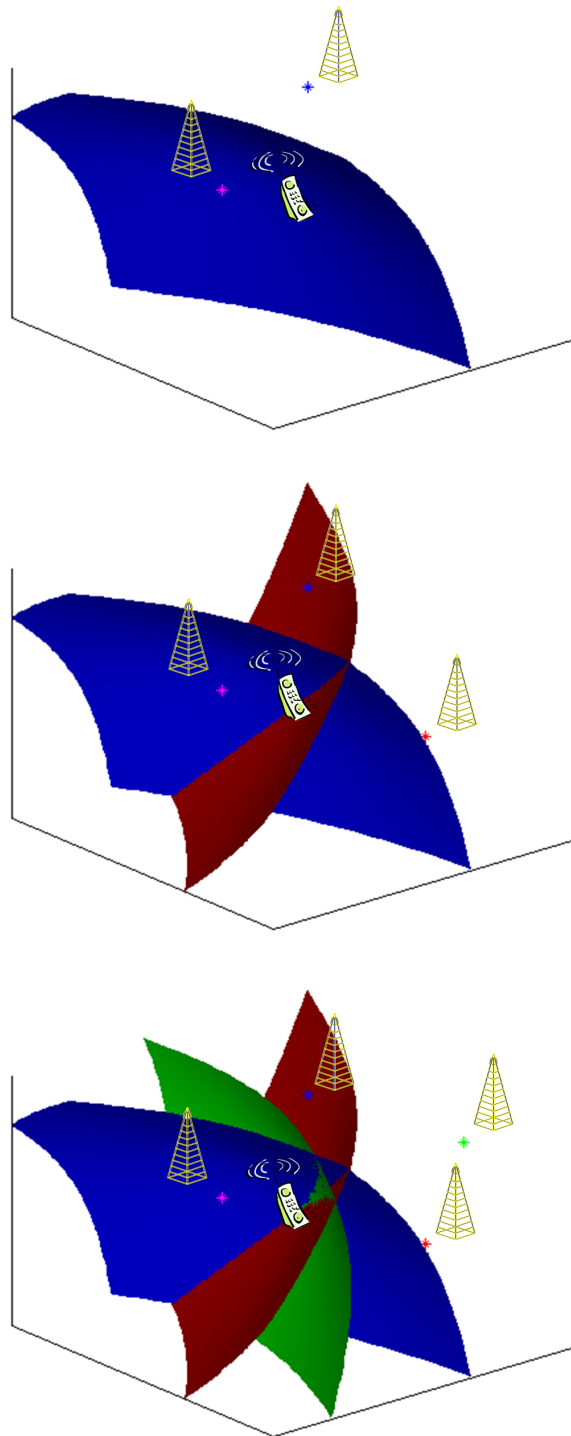


Figure 2.6: Hyperboloids in TDOA positioning.

Chapter 3

Standards

3.1 Third Generation Partnership Project

3GPP was established in 1998 to develop specifications for the third generation cellular networks. The scope includes both core network and radio access parts. Later on the scope was extended to cover the maintenance and development of GSM and its advanced radio access technologies including GPRS (General Packet Radio Service) and EDGE. The current scope also includes LTE, its radio access part E-UTRAN and the future LTE-A (LTE Advanced). [4]

This section provides insight to the positioning features, architectures and protocols within three 3GPP radio networks GERAN, UTRAN and E-UTRAN. GERAN specifications deal with the air interface and the MS (Mobile Station) part of the GSM/EDGE networks, whereas UTRAN is the air interface part of UMTS. UMTS supports a variety of UTRANs including UTRA-FDD and UTRA-TDD (Time-Division Duplex). UTRA-FDD is also known as WCDMA and the 1.28-Mchip/s UTRA-TDD as TD-SCDMA (Time-Division Synchronous CDMA). Finally, E-UTRAN refers to the radio access part of LTE. Although the emphasis is on the positioning protocols between the UE (or MS in the GERAN vocabulary) and the positioning entity in the network, an overview of the overall LCS architecture is also given for each RAN. Authors contribution to the positioning standards is limited to the air interfaces.

3.2 3GPP Location Services

The 3GPP work including location work is organized in three stages. The Stage 1 documents specify the service-level requirements for the agreed features and the Stage 2 defines the technical realization (architecture) based on the Stage 1 requirements. The Stage 1 work is preceded by a feasibility study and by an approval for the work item under which the work can be done. The feasibility study is sometimes called Stage 0. Finally, the Stage 3 documents define, for example, the actual signaling protocols. Although this work concentrates on the Stage 3 specifications, the Stage 1 and 2 aspects for the LCS in 3GPP networks are summarized for background.

3.2.1 Stage 1 description of LCS

The 3GPP TS 22.071 [78] specifies the Stage 1 requirements for the LCS in GERAN, UTRAN and E-UTRAN. From the perspective of the current work the most important aspects are the supported positioning technologies, the QOS (Quality-Of-Service) requirements as well as the flow and the location of the position calculation.

The supported positioning technologies include GNSSs (GPS, GLONASS, QZSS, Galileo), SBASs (WAAS, EGNOS, GAGAN, MSAS) as well as RAN-specific positioning methods including hyperbolic and CID/ECID methods.

The QOS requirements depend on the application utilizing the position information. Whereas local weather can be provided with position information having an accuracy in the order of kilometers (CID/ECID), navigation and turn-by-turn guidance requires 10-meter, or better, accuracy (AGNSS).

For the location of the position calculation the Stage 1 document recognizes two options. Positioning can either be UE-based or UE-assisted. In UE-based positioning the UE receives (pull or push) assistance data from the network unless the UE is to work in the standalone mode. Having received the assistance data the UE can position itself autonomously.

The other option is the UE-assisted mode, in which the UE reports the radio measurements to the network element, which performs the position determination. The UE resources need only be contributed to obtaining a snapshot of measurements and the burden of position calculation is on the network. In the UE-assisted case the network typically provides the UE with assistance that helps with obtaining the measurements. The bandwidth consumption of this reference assistance combined with returned measurements is significantly smaller than that of providing, say, ephemerides. Hence, the UE-assisted mode may be preferred over UE-based in bandwidth-constrained circumstances, such as control plane. Historically the UE-assisted mode has also been important due to the restrictions in the UE HW computational capabilities.

Yet another aspect of LCS is the origin of the LR (Location Request). Three types of LRs identified are MO-LR (Mobile-Originated LR), MT-LR (Mobile-Terminated LR) and NI-LR (Network-Induced LR). The MO-LR is simply an LR that the UE sets up to obtain its location and optionally velocity. The MO-LR can also be used by the UE to obtain assistance data and then continue positioning session for an extended duration without network interaction. Furthermore, the MO-LR can be used by the UE to fetch the assistance broadcast [79] ciphering keys.

MT-LR and NI-LR are both LRs triggered by an LCS client not residing in the UE. The distinction between the two is that the MT-LR and NI-LR originate from LCS clients external and internal, respectively, to the network. For example, if the UE is within GERAN and the LR originates from GERAN the LR is an NI-LR. A typical example of a NI-LR is emergency call positioning, in which the serving network initiates the positioning session as it detects an emergency call. An example of a MT-LR is a web-service for checking the family member locations. In such a case the LR originates from outside the network and the LR is, hence, a MT-LR. [8]

3.2.2 Stage 2 description of LCS

The 3GPP technical specifications 43.059 [80], 25.305 [81] and 36.305 [82] define the Stage 2 for positioning functionalities in the GERAN, UTRAN and E-UTRAN,

respectively. Moreover, the 3GPP location standards also cover the core network LCS specifications including 3GPP TS 23.271 [8] that covers the whole LCS system functional model.

3.2.2.1 GERAN

The Stage 2 GERAN LCS specification identifies four positioning methods for GERAN:

- AGNSS - see Chapter 2.2
- E-OTD - see Chapter 2.3.3
- RSS/TA-assisted CID/ECID - see Chapters 2.3.1 and 2.3.2
- U-TDOA - see Chapter 2.3.3

The first two methods can be performed in both MS-based and MS-assisted fashion while the CID/ECID is MS-assisted only. Also, for the first two methods the MS needs to obtain assistance data from the network (MS-based) or return signal measurements to the network (MS-assisted). For ECID the MS may need to provide the network with NMR (Network Measurement Result) that includes the RSS measurements for the serving and neighboring cells. Note that NMR is implicitly reported by the MS to the network for cell re-selection and handover purposes and, hence, no additional network load is induced by using NMR for positioning purposes. The performance of AGNSS and E-OTD may be improved by utilizing ECID (TA-assisted, [80]) to obtain a more accurate reference location. U-TDOA is a network-based method for which the MS may be instructed to briefly transmit at maximum power for maximum hearability.

Figure 3.1 shows the LCS architecture in GERAN. The location requests from the core network are received by a BSC (Base Station Controller) that is a part of a BSS (Base Station Subsystem). The requests are directed to an SMLC (Serving Mobile Location Center) that handles the requests. An SMLC can be integrated in a BSC or is a separate entity. An SMLC has the responsibility for coordinating

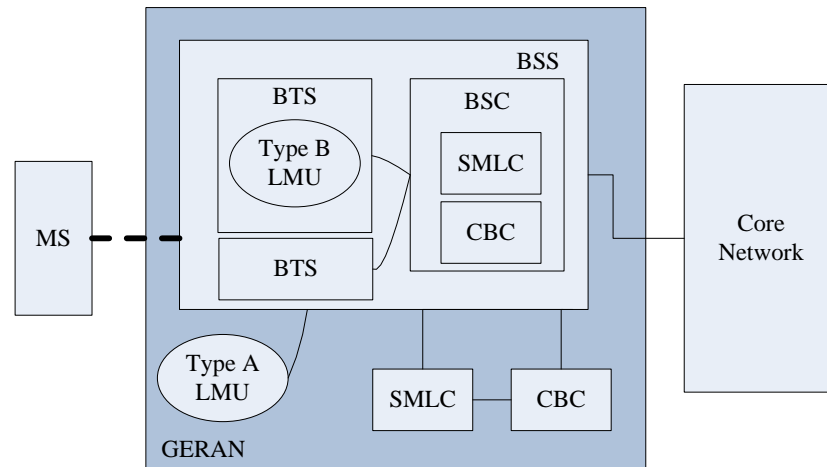


Figure 3.1: GERAN LCS architecture.

the overall resources needed for locating the MS. An SMLC provides the MS with assistance data and calculates the MS position estimate in the MS-assisted case.

An SMLC may control one or more LMUs. An LMU can either be integrated (type B LMU) into a BTS (Base Transceiver Station) or be a separate GERAN entity (type A LMU). An LMU has a responsibility for performing signal measurements that support positioning. The resulting assistance data is common to all the MSs in a certain geographical area. As an example, an LMU performs radio interface timing measurements resulting in RTD and/or ATD (Absolute Time Difference) information. RTD information includes the relative BS timings and can be utilized in E-OTD as explained in Chapter 2.3.3. ATD information, on the other hand, includes the BS timings with respect to some reference time scale, such as GPS time. Such information can be utilized as AGNSS fine time assistance as explained in Chapter 2.2.4. Finally, an SMLC may interact with a CBC (Cell Broadcast Center) to broadcast assistance data. In the assistance data broadcast a CBC transmits, amongst other information, GPS and E-OTD assistance. The broadcast of assistance data from a CBC is discussed in Chapter 3.2.3.2.

The architecture implies a protocol stack shown in Figure 3.2 with the SMLC-BSC and BSC-MS signaling as shown. The Stage 3 SMLC-MS interface is defined in 3GPP TS 44.031 [20] that is discussed in Chapter 3.2.3.1. Moreover, BSSAP-LE

(BSS Application Part LCS Extension, [83]) defines the information elements used by an MS to request assistance data from an SMLC. These request elements are carried within the SS (Supplementary Services) protocol between the MS and the network [84].

3.2.2.2 UTRAN

The UTRAN Stage 2 LCS specification defines the the following positioning methods for UTRAN:

- AGNSS - see Chapter 2.2
- OTDOA-IPDL - see Chapter 2.3.3
- RTT/RSS-assisted CID/ECID - see Chapters 2.3.1 and 2.3.2
- U-TDOA - see Chapter 2.3.3

As in GERAN the first two methods are suitable for both UE-based and UE-assisted use and the third is only for UE-assisted. U-TDOA, on the other hand, is a network-based method without interaction with the UE except for the possible instruction for the UE to transmit at full power for a short duration.

The OTDOA-IPDL utilizes the UE time difference measurements between the Node Bs as well as RTD/ATD information provided by the LMU on the Node B timings. The IPDL-part (Idle Period DownLink) refers to the time slots during which the serving Node B does not transmit. In the UTRA-FDD the Node Bs transmit at the same frequency resulting in the hearability problem for the UE, because the serving Node B transmission masks the neighboring Node Bs, when

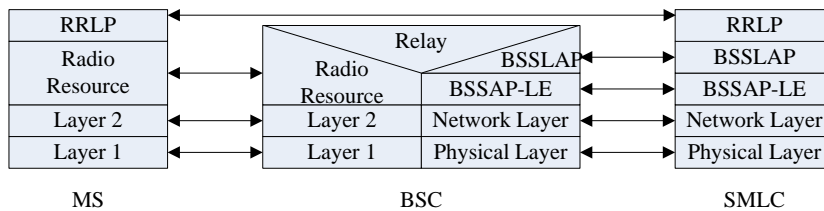


Figure 3.2: GERAN LCS protocol stack.

the UE is close to a Node B. Hence, the network may instruct the UE to perform the TDOA measurements during a given idle period so that the UE can also hear the neighboring Node Bs.

The hearability problems lead to the position determination yields being as low as 50-70% in urban environment in UMTS networks [76]. With IPDL the success rates improve to approximately 90%. Moreover, multipath limits the performance and OTDOA has not been found to fulfill the FCC E911 requirements. Another point for criticism in OTDOA-based methods is the need for the deployment of LMUs in the network increasing the infrastructure costs. Due to the performance and cost issues the OTDOA has not been commercially deployed in UMTS. Likewise the commercial uptake of E-OTD has been very limited. However, it should be noted that the concept of virtual LMUs introduced in Chapter 2.3.3 may be used to relax the infrastructure requirements for OTDOA. [85]

Figure 3.3 shows the UTRAN LCS architecture. In high level an S-RNC (Serving Radio Network Controller) receives the LR from the core network or intra-RAN. RNCs in general manage the UTRAN resources including Node Bs, LMUs and SASs (Standalone SMLC). Positioning works either in RNC-centric or SAS-centric mode the difference being that in the SAS-centric approach the SAS has a control over the positioning procedures. Note that whereas in the GERAN LCS architecture there is always an SMLC, in the UTRAN the S-RNC can assume the role of the SMLC including the control of positioning method, position calculation and providing assistance data to the UE. Even with an SAS in the network the S-RNC still has the responsibility for the overall location resource management.

Another role of RNC is to function as a C-RNC (Controlling RNC), for example, to control the uplink and downlink signal power levels for UE positioning, IPDL mechanism for TDOA and to control the broadcast of assistance information similarly to the GERAN broadcast. Moreover, the C-RNCs provide the S-RNC/SAS with positioning-related information. For example, a C-RNC controlling the LMU that receives the UE signal in U-TDOA will report the measurement to the S-RNC/SAS.

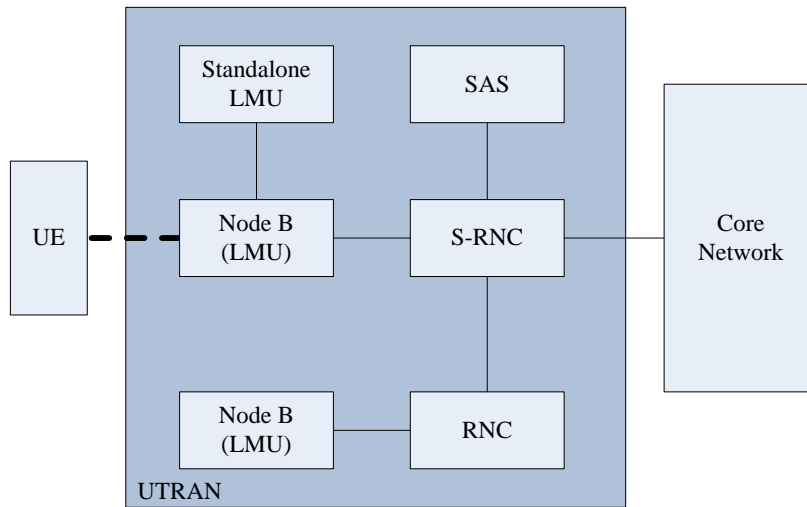


Figure 3.3: UTRAN LCS architecture.

The Stage 3 of UTRAN LCS is defined in RRC (Radio Resource Control) protocol [86] with termination points in the UE and RNC. Therefore, in the SAS-centric mode the SAS will provide the RNC the assistance data to be forwarded to the UE. Also, in the SAS-centric mode the S-RNC simply forwards the location requests to the SAS and provides the SAS with information needed in positioning such as the Node B information guide. In addition, the SAS, if it exists, is responsible for performing U-TDOA positioning on request by the S-RNC. Furthermore, the SAS will handle the simultaneous UE positioning requests.

The role of the LMU in UTRAN is similar to the GERAN LMU with the addition that a UTRAN LMU is also defined to measure inter-RAN timings. This refers to an LMU making timing measurements between Node Bs and BSs of external radio systems such as GERAN. A UTRAN LMU may exist as a standalone or as an associated version. An associated LMU is associated with a Node B and potentially utilizes its radio apparatus. The standalone LMU, on the other hand, as the name implies is not associated with a Node B, although a standalone LMU uses similar signaling as a Node B to communicate with its C-RNC.

3.2.2.3 E-UTRAN

3GPP TS 36.305 [82] specifies the following positioning methods for E-UTRAN:

- AGNSS - see Chapter 2.2
- UE-assisted TDOA - see Chapter 2.3.3
- TA/RSS/AOA-assisted CID/ECID - see Chapters 2.3.1 and 2.3.2

The three first methods are quite similar to the GERAN/UTRAN methods with a few exceptions. Firstly, the TDOA will only be available in the UE-assisted mode. Secondly, the AOA-assisted ECID is practically a new positioning method to the standardized cellular systems. Previously the AOA measurement has only been defined for 1.28 Mcchip/s UTRA-TDD [87]. AOA measurements can only be performed by eNode Bs (Node B in E-UTRAN vocabulary) [82] and currently it is not planned that AOA measurements would be routed to the UE.

In the 3GPP RAN2 working group discussions the TDOA-based methods for E-UTRAN have become under criticism for the same reasons as OTDOA-IPDL in UTRAN (see Chapter 3.2.2.2). Therefore some companies have proposed considering improved network-based technologies for LTE in addition to the OTDOA-based methods. Such technologies might include fingerprinting considered in Chapter 4.2.5. [85]

Figure 3.4 shows the E-UTRAN LCS architecture with the UE, eNode B, MME (Mobility Management Entity) and E-SMLC (Evolved SMLC). An MME is an entity in the E-UTRAN network taking care of, for instance, mobility between E-UTRAN and other 3GPP access networks and roaming. An MME also receives the location requests from the core network and forwards them to an E-SMLC. Alternatively, an MME may itself commence positioning request due to, say, emergency call. However, also in this case the location determination responsibility is in the associated E-SMLC.

An E-SMLC has the responsibility for the UE positioning as well as for providing the UE with assistance data. An E-SMLC therefore interacts with the UE and with the serving eNode B in order to collect the required radio measurement data

for positioning. Obviously also LMUs are required in the infrastructure given that LTE is an asynchronous network. However, the Release 9 of 3GPP TS 36.305 does not standardize the LMU functionality, but leaves it as an implementation aspect.

The protocol stacks implied by the architecture are described in Figure 3.5 and Figure 3.6. Figure 3.5 shows LPP (LTE Positioning Protocol) being a protocol between a UE and E-SMLC. LPP is transparent to the MME and eNode B that only relay the LPP signaling in the various lower level protocols. The lower level protocols include IP, SCTP (Stream Control Transmission Protocol), PDCP (Packet Data Convergence Protocol), LCS-AP (LCS Application Protocol), RRC and the S1-AP (S1 Application Protocol). LPP carries positioning instructions, assistance data and positioning capabilities information.

Figure 3.6, on the other hand, shows the LPPa (LPP Annex) protocol between an E-SMLC and eNode B. Again, LPPa is transparent to the MME and carries measurement instructions and measurement results between an E-SMLC and eNode B. Such measurements may include, for example, AOA measurements.

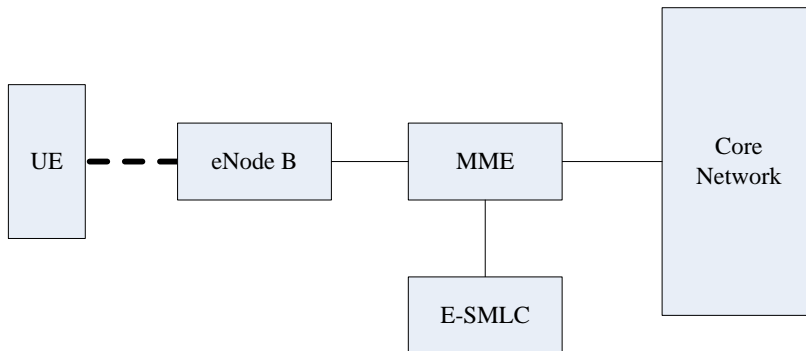


Figure 3.4: E-UTRAN LCS architecture.

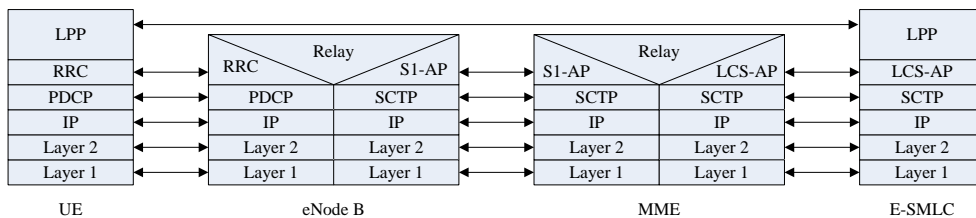


Figure 3.5: E-UTRAN LCS protocol stack: LPP.

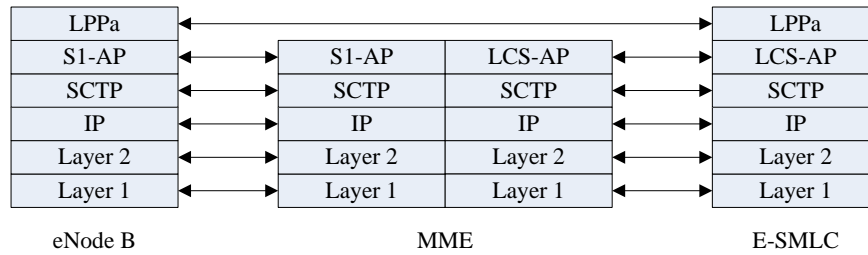


Figure 3.6: E-UTRAN LCS protocol stack: LPPa.

3.2.2.4 GERAN - UTRAN - E-UTRAN interworking

An external LCS client does not have information on to which network a UE is connected. Hence, in the LCS architecture there is a GMLC (Gateway Mobile Location Center), which is the first node an external LCS client accesses in the network. A GMLC contains the functionality to support LCS and within the network operated by an operator there may be more than one GMLC. The Requesting GMLC, to which the LCS client connects, routes the request to the GMLC associated with the serving node of the UE to be positioned (Visited GMLC). The UE may reside in GERAN, UTRAN or E-UTRAN access network. A GMLC thus provides seamless LCS across the various 3GPP access types. A GMLC is also responsible of the privacy checks associated with location requests. [8]

3.2.3 Stage 3 description of LCS

3.2.3.1 GERAN RRLP

3GPP TS 44.031 [20] defines RRLP for GERAN. RRLP is a highly important positioning protocol due to it being the lead for the UTRAN positioning protocol [86]. Major parts of RRLP are also re-used in LPP [88] [89]. Moreover, in addition to its use in the control plane RRLP is also widely used in the user plane as a sub-protocol to SUPL [6] [7].

RRLP provides the functionality required to request and return position or measurements, to provide the MS the necessary assistance data, to handle the protocol errors as well as to exchange positioning technology capabilities between

the MS and SMLC. The specific components in RRLP are

- Measure Position Request
- Measure Position Response
- Assistance Data
- Assistance Data Acknowledgment
- Protocol Error
- Positioning Capability Request
- Positioning Capability Response

Note that the request of assistance data from the MS to the SMLC is not defined in RRLP, but in 3GPP TS 49.031 [83] and is carried in the SS protocol [84] between the MS and the network. Moreover, the MS cannot request additional assistance data in the same RRLP session. However, the MS may send the SMLC a protocol error after the assistance data delivery to indicate a problem in the delivery or in the quality of assistance data. The SMLC may then decide to re-send the assistance data to the MS or to terminate the RRLP session. SMLC may modify the assistance data to be sent, but the MS cannot explicitly ask for certain assistance data in this case. Finally, note that RRLP does not support parallel positioning requests or transactions, but the current RRLP session is terminated by the MS, if a new positioning session is started by an SMLC.

Positioning methods and method types in RRLP

The positioning methods supported by RRLP include E-OTD and (A)GNSS, their hybrid or free choice. Hybrid refers to either using E-OTD and GNSS measurements in the same solver or using one method as a fallback to the other. Note that in GERAN the SMLC controls the positioning session and instructs the MS to use a specific positioning method.

The method types supported are MS-assisted (position estimation in the SMLC), MS-based (position estimation in the MS), MS-based preferred, but MS-assisted

allowed and MS-assisted preferred, but MS-based allowed. Again, the SMLC controls the method types.

Assistance data delivery in RRLP

RRLP allows for transferring assistance data for E-OTD, AGPS and AGANSS (Assisted GANSS). The specific GNSSs supported are GPS, Modernized GPS (L1C, L2C, L5 signals), GLONASS, QZSS, SBASs (WAAS, EGNOS, MSAS, GAGAN) and Galileo. Note that the upcoming GLONASS CDMA is not supported due to the lack of ICD.

The RRLP assistance data structure is shown in Figure 3.7. The assistance data can be divided into three parts. The first part is the assistance for E-OTD. The E-OTD assistance contains a list of BTSs the SMLC expects the MS to hear and their predicted OTDs at the MS location. This information assists the MS to concentrate measurement resources on specific channels (frequencies). In the MS-assisted case the MS returns the measured OTDs to the SMLC. In the MS-based case, however, the MS receives also the accurate RTDs for the BTSs as well as the BTS locations. This information is sufficient for the MS to deduce its position given that enough measurements were obtained and the geometry associated with the BTSs is appropriate as shown in Chapter 2.3.3.

The two other assistance data parts deal with Assisted GNSS. It is, however, divided into GPS-specific and GANSS-specific parts. This division is due to the backwards compatibility reasons. The GPS-branch was defined already in the RRLP Release 98, whereas the GANSS-branch was added only to the RRLP Release 7 [21]. In order to guarantee that the newer RRLP versions work with legacy UEs the GPS branch was left untouched, although the GANSS-branch could in principle support legacy GPS. Note that the modernized GPS -specific items are in the GANSS-branch.

The items in the GANSS-branch were outlined by the authors in [P4] and [P5] describing the overall and navigation model structures, respectively. It should be noted that these publications were written at the time when the RRLP Release 7 was being defined. The author and the co-authors have significantly contributed

to the structure and contents and of the RRLP GANSS-branch.

The GANSS-branch is divided into Common and Generic Assistance. Common Assistance refers to the assistance data that is independent of GNSSs. Such data include reference location, reference time, ionosphere models and EOPs (Earth-Orientation Parameters).

Reference location for AGNSS may be based on the coordinates of the serving BTS. As shown in Chapter 2.2.4 already the modest reference location accuracy of 3 km ($\mp 10 \mu s$) is sufficient for AGNSS. Moreover, cells are typically significantly smaller especially in urban areas.

Reference time is given with respect to the cellular frame timing of the serving BTS. The reference time field associates the GST with the specific bit in the serving BTS broadcast. In GERAN the longest frame structure is the hyperframe that consists of 2048 super frames that further consists of $51 \cdot 26 = 1326$ TDMA frames. Hence the hyperframe includes 2715648 TDMA frames. The TDMA frame number is given as modulo $2^{21} = 2097152$ frames in RRLP. Within the TDMA frame

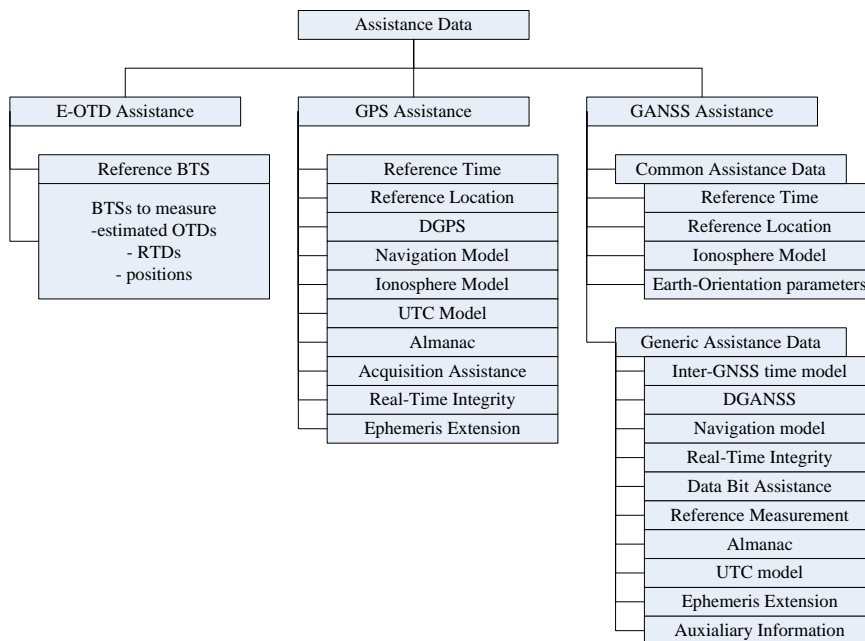


Figure 3.7: RRLP assistance data structure.

the time slot (0-7) and the specific bit (0-156) within the slot are defined. Also the drift of the cellular time with respect to the GST is given. The accuracy of fine time assistance is defined as $\mp 10 \mu\text{s}$ in the AGPS minimum performance requirements [90].

The third element in the Common Assistance are the ionosphere models that are discussed in Chapter 2.1.3 and further in Chapter 4.2.3. The ionosphere models available are the Klobuchar and NeQuick models as described in [13] and [14], respectively. The Klobuchar model is available in two flavors, one type for the global model and the other for the Japan area as defined in [16]. The final element in the common part are the EOPs that are needed to map the SV locations from the ECI (Earth-Centered Inertial) to the ECEF (Earth-Centered Earth-Fixed) coordinate system, and vice versa, for high-accuracy purposes.

The Generic Assistance includes GNSS-specific items and the Generic Assistance element repeats for each GNSS for which the assistance is provided. The items in the Generic Assistance are summarized in Table 3.1. Of these items Navigation model, Reference Measurements and Ephemeris Extension are discussed in detail.

Navigation Model. Almost every GNSS uses its own format for the navigation model. The Quasi-Keplerian orbit model utilized in legacy GPS in the GPS NAV broadcast is reused in Galileo, but the clock models differ in GPS NAV and Galileo. The navigation model utilized in GLONASS is unique and not used by any other GNSS. Further, modernized GPS and QZSS share the same GPS CNAV high-accuracy Quasi-Keplerian orbit and clock models. Finally, SBASs utilize their own format suitable for representing SV positions in geostationary orbits.

The navigation model in RRLP is called multi-mode navigation model due to it de-coupling orbit and clock model parameterizations from each other. The reasoning behind the approach is that it is expected that in the near future orbit and clock information utilized in the assistance data services originates not from the GNSS broadcasts, but from terrestrial sources. These include GNSS networks, such as [59], that track the SVs and predict their orbit and clock behavior days/weeks ahead. It is expected that this leads to the harmonization of the navigation model parameterizations across the GNSSs. Whereas now each GNSS utilizes its own

Table 3.1: Items in the Generic Assistance element in RRLP.

Item	Explanation
Inter-GNSS time model	Almost every GNSS uses its own time scale (see Chapter 4.1). These models provide mappings between the different GSTs.
DGANSS	Differential corrections to pseudorange measurements as well as the rate of change for the corrections
Navigation model	Orbit and clock models - see discussion
Real-Time Integrity	Information on the health of the SV broadcasts
Data Bit Assistance	Bit sequences for data wipe-off as explained in Chapter 2.2.5
Reference Measurements	SV signal reference measurements for the MS-assisted AGANSS - see discussion
Almanac	Coarse long-term orbit and clock models
UTC model	Model to relate GST to the UTC time scale
Ephemeris Extension	Accurate long-term orbit and clock models - see discussion
Auxiliary Information	Information on the SV signal capabilities and GLONASS frequency channel allocations

native parameterization, the data resulting from the orbit and clock predictions is independent of the GNSS. Hence, it is straightforward to represent the navigation data for all the GNSSs in a single chosen format. For example, it is expected that the Quasi-Keplerian parameterizations in GPS NAV or GPS CNAV will be the primary formats for representing the SV orbits for all the GNSSs. In [P6] the authors in fact show that the GPS NAV is a suitable parameterization for the GLONASS navigation data. RRLP is fully capable of supporting this potential harmonization. [P5] [S4] [91]

Reference Measurements. The purpose of the reference measurements is to provide the MS accurate information on the SV signals it can detect. The

information includes the SV position in the sky (elevation and azimuth), expected code phase, code phase search space, expected Doppler frequency and the Doppler rate of change as well as uncertainty. Reference measurements accelerate the acquisition of the SV signals by the MS, because the three-dimensional search space is significantly reduced using the assistance as shown in Chapter 2.2.4. Reference measurements are used in the MS-assisted mode, in which the SMLC gives the MS the reference measurements as well as instructions to return the SV signal measurements. The SMLC then determines the position based on the returned measurements. The reference location for the measurements may be obtained, for instance, through CID/ECID or E-OTD. Reference Measurements is called Acquisition Assistance in the GPS-branch.

Ephemeris Extension. Global GNSS reference networks provide information on the SV positions and velocities for the purposes of predicting satellite orbits several days or even weeks in advance. The concept of extended ephemeris refers to providing the MS the orbit/clock information for the SVs for up to 512 hours (approximately three weeks) in advance in order to accelerate the first fix. Having valid navigation model data in the UE shortens TTFF and improves user experience. The time saving results from the MS not needing to connect to the SMLC in order to fetch the orbit/clock assistance each time the positioning session is commenced [63]. The arrangement also potentially reduces the assistance server load. RRLP includes one flavor of this technology called Ephemeris Extension. The format is based on having a reference navigation model to which delta information are provided. The delta information are summed cumulatively to the reference model in order to reconstruct the full navigation model at a given time. In addition, the protocol definition also includes the possibility for the MS to perform a validity check for the extension data the MS has. Ephemeris Extension in RRLP is highly bit-optimized for the control plane.

Positioning Capability Exchange

The purpose of the capability exchange is to provide the SMLC capability to request the MS its assistance data, GNSS signal, positioning method and method

type capabilities. Receiving this information allows the SMLC to optimize the assistance data it provides to the MS. Also, the mechanism ensures that the SMLC does not provide the MS with assistance data the MS cannot take advantage of.

Measurement Request and Response

The Measurement Request allows the SMLC to request position or measurements from the MS. The request specifies the positioning methods, method types as well as the GNSS signals to be utilized, if applicable. Moreover, the request may already include assistance data. For example, if the SMLC requests E-OTD measurements it is natural for the SMLC to provide the MS directly the list of potentially hearable BTSs as well as their expected OTDs.

3.2.3.2 GERAN LCS broadcast

Whereas RRLP is a point-to-point protocol between the MS and SMLC, 3GPP TS 44.035 [92] defines point-to-multipoint broadcast for the assistance data. The broadcast originates from the CBC as indicated in Chapter 3.2.2.1. The broadcast allows the MS to have a valid set of positioning assistance data at all times without setting up a dedicated positioning session with the SMLC. This possibility shortens TTFF. The commercial deployment of cellular broadcast has, however, been limited due to the low number of users per cell making the dedicated point-to-point MS-SMLC connection more cost effective.

The broadcast includes the items required for E-OTD and AGPS. The E-OTD broadcast includes information on the reference BTS, neighboring BTSs, BTS locations as well as on the RTDs between the BTSs. The AGPS broadcast includes reference time, reference location, health information on the GPS SVs, navigation models and almanac. Moreover, DGPS (Differential GPS) corrections are broadcast. DGPS information is highly suitable for broadcast, because the same set of data is applicable for a large area, but needs to be updated fairly frequently. [92]

3.2.3.3 UTRAN RRC

3GPP TS 25.331 RRC defines the radio resource control protocol for the UE - UTRAN interface. RRC termination points are at the UE and S-RNC [93]. RRC provides, amongst other items, cell selection and re-selection, UE measurement reporting as well as the control of the measurement reporting. [86]

RRC provides similar functionality for positioning a UE in UTRAN as RRLP does for positioning an MS in GERAN. It should, however, be noted that whereas RRLP is solely a positioning protocol with termination points at MS and SMLC, RRC carries in addition to positioning payload also a plethora of other data.

The GPS and GANSS assistance data structures in RRC are almost similar to RRLP due to the RRC items being copies from RRLP. An example of a difference is the UE assistance data capabilities, which exists in RRLP, but not in RRC.

In fine time assistance the principle is the same as in RRLP, but the GST - cellular frame time is given in terms of the UTRAN frame structure. The physical layer in UTRAN is arranged into 4096 System Frames each having a length of 10 ms. The fine time assistance is given by indicating the GST at the beginning of the given System Frame at the resolution of 250 ns.

The downlink positioning method in UTRAN is the OTDOA-IPDL. Similarly to RRLP, RRC provides the UE with information on the reference Node B, neighboring Node Bs and on the expected SFN-SFN (System Frame Number) differences in order to reduce the search space. In UE-based OTDOA the Node B locations as well as accurate SFN-SFN time differences (RTDs) are given to the UE. RTT measurements may also be provided to the UE in the UE-based case.

Another difference is the existence of the IPDL parameters. Idle periods in the serving Node B transmission improve the hearability of neighboring cells as explained in Chapter 3.2.2.2. Idle period sequences are transferred to the UE in RRC and the interpretation of the parameters is given in 3GPP TS 25.214 [94]. Idle periods are generated in the downlink so that during an idle period the transmission from the Node B is seized on all the channels. Idle periods may be given in a burst so that between bursts there are periods during which no idle periods occur. Alternatively the idle periods may occur continuously in the frame structure.

3.2.3.4 E-UTRAN LPP

LPP is currently under development in 3GPP RAN2 and is defined in 3GPP TS 36.355 [95]. Despite the Stage 3 being in its early stages, design principles can be outlined based on the Stage 2 documentation [82]. LPP supports three different types of transactions:

- Exchange of positioning capabilities
- Transfer of assistance data
- Transfer of location information (measurements or position estimate)

Comparing the LPP transaction types to the RRLP components it can be concluded that the LPP capabilities will be very close to those in RRLP. The exchange of positioning capabilities is effectively mandated by the variety of assistance data formats and GNSSs supported in order to guarantee effective assistance data delivery. The capability exchange results in the UE knowing which data it may obtain and the E-SMLC knowing which data it can provide to the UE.

The LPP assistance data content will also be very close to that in RRLP. It is, however, recognized that because E-UTRAN is a new RAN without legacy implementations LPP can be defined from the first principles. For example, the GPS-GANSS branch division in RRLP and RRC is due to the requirement for backwards compatibility for legacy implementations. LPP does not have this restriction and will have a single AGNSS-branch.

The AGNSS-branch contents will be almost identical to those presented in Figure 3.7 for the RRLP GANSS-branch. However, LPP will not support ephemeris extension due to bandwidth restrictions in the control plane. In contrast to RRLP, LPP will also include an element for requesting assistance data. Moreover, LPP will also support requesting additional assistance data within the same session as well as having multiple simultaneous transactions.

The E-OTD branch in RRLP is replaced by the OTDOA-branch in the LPP. The branch will include elements to carry the list of expected eNode Bs and expected time differences for improved sensitivity. Moreover, LPP will carry RSS measurements from the UE to the E-SMLC for ECID.

One intriguing aspect of LPP is acknowledging its potential re-use outside E-UTRAN. The Stage 2 specification [82] states that

In contrast to GERAN and UTRAN, the E-UTRAN positioning capabilities are intended to be forward compatible to other access types and other position methods, in an effort to reduce the amount of additional positioning support needed in the future. This goal also extends to user plane location solutions such as OMA SUPL, for which E-UTRAN positioning capabilities are intended to be compatible where appropriate.

This means that whereas RRLP was designed only keeping GERAN in mind, LPP will be designed from the beginning to be forward compatible. The intention is to enable the use of the LPP in the future RANs, such as LTE-A. Moreover, RRLP is currently widely used in the user plane solutions including OMA SUPL [6] [7]. However, RRLP has been noticed to lack in capabilities [96] and OMA has been forced to fix the RRLP insufficiencies in SUPL in order to provide the required features in user plane. LPP should solve the issues by taking the user plane use into account already from the beginning. In [82] it is actually noted that the UE and E-SMLC can be any general client-server pair and, hence, LPP is not necessarily strictly bound to the E-UTRAN LCS architecture, but can for example be seamlessly used in the user plane over the SUPL bearer. Moreover, LPP high level structure is envisioned to be such that it is possible for other fora to extend the LPP [97].

Finally, 3GPP TS 36.305 also states that

To avoid creating new positioning protocols for future access types developed by 3GPP, and to enable positioning measurements for terrestrial access types other than E-UTRAN, the LPP is in principle forward-compatible with other access types, even though restricted to E-UTRAN access in this specification.

This is an additional interesting point. This holds that LPP can also carry measurements made from other air interfaces including GERAN and UTRAN

in order to provide a single, converged future-proof positioning protocol. Other potential additional access types includes IEEE 802.11 WLAN - see Chapter 4.2.

3.2.4 Minimum performance requirements

3GPP TS 45.005 [90] and 51.010-1 [98] define the minimum performance requirements and the corresponding test cases, respectively, for AGPS devices in GERAN under certain conditions. In UTRA-FDD the corresponding specifications are 3GPP TS 25.171 [64] and 34.171 [99], respectively. Note that the existence of such requirements is in contrast to stand-alone GPS devices that have no standardized requirements for their performance.

The specifications define the performance requirements that the MS/UE must fulfill in certain scenarios. The scenarios include

- Sensitivity with fine and coarse time assistance
- Nominal accuracy
- Dynamic range
- Multipath
- Moving scenario with a periodic update of position

Each test case defines the number of SVs, GPS constellation parameters (HDOP, Horizontal Dilution Of Precision), radio channel model, the type of assistance the UE receives and the SV signal levels. In addition, the pass criteria are also defined. Pass criteria include accuracy, yield and response time. Yield refers to the ratio of cases, in which the position response was delivered successfully in terms of defined accuracy and response time limits. To illustrate the test cases Table 3.2 shows the sensitivity test case with fine time assistance with the pass criteria given in Table 3.3. Currently 3GPP is working with similar minimum performance requirements for AGANSS.

Table 3.2: Sensitivity test with fine time assistance. AWGN stands for Additive White Gaussian Noise.

Parameter	Units	Value
Number of generated satellites	-	8
HDOP Range	-	1.1 to 1.6
Propagation conditions	-	AWGN
GPS Coarse time assistance error range	seconds	∓ 2
GPS Fine time assistance error range	μs	∓ 10
GPS Signal level for all satellites	dB ref 1 mW	-147

Table 3.3: Pass criteria for the sensitivity test with fine time assistance.

Success rate	2D position error	Max response time
95%	100 m	20 s

3.3 Open Mobile Alliance

Open Mobile Alliance was formed in 2002. The participants of this standardization forum include mobile operators, device vendors, network equipment manufacturers and test tool companies. OMA aims at developing enabler specifications that are end-to-end solutions covering the whole value chain. [5]

3.3.1 Secure User Plane Location protocol

3.3.1.1 Releases 1 and 2

The previously considered RRLP, RRC and LPP are positioning protocols for control planes of the cellular networks - they are integral parts of the cellular networks. However, in addition to control plane solutions there are also positioning protocols for the user plane (IP-network). The OMA location technology standards offering includes the OMA SUPL that is the location service and positioning protocol for the user plane.

The role of SUPL is two-fold. On one hand it provides services including triggered services, in which the SET (SUPL-Enabled Terminal) may be instructed

to report its location to the SUPL server in case a certain spatial or temporal criterion is fulfilled. These are called geographic and periodic triggers, respectively. SUPL also provides authentication of the SET, security, privacy and charging of services through other enablers defined by OMA, 3GPP or other standardization fora. Therefore, the OMA LCS architecture can be considered to be a complete end-to-end solution.

On the other hand SUPL also provides signaling for the actual positioning of the SET through the re-use of 3GPP-defined positioning protocols including RRLP, RRC and LPP. This is indicated in Figure 3.8 showing the SUPL protocol stack for Releases 1 and 2. The transport medium for SUPL is the TCP/IP (Transport Control Protocol) added with TLS (Transport Layer Security). The ULP (User Plane Location protocol) is the SUPL service protocol. ULP encapsulates the 3GPP positioning protocols as sub-protocols to the ULP. Note that since RRLP does not include the assistance data request it is defined in the ULP-layer.

The SUPL Release 1 [6] was finalized in 2007. The Release 1 included support for GSM, WCDMA and CDMA bearers including the NMRs and the RAN-specific positioning protocols. The TIA-801 [100] shown in Figure 3.8 is the positioning protocol for the CDMA networks.

The SUPL Release 2 [7] is currently in the Candidate Release phase and is expected to be completed early 2010. The SUPL Release 2 is backwards compatible with the SUPL Release 1, but also introduces a wide variety of improvements to the Release 1. Firstly, the bearer support was extended to include I-WLAN, WiMAX, LTE and UMB (Universal Mobile Broadband) along with their NMRs in the ULP layer. The SUPL Release 2 also includes the support for LPP as shown on the right hand side in Figure 3.8. Another improvement in the positioning technologies

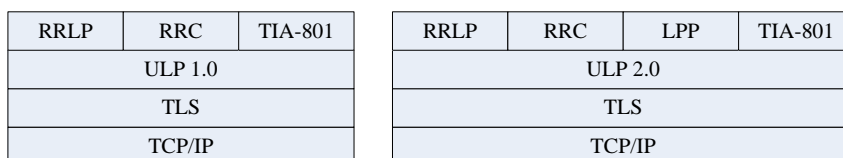


Figure 3.8: OMA SUPL protocol stack. Left: SUPL Release 1, Right: SUPL Release 2.

is the addition of fine time assistance for UTRAN in the ULP layer. This allows utilizing RRLP as the positioning protocol even in the case the SET would use UTRAN as the air interface. However, in [P7] the authors show that mixing the RAN-specific positioning protocols and different RANs has several disadvantages including breaking down the protocol layering.

The service side additions in the SUPL Release 2 include geographical and temporal events. Geographic triggers include cases for entering or leaving an area or being inside or outside of a given area. The events may be utilized to trigger other services including commercials. Another addition is the SET location delivery to a third party and the retrieval of the location of another SET for friend-finder type applications.

Figure 3.9 introduces a simplified OMA LCS Architecture. The architecture shows the major entities including SLP (SUPL Location Platform), SMSC (Short Messaging Service Center), WAP PPG (Wireless Application Protocol Push Proxy Gateway) and the SUPL Agent.

An SLP can act in three roles as H-SLP (Home SLP), V-SLP (Visited SLP) or E-SLP (Emergency SLP). Each SET has a dedicated H-SLP in the home network. The SET can be assigned a V-SLP to assist in positioning and E-SLPs are utilized

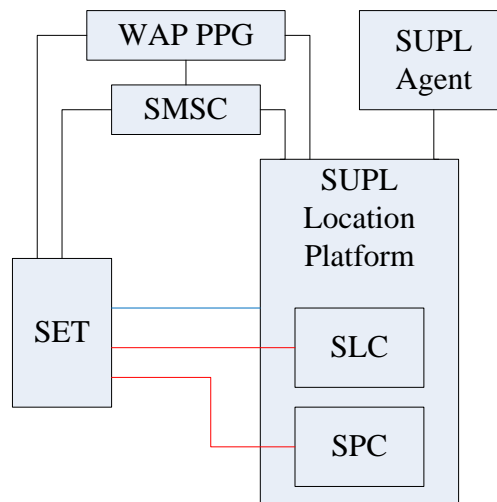


Figure 3.9: OMA LCS architecture.

in emergency call positioning. SLP has two functional elements called SLC (SUPL Location Center) and SPC (SUPL Positioning Center). SLC handles, amongst other items, security, roaming, charging and privacy. SPC, on the other hand, is responsible for the retrieval and delivery of positioning assistance data. The position determination may be divided between the SLC and SPC with SLC taking the responsibility for translating the Location ID, such as the BS information, to geographical coordinates for reference location purposes and SPC handling the other functions. The SET connects either to the SLP (proxy mode) or directly to the SLC/SPC (non-proxy mode in CDMA/CDMA2000 networks). The proxy and non-proxy modes are shown in blue and red lines in Figure 3.9, respectively.

In the SUPL framework the positioning session can either be SET-initiated or network-initiated. In the SET-initiated case the SET connects to the SLP (proxy-mode) and, for example, retrieves the required assistance data from the SLP. However, in the network-initiated case the SET must be notified so that it knows to set up the data connection to the SLP. The channels to deliver the notification are, for example, via an SMS (Short Messaging Service) through the SMSC or over WAP (Wireless Application Protocol) through the WAP PPG. In an exemplary case of the network-initiated session a SUPL Agent (a service requesting the location of the SET) external to the SET requests the SLP to position the SET. Having received the request the SLP sets up a network-initiated session with the SET using an SMS, positions the SET and returns the location to the SUPL Agent. The network-initiated sessions can, for example, be utilized in various commercial services, such as a child finder, as well as in lawful interception and positioning of emergency calls.

3.3.1.2 Upcoming Release 3

In January 2009 the OMA TP (Technical Plenary) approved a Work Item for the SUPL Release 3 [101]. The Work Item allows working with the items summarized in Table 3.4.

Currently the working group is finishing the work with the requirements specification for the SUPL Release 3 [102]. From the positioning technology

point-of-view the most significant additions will be the streaming of location and positioning information between two or multiple SETs via an SLP or between a SET and an SLP. Another important addition is the possibility to push assistance data change notifications from the SLP to the SET. These items will be discussed further in Chapter 4.2.

Table 3.4: SUPL Release 3 Work Item contents.

Item	Explanation
Improved location for IP emergency calls	Includes, for instance, SET-initiated emergency call positioning and latency reductions
Improved location performance	High-accuracy GNSS methods including PPP and RTK require both new assistance data types and new protocol features such as continuous periodic delivery of assistance data. Other potential improvements include the broadcast of assistance data over the IP network and assistance data change notification push to the SET. See Chapter 4.2 for more information.
Triggered location enhancement	New trigger types including equidistant trigger, SET-to-SET triggers (proximity), velocity and hybrid triggers.
Improved indoor location accuracy	Indoor location is highly relevant due to people spending majority of their time indoors. The challenge can at least partly be addressed by radiomap and fingerprint -based technologies. See Chapter 4.2.5.
Continued on Next Page...	

Item	Explanation
SET to SET location	SUPL Release 2 supports locating another SET. SUPL Release 3 will add the possibility to establish a data pipe between two SETs for continuously exchanging position and measurement data.
Authentication enhancements	Currently SUPL only supports authenticating SETs with a subscription to a 3GPP network. However, the emerging UE types including netbooks may not have a SIM (Subscriber Identification Module) that the currently used authentication mechanism requires.
Additional access networks	Includes, for instance, WLAN access. The SUPL Release 2 supports 3GPP I-WLAN, but not private WLAN access points due to security and authentication reasons.
Support for extended location information	Includes assistance for emerging sensor types including barometers. Also covered are descriptions for the state-of-motion, such as walking, cycling, driving etc. This information can be used to tailor LBS including commercials based on the type of motion.

Chapter 4

Way forward

This section discusses items that can be improved in the current and future assistance standards. Moreover, novel positioning techniques and complementing data are discussed in the context of how they could be applied in the positioning and assistance standards.

4.1 Harmonization of AGNSS information

The RRLP Release 8 [20] is a significant step towards convergence between GNSSs in the AGNSS standards. To exemplify, code phase measurements (reported and reference) are represented in the Release 8 in the units of time instead of GNSS-specific chip-units. Another example of the achieved harmonization between GNSSs in RRLP (as well as in RRC) is the representation of the GST.

Natively each GNSS uses its own time scale and parameterization of time. The different GSTs and their representations are summarized in Table 4.1. However, in RRLP the parameterization of the GST is independent of the GNSS. In the GANSS-branch the GST is a combination of GNSS Day and GNSS TOD (Time-Of-Day). The day count is 13 bits ranging $[0, 8191]$ days, i.e. approximately 22 years and TOD can be expressed in, for example, hours or seconds depending upon the need. Although the actual parameterization has been harmonized, the origin of the GST is still GNSS-dependent. The mapping between the RRLP GST

Table 4.1: GNSS System Times. TOW for Time-Of-Week. TAI for International Atomic Time.

GNSS	Time Scale	Parameterization	Origin
GPS	GPS System Time maintained within $\mp 1 \mu\text{s}$ to UTC(USNO) modulo 1 s	Week count (0-1023) and TOW in seconds	00:00:00 Jan 6 th 1980 UTC(USNO)
GLONASS	GLONASS System Time maintained within $\mp 1 \text{ ms}$ to UTC(SU)+3h	Day number within a 4-year period and TOD in seconds	–
QZSS	QZSS System Time controlled with respect to GPS	See GPS	See GPS
WAAS	Synchronized to GPS time within 50 ns	TOD in seconds	–
Galileo	Galileo System Time maintained within 50 ns to TAI - 19 s	Week count (0-4095) and TOW in seconds	00:00:00 UT Aug 22 nd 1999

and the native parameterization is, therefore, straightforward. The benefits of the arrangement include the simplification of the standard, because the same format can be utilized for every supported GNSS (listed in Table 4.1).

Code phase measurements and GST representation are examples of the harmonization achieved thus far. However, supporting a plethora of different representations in other data types is inevitable due to the need in some cases to support GNSS-specific native parameterizations as defined in the respective ICDs. As an example, RRLP includes native orbit and clock model representations for each GNSS, although they can be used interchangeably as explained in Chapter 3.2.3.1.

However, as discussed in Chapter 3 and in [P7] the expectation is that in the short-term the GNSS reference network services, such as IGS [59] and CODE (Center for Orbit Determination in Europe, [103]), will start providing AGNSS

data not originating from the GNSS broadcasts. These services track SVs and use other sources of information to extend the applicability of data including orbit and clock models. In such a case the resulting data is not in a GNSS-specific format, but the produced data can easily be mapped to the single standardized format used for all the GNSSs. These harmonization possibilities are discussed next.

4.1.1 SV clock models

Each GNSS SV broadcasts a model that relates the onboard SV time to the GNSS System Time as discussed in Chapter 2.1.3. While GPS, QZSS and Galileo use 2nd order polynomials, GLONASS and WAAS utilize 1st order polynomials. The number of bits and scaling factors for polynomial coefficients may also vary from GNSS to GNSS as shown in Table 4.2.

In addition to the polynomial terms the GPS and GLONASS clock models include the group delay parameter t_{gd} between the L1 and L2 signals - specifically in the case of GPS t_{gd} denotes the group delay between L1 P(Y) and L2 P(Y) signals. The Galileo Open Service broadcast, on the other hand, includes the group delay parameters between E5a/E5b and E1 signals. In addition, in the case of GPS CNAV broadcast (L1C, L2C, L5) the clock model also includes other inter-signal group delays. However, because this information is strictly GPS-specific it shall

Table 4.2: SV clock model in different GNSSs (bit counts / scale factors) and mappings to the harmonized SV clock model. QZSS utilizes both GPS NAV and GPS CNAV parameterizations.

	NAV	CNAV	SBAS	GLONASS	Galileo
Native a_0	22b / 2^{-31}	26b / 2^{-35}	12b / 2^{-31}	22b / 2^{-30}	31b / 2^{-34}
Native a_1	16b / 2^{-43}	20b / 2^{-48}	8b / 2^{-40}	11b / 2^{-40}	21b / 2^{-48}
Native a_2	8b / 2^{-55}	10b / 2^{-60}	–	–	6b / 2^{-59}
Native t_{gd}	8b / 2^{-31}	13b / 2^{-35}	–	5b / 2^{-30}	10b / 2^{-32}
Scaling of a_0	2^4	1	2^4	2^5	2^1
Scaling of a_1	2^5	1	2^8	2^8	1
Scaling of a_2	2^5	1	–	–	2^1
Scaling of t_{gd}	2^4	1	–	2^5	2^3

not be considered in the harmonization.

The SV clock model representations may be harmonized across the GNSSs by defining an SV clock model

$$\delta t_{GST}^{SV}(t_{GST}) = a_0 + a_1(t_{GST} - t_{GST}^{ref}) + a_2(t_{GST} - t_{GST}^{ref})^2 \quad (4.1)$$

with a_0 , a_1 and a_2 defined in Table 4.3.

The bit counts and scaling factors in the harmonized model have been chosen such that the native formats may be mapped to the harmonized model without loss of information. Mapping can be accomplished by re-scaling the coefficients received in the GNSS broadcast. The scalings required for native parameterizations are summarized in Table 4.2.

4.1.2 GST - UTC models

Each GNSS broadcasts a model that relates its specific GNSS System Time to the specified UTC time scale. The various UTC standards utilized by different GNSSs are summarized in Table 4.4. Moreover, GNSSs utilize different parameterizations for the GST-UTC model. Whereas GLONASS utilizes a 0th-order, GPS NAV and Galileo use a 1st-order and GPS CNAV uses a 2nd-order polynomial. The native representations with bit counts and scale factors are summarized in Table 4.5.

The GST-UTC models may be harmonized across the GNSSs by defining a UTC time model through a polynomial

$$t_{UTC}(t_{GST}) = t_{GST} - \left(\Delta t_I + a_0 + a_1(t_{GST} - t_{GST}^{ref}) + a_2(t_{GST} - t_{GST}^{ref})^2 \right), \quad (4.2)$$

Table 4.3: Harmonized SV clock model. All the parameters are 2's complement with the MSB being the sign bit.

Parameter	Bit count	Scale factor	Units
a_0	32	2^{-35}	s
a_1	21	2^{-48}	s/s
a_2	13	2^{-60}	s/s^2
t_{gd}	13	2^{-35}	s

Table 4.4: UTC standards.

System	Reference UTC	Maintained by
GPS/WAAS	UTC(USNO)	US Naval Observatory
GLONASS	UTC(SU)	Russia
QZSS	UTC(NICT)	National Institute of Information and Communication Technology
Galileo	To be decided	To be decided

Table 4.5: UTC-GST models and mappings to the harmonized UTC model. QZSS utilizes both GPS NAV and GPS CNAV parameterizations. WAAS utilizes GPS NAV parameterization.

	GPS NAV	GPS CNAV	GLONASS	Galileo
Native a_0	32 b / 2^{-30}	16 b / 2^{-35}	32 b / 2^{-31}	32 b / 2^{-30}
Native a_1	24 b / 2^{-50}	13 b / 2^{-51}	–	24 b / 2^{-50}
Native a_2	–	7 b / 2^{-68}	–	–
Native $\Delta t'_I$	8 b / 1	8 b / 1	–	8 b / 1
Scaling of a_0	2^5	1	2^4	2^5
Scaling of a_1	2^1	1	–	2^1
Scaling of a_2	–	1	–	–
Scaling of $\Delta t'_I$	1	1	–	1

where the polynomial coefficients are defined in Table 4.6. The term Δt_I is the integer number of seconds between the GST and UTC. In case of GPS, QZSS, SBAS and Galileo this corresponds to the leap second count between UTC and GPS/Galileo System Time (currently 15 s) and is carried in the parameter $\Delta t'_I$ in the harmonized model. Because in the case of GLONASS Δt_I is constant 10800 s (3 hours), the parameter $\Delta t'_I$ need not be carried for the GLONASS GST-UTC model. Again the native parameterizations may be mapped to the harmonized model by re-scaling. The re-scaling factors are defined in Table 4.5.

The harmonized UTC model also needs to specify to which UTC standard the GST is being mapped to. It might also be beneficial to standardize the UTC standard to be used. In addition, the time of the next leap second should be

Table 4.6: Harmonized UTC model. All the parameters are 2's complement with the MSB being the sign bit.

Parameter	Bit count	Scale factor	Units
a_0	37	2^{-35}	s
a_1	25	2^{-51}	s/s
a_2	7	2^{-68}	s/s ²
$\Delta t'_I$	8	1	s

included in the model, if required. For example, in the case of GPS the GST-UTC relation is discontinuous at the leap second boundary, whereas GLONASS system time follows UTC at the boundary as indicated in Table 4.1.

4.1.3 Inter-GNSS time models

The inter-GNSS time model maps one GST to another. Again the GNSSs utilize various GNSS-specific parameterizations, but they all are 0th-, 1st- or 2nd-order polynomials with the bit counts and scale factors defined in Table 4.7.

The inter-GNSS time models may be harmonized by defining a polynomial

$$t_{GST_1}(t_{GST_0}) = t_{GST_0} - \left(\Delta t_I + a_0 + a_1(t_{GST_0} - t_{GST_0}^{ref}) + a_2(t_{GST_0} - t_{GST_0}^{ref})^2 \right) \quad (4.3)$$

with bit counts and scale factors for the coefficients defined in Table 4.8. Information to calculate Δt_I is either carried in the parameter $\Delta t'_I$ in the harmonized model or is known to be zero. The term Δt_I is zero, when describing GPS-QZSS-SBAS-Galileo GST offsets. This is because GPS, QZSS and SBAS clocks are referenced to the GPS GST as indicated in Table 4.1. Moreover, in the beginning of the Galileo GST it was set 13 s ahead of UTC (19 s behind TAI). Hence, the integer second difference between GPS and Galileo GSTs is zero. Thus $\Delta t'_I$ need not be carried for these inter-GNSS time models.

However, in case the other GST is GLONASS the term $\Delta t'_I$ needs to be carried to the UE, because GLONASS directly follows UTC while the other time scales are continuous. Hence, when representing the offset between GLONASS and GPS/QZSS/SBAS/Galileo GSTs $\Delta t'_I$ reflects the leap second count between GPS/Galileo GSTs and UTC.

Table 4.7: Mappings from the native GNSS - GNSS time models to the harmonized format. QZSS utilizes the GPS CNAV parameterization.

	GPS CNAV	GLONASS	Galileo
Native a_0	16 b / 2^{-35}	22 b / 2^{-30}	16 b / 2^{-35}
Native a_1	13 b / 2^{-51}	–	12 b / 2^{-51}
Native a_2	7 b / 2^{-68}	–	–
Scaling of a_0	1	2^5	1
Scaling of a_1	1	–	1
Scaling of a_2	1	–	–

Table 4.8: Harmonized GNSS-GNSS time model. All the parameters are 2's complement with the MSB being the sign bit. Parameter $\Delta t'_I$ is included only, when the other GST is GLONASS. In such case the term Δt_I is captured by adding 10800 s to $\Delta t'_I$.

Parameter	Bit count	Scale Factor	Units
a_0	27	2^{-35}	s
a_1	13	2^{-51}	s/s
a_2	7	2^{-68}	s/s^2
$\Delta t'_I$	8	1	s

If the inter-GNSS time model is received in a GNSS-specific format it needs to be re-scaled to the harmonized format. The mappings for the parameters are defined in Table 4.7. No scaling is required for the GPS CNAV and Galileo formats as shown. The UTC - GPS/Galileo GST leap second count can be obtained from the UTC model. The harmonized inter-GNSS time model has been proposed for LPP.

4.1.4 Fine time assistance

The significance of fine time assistance was discussed in Chapter 2.2.4 and the GERAN- and UTRAN-specific representations in Chapters 3.2.3.1 and 3.2.3.3, respectively. In fine time assistance the GST is bind to the cellular time by expressing the GST at the given cellular time expressed in terms of frames and sometimes also bits. These representations are currently GERAN- and UTRAN-specific in RRLP and RRC, respectively.

One option to harmonize fine time assistance across GNSSs and RANs is to state the GST and the time passed from the beginning of the longest frame structure at the given GST.

As discussed in Chapter 3.2.3.1 the longest frame structure in GERAN is the hyperframe that includes 2715648 TDMA frames. Each frame is divided into 8 time slot each lasting 15/26 ms [104]. Therefore, the hyperframe length is 12533.76 seconds that can be presented using 14 bits (full seconds). Each time slot contains 156.25 bits [104] resulting in the bit duration of 3.96 μ s. The resolution of fine time assistance in RRLP is one bit [20] and, hence, the fractional part must exceed this resolution. Expressing this requires 19 bits, since $2^{-19} \approx 1.9 \mu$ s. The drift in RRLP is expressed with 7 bits and scale factor of 2^{-30} such that the range is approximately $\mp 6 \cdot 10^{-8}$ s/s.

In UTRAN the cellular time is expressed in terms of SFN running from 0 to 4095 and each radio frame is 10 ms in duration [105]. In UTRAN the longest frame structure is, therefore, 40.96 seconds (4096 radio frames). This can be represented with 6 bits (full seconds). The required 250-ns resolution [86] can be achieved with 22 bits, because $2^{-22} \approx 238$ ns.

The drift between the Node B time and GST is expressed in terms of UTRAN chips. The possible values are $\{0, \mp 1, \mp 2, \mp 5, \mp 10, \mp 15, \mp 25, \mp 50\}$ in units of 1/256 chips per seconds. The chip rate is 3.84 Mchips per second [70]. The maximum drift rate is, hence, approximately $\mp 5 \cdot 10^{-8}$ s/s, which is essentially the same as in GERAN.

In E-UTRAN the cellular time is expressed in terms of 10-ms radio frames that are numbered from 0-1023 (10 bits) [106]. The longest frame structure, hence, is 10.24 seconds and the resolution as well as the drift rate have been proposed to be the same as in UTRAN [89]. Therefore, GERAN and UTRAN considerations already provide sufficient ranges and resolutions for the parameters to be used in the E-UTRAN fine time assistance as well.

Uncertainty in the fine time assistance is presented as in the RRLP by a 7-bit index k ranging [0, 127]. The mapping of the index k to the uncertainty figure is given in [20] as

$$t_{uncertainty}(k) = 0.0022 \cdot ((1 + 0.18)^k - 1) \mu s \quad (4.4)$$

allowing representing the uncertainty range [0,2.96) s. However, in LPP the coefficients have been proposed to be modified from 0.0022 and 0.18 to 0.5 and 0.14, respectively to represent range [0,8.43) s. The reason for change is the unrealistically small resolution of time uncertainty in RRLP [89]. The LPP representation may be utilized in the harmonized fine time model.

To summarize, the harmonized fine time assistance for GERAN, UTRAN and E-UTRAN radio interfaces can be expressed with the information presented in Table 4.9. If the GST needs to be mapped to the bit or radio frame boundary, this can be easily accomplished utilizing the drift rate. Note that the bit count in the day number has been increased to 15 bits to cover range [0,32767] days corresponding to approximately 89 years. The harmonized fine time assistance has been proposed for LPP.

4.1.5 Performance aspects

Harmonization of the discussed models necessarily leads to a non-bit-optimized representation of the models. Table 4.10 summarizes the penalties introduced. In the table the figures indicate how many bits less the native representation would use. Note that the penalties have been calculated by taking into account that if native models are to be mapped to the harmonized representation, not all the

Table 4.9: Harmonized fine time assistance.

Parameter	Bit count	Scale factor	Units
GNSS day	15	1	day
GNSS Time-of-day	17	1	s
Cellular time	14	1	s
Cellular time fractional part	22	2^{-22}	s
Drift	7	2^{-30}	s/s
Uncertainty	7	–	–

parameters in the harmonized model are needed. For example, in the case of ASN.1 (Abstract Syntax Notation) coding the parameter inclusion in the information element can be optional.

On average the use of harmonized models results in a few tens of bits overhead in the bit consumption. The most profound penalty arises from the overhead in the SV clock model, which is provided for every SV for which assistance is provided. However, given that the penalties are in the order of hundreds of bits per assistance transaction it may be argued that the simplification introduced by using a single model for all the GNSSs is more significant than the minor increase in the bit consumption. This is especially true nowadays, because the assistance services are more and more relying on the user plane. In the bandwidth-limited control plane deployments it may still, however, be advisable to use the bit-optimized native models.

Harmonized fine time assistance was shown to take 82 bits in Table 4.9. In the cellular standards fine time assistance takes 76 and 75 bits in RRLP and RRC, respectively. Therefore, penalties are 6 and 7 bits with respect to GERAN and UTRAN native representations, respectively. Note that these counts neither take into account the bits used for representing to which GST the model maps the cellular time nor the identification of the BS, of which time is considered. The

Table 4.10: Penalty due to model harmonization. The table shows the number of bits the use of native models would save. In inter-GNSS time models if the other GST is GLONASS the size of the harmonized model increases by eight bits to account for the parameter $\Delta t'_j$. In such a case the penalty increases by eight bits for all the systems.

	SV clock	GST-UTC	Inter-GNSS
Harmonized	79 bits	77 bits	47 bits
GPS NAV	-25 bits	-6 bits	–
GPS CNAV	-10 bits	-33 bits	-11 bits
SBAS	-33 bits	-6 bits	–
GLONASS	-28 bits	-5 bits	-5 bits
QZSS	-25/-10 bits	-6/-33 bits	-11 bits
Galileo	-11 bits	-6 bits	-12 bits

increase in the bit consumption is so insignificant that the generalized fine time assistance is preferred over the various native representations.

4.2 Missing information elements in positioning standards

4.2.1 Navigation models

As discussed in Chapters 3.2.3.1 and 3.2.3.3 the current AGNSS standards support the native navigation model parameterizations defined in the respective ICDs. As noted, the GANSS-branch in RRLP includes five different orbit parameterizations: Keplerian, NAV Keplerian, CNAV Keplerian, GLONASS and SBAS ECEF parameterizations. These are required to support native formats of Galileo, legacy GPS, modernized GPS, GLONASS and SBAS systems. The standards also allow for mixing the parameterizations and the GNSSs in order to provide navigation models in non-native formats and use the same parameterization for all the applicable GNSSs.

By moving away from using different parameterizations for different GNSSs and from utilizing data from the GNSS broadcasts brings advantage in terms of harmonized performance. Firstly, using the same parameterization removes the possible performance differences due to the actual parameterization. Performance issues may include accuracy, but also the validity period of the model. As an example, GPS NAV is typically valid for four hours, but GLONASS orbit model only for half-an-hour. Secondly, using, say, IGS data improves accuracy over the broadcast models. The IGS ultra-rapid 48-h orbit/clock prediction for GPS has 5-cm 1D RMS accuracy in spatial and 3-ns RMS accuracy in temporal domains [59]. In contrast, the broadcast orbit accuracy is in the order of few meters RMS for GPS [P6]. The clock model harmonization suffers from the variety of clock qualities onboard SVs.

Publication [P6] shows that GPS NAV parameterization can be used for GLONASS SVs as well. The authors find that reparameterizing IGS orbit data

for GLONASS into GPS NAV format typically results in a few decimeter fit residual in maximum. Using GPS CNAV can be expected to result in even smaller residuals. However, the study in [P6] does not address, if GLONASS orbits can be described in GPS CNAV parameterization. Moreover, the same applies to other GNSSs including SBAS, Galileo and QZSS. For example, in the case of SBAS it can directly be seen that the field for expressing the semi-major axis in GPS CNAV is not applicable for SBAS SVs as such.

Given these considerations and the benefits of the harmonization it might be advisable to study the GPS CNAV format in order to see, how suitable it is for other systems than GPS. Having done this a potentially modified version of GPS CNAV might be included in the standards, which model would then become the de-facto navigation model in the AGNSS standards. The reasons for considering GPS CNAV include that it is compact in terms of bit consumption, but allegedly also very accurate in describing the orbits given that the original orbit data is accurate. The improved performance of GPS CNAV over GPS NAV arises from, for example, various resolution improvements as well as new rate-of-change parameters for certain parameters including the semi-major axis.

It is also possible to consider the inclusion of SP3-type (Standard Product 3, [107]) representation of the SV navigation model into the AGNSS standards. The SP3 representation is a set of SV position-velocity-time records at given epochs, which records are interpolated between the epochs [108]. However, the SP3-type approach is more bandwidth-demanding compared to GPS CNAV. Assuming 15-minute epoch spacing, 32 bits for each of the three position coordinates, 16 bits for the SV clock bias and 4-hour worth of data, the SP3 consumes approximately 1900 bits per SV, while GPS CNAV consumes only one fourth of this. Another approach to consider is providing the UE the polynomial coefficients that result from the interpolation process [108]. Before proposing these (SP3, polynomial coefficients) to the standards, an analysis of their advantages over, say, GPS CNAV as well as performance in terms of accuracy and bandwidth consumption must be made.

4.2.2 Navigation model request

A tangible deficiency in the current standards is related to the assistance data request. Assume that the UE has a GPS+GLONASS dual-AGNSS capability and the UE can utilize the native GPS NAV and GLONASS navigation models as well as GPS NAV model for GLONASS. Next assume that the UE wishes to utilize GPS NAV model for both GPS and GLONASS. The UE sends an assistance data request to the assistance server. In the case that the server has GPS NAV models for GPS and GLONASS SVs the server returns the models as requested. However, in the case that the assistance server does not have GPS NAV models for GLONASS SVs the server might return an error. Alternatively, a defined fallback might be sending the native models to the UE.

The above example illustrates that in the case the UE is capable of utilizing several navigation models for a given GNSS, the assistance data request/delivery process becomes challenging. This is due to the UE not knowing, which data the server has, when the UE requests assistance. Moreover, the situation becomes even more challenging as there are more GNSSs and more navigation model types. In addition, for example the SP3 is not native to any GNSS so introducing the SP3 support would complicate the situation further.

One option to solve the problem is to re-define the request from an explicit request for some model type to a preference list -based solution. With the preference list the UE requests models from the AGNSS server in the order of descending preference. In the example above the highest record in the request for GLONASS would have indicated GPS NAV parameterization and the second record GLONASS native format. The arrangement guarantees that the UE always receives data in the format the UE can utilize and also that the UE gets the most preferred model the AGNSS server has available.

Implementing a preference list -based request is straightforward in the AGNSS standards. The different parameterizations only need to be assigned enumerated values that the UE then stacks in the request list in the order of preference. The approach has been proposed for LPP [109].

4.2.3 Ionosphere models

The globally averaged 1σ magnitude of ionospheric delay is 7 meters [31]. Under the ionosphere storm conditions the delay may be in the order of a hundred meters [110]. Hence the accuracy of the ionosphere model has profound effect on the GNSS performance.

The GPS broadcast [13] includes a Klobuchar ionosphere model that compensates 50-60% of the group delay [111]. Moreover, multi-frequency GNSS UEs may estimate the ionospheric error due to the dispersive nature of the charged medium in ionosphere. The dispersive nature can be seen by considering the Appleton-Hartree model for the refractive index n in homogeneous plasma. Assuming non-interacting electrons that are not subject to external magnetic field the model reduces to [111]

$$n^2 = 1 - \frac{f_N^2}{f^2} \quad (4.5)$$

with f and f_N^2 being the signal and plasma frequencies, respectively. Plasma frequency (frequency of charge density oscillations) is defined by

$$f_N^2 = N_e \frac{q_e^2}{4\pi^2 \epsilon_0 m_e} = 80.6 \frac{\text{m}^3}{\text{e}^-} N_e \text{ Hz}^2, \quad (4.6)$$

where N_e is the electron density given in units of e^-/m^3 , q_e and m_e the electron charge and mass, respectively, and ϵ_0 the permittivity of free space. Writing $n = \sqrt{1 - \frac{f_N^2}{f^2}} = 1 - N_I$ and expanding yields for the first-order for the ionospheric refractivity N_I

$$N_I = \frac{f_N^2}{2f^2} + \dots \approx \frac{40.30}{f^2} N_e. \quad (4.7)$$

Now, the group (n_g) and phase (n_ϕ) refractive indices are related by

$$n_g = n_\phi + f \partial_f n_\phi \approx 1 + N_I, \quad (4.8)$$

in which the last result is obtained by substituting Equation 4.5 for n_ϕ and expanding for the first order. Evaluating the group and phase velocities for the first order yields

$$\begin{aligned}
c_g &\approx c(1 - N_I) = c \left(1 - \frac{40.30}{f^2} N_e \right) \\
c_\phi &\approx c(1 + N_I) = c \left(1 + \frac{40.30}{f^2} N_e \right)
\end{aligned} \tag{4.9}$$

showing that the group velocity is smaller than the phase velocity. Hence, the code phase and carrier phase measurements are delayed and advanced, respectively, due to ionosphere. This is called code-carrier divergence. The result also shows the dispersive (frequency-dependent) nature of ionosphere and its f^{-2} -dependency.

The dispersive nature of ionosphere allows for dual- and triple-frequency GNSS users to estimate/remove the ionosphere delay/advance. However, the IF linear combination leads to magnified noise and the net effect is still a 1-meter RMS delay residual [56]. Moreover, when using the GPS (Klobuchar) or Galileo (NeQuick) broadcast ionosphere models, the delay residual is still potentially tens of meters [56] for single-frequency users. Therefore, it is beneficial to consider more advanced ionosphere data for AGNSS standards in addition to the currently carried models even in the case the UE had a multi-frequency receiver.

One option is to carry ionosphere zenith delays in a grid as in the SBAS broadcast. The SBAS broadcast includes ionospheric delay data at discrete grid points. The zenith ionosphere delay and the estimated delay error are given for each grid point. The zenith delay is mapped to the slant (along the signal path) delay by taking into account the SV signal ionosphere pierce point as well as the effective ionosphere thickness. [25]

The WAAS performance can be considered as the accuracy limit achievable with monitoring networks as of now. When approaching the next solar maximum predicted to take place in May 2013 [112] it is interesting to discuss the performance under the storm conditions. In [110] the authors report WAAS performance under severe ionosphere storm conditions, in which zenith range delays up to 70 m were observed. The authors deduce that the WAAS ionosphere delay data had at its worst 10-m errors in the zenith range delay model. Although not quantized in the study, it can, however, be anticipated that the Klobuchar model performance would have been significantly worse under the same conditions.

Although the SBAS approach shows good performance, its problem is the requirement to constantly retrieve the data from the SBAS broadcast - or alternatively over the data channel, if such ionospheric grids could be carried over the standardized positioning protocols. The former solution suffers from the need to decode data payload constantly without the possibility for power saving in the UE. The latter solution suffers from using the data channel constantly reducing battery life and inducing data transfer costs.

Therefore, it can be deduced that the real-life feasibility of advanced ionosphere model depends on capabilities to forecast ionosphere behavior. Therein, for example, publication [113] introduces a forecasting technique, which appears to produce 10%-accurate 24-h predictions in mild ionosphere conditions. In absolute terms the reported forecasting errors are approximately $3 \cdot 10^{16} \frac{e^-}{m^3}$ corresponding to roughly 0.5-meter range error in L1 [114]. Obviously, worse performance is expected under storm conditions. Having ionosphere delay forecast allows then providing the UE ionosphere delay assistance for an extended period in some suitable parameterization. Ionosphere grid is one such parameterization. Another one that could be considered is the semi-physical regional TEC (Total Electron Content) model NTCM1 (Neustrelitz TEC Model 1) developed for Europe [115]. NTCM1 is defined by

$$TEC = \sum_{i=1}^5 \sum_{j=1}^3 \sum_{k=1}^2 \sum_{l=1}^2 H_i(TOD) Y_j(DOY) L_k(\varphi, \psi, TOD, DOY) S_l(F10), \quad (4.10)$$

where $TEC = \int N_e(s) ds$ and is integrated along the signal path, $H_i(TOD)$ the diurnal variation that is dependent on TOD, $Y_j(DOY)$ seasonal variation that is dependent on DOY (Day-Of-Year) and $L_k(\varphi, \psi, TOD, DOY)$ is the solar Zenith angle dependence with φ being latitude and ψ longitude. Finally, $S_l(F10)$ is the solar activity dependence that is a function of the $F10$ 10.7-cm solar flux index. The improved NTCM2 takes into account the geomagnetic latitude also [114].

The underlying 60-parameter model provides large scale TEC values taking into account daily and yearly changes as well as location and the sun activity. The model is refined with current measurements in such a way that the resulting

NTCM provides measured values at some locations and model values at other locations. The strength of the approach is that even in the case of small number of observations the model provides reasonable ionosphere corrections to the users [116]. Furthermore, the prospects are that the model evolves to characterize the 3D behavior of TEC [117]. An exemplary use of NTCM in assistance standards could include providing the full model daily or so and then providing model parameter deltas so that an extended model applicability period is achieved.

Both 2D/3D ionosphere delay grids and semi-statistical models are potential candidates for the inclusion into the future AGNSS standards, because of their clear potential to significantly improve the positioning performance especially in the single-frequency case. The availability of reliable, high-quality ionosphere predictions is a key enabler for this improvement to take place. The models itself are straightforward to implement in the AGNSS protocols and requesting such assistance from the AGNSS server requires knowing only the coarse location of the UE. From the assistance service provider point-of-view one of the issues with ionospheric delay data, however, is that it may be difficult to predict the period of applicability for the data. Hence, the UE must either periodically check for the availability of updates or alternatively the UE can subscribe to assistance data change notifications (see Chapter 4.2.8).

4.2.4 Troposphere models

An overview of troposphere-related effects in GNSS was given in Chapters 2.1.3 and 2.1.4. The effect is most typically mitigated using statistical models in the UE that only depend on the information that the UE has by default. These include time and the UE orthometric height h . However, there are also more advanced models that, however, require external input to the AGNSS receiver. Such information does not exist in the current assistance standards.

The STD (Slant Total Delay) experienced by the SV signal due to the troposphere is given by

$$STD = ZHD \cdot m_h(\varphi, \psi, h, \nu) + ZWD \cdot m_w(\varphi, \psi, h, \nu), \quad (4.11)$$

where ZHD and ZWD are the Zenith Hydrostatic Delay (dry) and Zenith Wet Delay, respectively, and m_h and m_w the hydrostatic and wet mapping functions to map the zenith delays to the slant delays. The mapping functions are functions of the UE location and the satellite elevation ν .

The hydrostatic and wet components can be expressed by

$$\begin{aligned} ZHD &= 10^{-6} \int_{h_0}^{h_T} N_h(h) dh \\ ZWD &= 10^{-6} \int_{h_0}^{h_T} N_w(h) dh, \end{aligned} \tag{4.12}$$

where N_h and N_w are the hydrostatic and wet refractivities, respectively. The integration takes place along the vertical from the reference altitude h_0 to the effective troposphere height h_T . Hydrostatic refractivity is fairly straightforward to integrate as N_h is only a function of total pressure and temperature. Ideal gas law may be applied. The N_w is more problematic due to its dependency on the spatial variability of the partial water pressure and temperature along the integration path. Models to approximate ZHD and ZWD based on the surface meteorological data including pressure, temperature and partial water vapor pressure do exist, but they often fail to predict the wet component sufficiently [57].

The residual troposphere delays are in the order of few decimeters, when there is no meteorological data available [56] [118]. The residual is attributable to the wet delay that the statistical models cannot account for adequately. However, because the wet component contributes only 10% of the total troposphere delay [31] the results without meteorological data are still reasonable.

However, whenever higher accuracy is required, it is advisable to estimate the troposphere delays either directly through measurements using GNSS receiver networks or from the weather forecasts [111]. For example, in [119] it is shown that significant (up to 78% in terms of standard deviation) improvements in the up-direction in the baseline solutions are achieved, when deriving tropospheric delay data from the numerical weather predictions as compared to using Saastamoinen's model and Niell's mapping function. Moreover, [118] shows that 1-cm RMS

troposphere delay residuals are achievable, when deriving zenith troposphere delay estimate from a dense GPS reference station network.

One option to transfer troposphere delay -related information is the format defined for QZSS [16]. The L1-SAIF broadcast from the QZSS SVs includes Zenith delays for a number of grid points within the QZSS coverage area. The UE stores the zenith delay for the relevant grid points, calculates the troposphere pierce points for the GNSS SV signals, maps the zenith delays to the slant delays and applies the corrections. The resulting slant delays are used to correct both the code phase and carrier phase measurements.

Mapping the zenith delays to the slant delays is not a trivial task either, but the mapping functions may assume arbitrarily complex forms. Niell's mapping function is commonly used and it assumes the form

$$\begin{aligned}
 m_h(\varphi, h, DOY, \nu) &= \frac{1 + \frac{a}{1 + \frac{b}{1+c}}}{\cos(\nu) + \frac{a}{\cos(\nu) + \frac{b}{\cos(\nu)+c}}} + h \left(\frac{1}{\cos(\nu)} - \frac{1 + \frac{a_h}{1 + \frac{b_h}{1+c_h}}}{\cos(\nu) + \frac{a_h}{\cos(\nu) + \frac{b_h}{\cos(\nu)+c_h}}} \right) \\
 m_w(\varphi, \nu) &= \frac{1 + \frac{a_w}{1 + \frac{b_w}{1+c_w}}}{\cos(\nu) + \frac{a_w}{\cos(\nu) + \frac{b_w}{\cos(\nu)+c_w}}},
 \end{aligned} \tag{4.13}$$

with h being the UE orthometric altitude and $a, b, c, a_h, b_h, c_h, a_w, b_w$ and c_w the model parameters that can be found in tables. a_w, b_w and c_w are functions of the latitude. a, b and c are functions of both latitude and day-of-year, while a_h, b_h and c_h are fixed. Other options for mapping functions include deriving the mapping function parameters directly from the meteorological data. An example of this class of mapping functions are VMFs (Vienna Mapping Functions) [120]. VMFs are based on continuous fractions similar to Niell's function. Both the hydrostatic and wet mapping functions have three parameters $a_h^{VMF}, b_h^{VMF}, c_h^{VMF}$ and $a_w^{VMF}, b_w^{VMF}, c_w^{VMF}$. For example, in [121] it is shown that it is feasible to derive mapping functions from numerical weather predictions.

The Niell's mapping function has the drawback that it is isotropic. However, the spatial variability (anisotropy) of water vapor is of much concern in accurate

Table 4.11: Contents of troposphere assistance.

Parameter	Contents
Reference location	The grid point (origin of ENU)
Period of applicability	Applicability begin and end times
Zenith delays and rates of change	ZHD , $\partial_t ZHD$, ZWD and $\partial_t ZWD$
Mapping function parameters	a_h^{VMF} , b_h^{VMF} , c_h^{VMF} a_w^{VMF} , b_w^{VMF} , c_w^{VMF}
Gradient terms	G_n and G_e
Barometric pressure	Reference pressure for obtaining reference altitude from barometer

GPS applications [111]. By expanding N_w in the ENU-coordinate (East-North-Up) system it is obtained

$$N_w(\underline{e}, \underline{n}, \underline{u}) = N_w(\underline{0}, \underline{0}, \underline{0}) + (\partial_{\underline{e}} N_w)^T \underline{e} + (\partial_{\underline{n}} N_w)^T \underline{n} + (\partial_{\underline{u}} N_w)^T \underline{u} \quad (4.14)$$

with \underline{e} , \underline{n} and \underline{u} being the delta vectors in the east, north and up directions, respectively. Integrating in height yields [111]

$$ZWD(\alpha, \beta) = ZWD_0 + \frac{1}{\tan(\beta)} (G_n \cos(\alpha) + G_e \sin(\alpha)) \quad (4.15)$$

with ZWD_0 being the zenith wet delay at the reference location, G_n and G_e being the appropriate gradients along north and east directions, respectively, and with α and β being the azimuth and elevation of the direction of travel in the ENU system. The gradients can, again, be estimated from the measurements or from the weather forecasts.

Having now considered the troposphere modeling and the required components, Table 4.11 summarizes the data needed to provide a UE troposphere assistance in terms of delays and mapping functions. The data should be provided for an appropriate area around the UE so that the grid points delivered to the UE cover the troposphere pierce points for the visible SVs.

In addition to data for estimating troposphere delay the troposphere assistance can be used to provide the UE reference altitude information through the use of a barometer. As explained in Chapter 2 having reference altitude reduces the number of SVs required in positioning and can be an essential additional measurement especially in urban environment. Barometer may also be used to stabilize the altitude solution as in [122].

The accuracy of altitude reference should be in the order of 10 meters. In terms of pressure this equals to approximately 120 Pa at sea level. Such accuracy is achievable with commercially available small MEMS-based (Micro-Electromechanical System) barometers as shown in [123]. In the publication the authors find a few Pa absolute differences in the pressure measurements in uncalibrated circumstances and approximately similarly-sized noise together corresponding to less than one meter error in altitude. Moreover, in [122] it was found that temporal gradients may be up to 100 Pa/h meaning that the age of the barometric assistance should not exceed one hour. Alternatively, if numerical weather forecasts are used, also the temporal gradient of pressure can be given. In conclusion, assuming a source for the reference pressure readings and a UE-integrated barometer giving reference altitude using pressure is feasible. The wide-scale uptake of barometers in mobile UEs is, however, subject to size and cost constraints.

The discussion shows that troposphere residual can be modeled and is retrievable through both measurements and numerical weather predictions that can importantly also provide data for an extended period. However, the effect of this residual is so insignificant that the inclusion of troposphere assistance into the standards should only be considered after addressing larger residual error sources including ionosphere and navigation models. Note that this conclusion does not apply to barometric pressure, of which use as a reference altitude improves the availability of GNSS-based positioning.

4.2.5 Radiomaps and fingerprints

The coverage area -based methods were discussed in Chapter 2.3.1 in the context of RANs and network-based positioning methods. The current standards lack the capability to carry coverage area data and, hence, the method is currently limited to either purely network-based or UE-assisted method types.

Coverage areas may be presented either as hard boundaries or as statistical models describing, say, the spatial distribution of the node users [124]. Also, coverage areas need not always be defined explicitly, but such information may also consist of the node location, maximum range, antenna direction and the beam width. Such a database of coverage areas and complementing information is called a radiomap. A fingerprint, on the other hand, is defined as a set that contains the position and radio characteristic records from a variety of radio networks [125]. The characteristics may include time delays, time differences between the wireless communication nodes, channel or signal quality indicators including power histograms, number and spread of rake finger traces and pulse shapes. Moreover, included may be information from multiple antennas (diversity receiver). A fingerprint may be generated by physical measurements, be based on radio channel modeling or their combination.

A fingerprint database is a set of fingerprints in a grid and the database may have wide or even global coverage. Formally, the fingerprint database F may be defined by [126]

$$\begin{aligned} F &= \{F_i \mid i \in \mathbb{F}\} \\ F_i &= \{\underline{p}_i, \{\underline{a}_{ij} \mid j \in N_i\}, \theta_i\}, \end{aligned} \tag{4.16}$$

where \mathbb{F} is the index set for all the database points, F_i the i^{th} fingerprint, \underline{p}_i the fingerprint location and θ_i other recorded information such as the orientation of the recording UE. Finally, \underline{a}_{ij} is the vector for the radio characteristics at the location \underline{p}_i and N_i the set of the communication nodes that is anticipated to be observed at the location \underline{p}_i .

Positioning with the fingerprint database is based on making an observation \underline{y} on the wireless communication nodes and either deterministically or statistically

obtaining a location estimate [125]. Deterministic methods typically involve combining the grid points through

$$\begin{aligned}\underline{x}^{UE} &= \sum_i w_i \underline{p}_i \\ \sum_i w_i &= 1,\end{aligned}\tag{4.17}$$

with w_i being the weight that can, for example, be formed by $w_i^{-1} = \|\underline{y} - \underline{a}_{ij}\|$ assuming the use of L_2 -norm. Other choices may include, for example, the L_∞ -norm. Typically only a subset of database points is considered, in which case the method is called the K-nearest neighbor method.

Probabilistic methods include the Bayesian approach in which one seeks for a solution to the problem

$$p(\underline{x}^{UE}|\underline{y}) = \frac{p(\underline{y}|\underline{x}^{UE})p(\underline{x}^{UE})}{p(\underline{y})},\tag{4.18}$$

where $p(\underline{x}^{UE}|\underline{y})$ is the distribution for the UE position conditioned with the measurement \underline{y} and $p(\underline{y}|\underline{x}^{UE})$ the likelihood of \underline{x}^{UE} given \underline{y} . Finally, $p(\underline{x}^{UE})$ is the prior distribution for the UE location and $p(\underline{y})$ the normalization factor.

While the prior is often taken as uniform distribution, the computation of the likelihood becomes the critical factor. For instance, in [125] the author considers various choices including histograms and Kernel functions to obtain the likelihood. In the histogram approach the likelihood is computed from the similarity between the histograms in the fingerprint database and the observed histogram. Histograms can be treated as distributions for measurements and can be compared using the probability density distance measures. In [127] the authors evaluate various distance measures including the Simandl norm that is found to perform the best in terms of mean error, when studying a case with WLAN APs in an office environment.

Kernel functions estimate the underlying measurement distribution based on the measured data and the histogram is approximated as the linear combination of the Kernel functions. A typical Kernel function is a Gaussian distribution. The

same probability distance measures can be utilized as in the case of histogram comparison. Kernel functions solve the problem of zero observations in a subset of histogram bins due to the incompleteness of the training data. [125]

In [126] the authors find that the histogram comparison benefits more than the Kernel method from a prolonged observation period per each grid point as well as from rotating the UE used in collecting the training data. This is likely due to the property noted previously that the Kernel method solves the problem of no probability mass in certain bins. The authors report 10-s observation period being sufficient. Also, it is observed that the quality of the position estimate does not improve significantly after five WLAN access points. RMS errors ranging from eight to eleven meters are reported for position.

Although the study [126] considered only WLAN and RSS measurements, the same approach is applicable to the cellular RANs and potentially also to various other measurement types presented. Fingerprinting is also called signal-of-opportunity positioning [P7] due to the potential to use any radio signal in positioning.

Radiomaps and fingerprint databases are inherently linked. Both the coverage area and fingerprint data can be in the same database and can be collected at the same time. In fact, coverage areas may be modeled based on the collected/modeled fingerprints. Moreover, in positioning the coverage area -based method may be considered to be computationally less expensive method due to not needing to integrate the distributions as in the fingerprint methods.

As discussed, the current standards lack the capability to carry radiomap information. Radiomaps may be introduced into the standards in terms of shapes specified in 3GPP TS 23.032 [128] including ellipses, ellipsoids or polygons. Regarding the hard boundaries and statistical models it is acknowledged that the matrix definition of an ellipse may also be interpreted as the covariance matrix of a two-dimensional Gaussian distribution. If required, the coverage area may also be defined as a weighted linear combination of several Gaussians. The radio networks included in the radiomap may be, but are not limited to, GERAN, UTRAN, E-UTRAN, WLAN and Bluetooth. The standard must also include a

radiomap request mechanism, mechanism to geographically partition the requests and a version control for delivering updates to the UEs.

Similarly to radiomaps the standards lack the capability to carry fingerprints from the UEs and fingerprint databases to the UEs. The ULP-layer of the SUPL Release 2 includes the possibility to report network measurement results for GERAN, UTRAN, E-UTRAN, CDMA, HRPD (High-Rate Packet Data), UMB, WLAN and WiMAX networks. However, SUPL is not intended for the fingerprint collection and, therefore, future standards should define methods to collect location-associated fingerprint measurements as discussed in [129] for fingerprint database and radiomap collection. Also, the downlink data in the form of a fingerprint database must be defined. As discussed, in the case of signal strength -records (histograms) one feasible option is to define the database in terms of Kernel functions at grid points. Similarly to radiomaps, the standard must also include a database request mechanism, mechanism to geographically partition the requests and a version control for delivering updates to the UEs.

Defining the measurement and database formats is straightforward by considering the NMRs defined for the respective radio networks. Moreover, a positioning method associated with fingerprinting and radiomaps must be defined so that the UE is provided an option for both UE-based (local radiomap or fingerprint database) and UE-assisted (calculation in the server) modes. Finally, regarding the fingerprint collection data may be embedded in the radiomap to instruct the UEs about the collection. The instructions may include information to accelerate, stop or monitor only selected radio access types. The instructions may also be tied to time or location.

One option is to introduce the support for fingerprint collection, fingerprint databases, radiomaps and the associated positioning methods in LPP [130]. This possibility is described in the E-UTRAN LCS Stage 2 [82] stating that

the UE Positioning architecture and functions shall include the option to accommodate several techniques of measurement and processing to ensure evolution to follow changing service requirements and to take advantage of advancing technology.

Although there is no standardized solution for radiomaps and fingerprints, it should be noted that there are various commercially deployed proprietary solutions in the market. Examples include Skyhook Wireless [131], Ekahau [132] and Polaris Wireless [133]. Skyhook Wireless is specialized in WLAN-based positioning and maintains a global database of WLAN access points and their locations. Data is collected both by employees as well as the service users that send back fingerprint data, whenever GPS is on in the UE. Ekahau, on the other hand, is specialized in high-accuracy WLAN-based positioning by using fingerprint-based technology. Their database collection methods include site surveying and radio channel modeling.

Polaris Wireless is a US-based company that provides equipment to the mobile operators for UE positioning. The Polaris technology is embedded in the RAN so that the fingerprint database is always up-to-date, because the positioning engine receives information on the changes in the radio landscape directly from the RAN. Hence, the potentially expensive burden of keeping the fingerprint database up-to-date is eased. In positioning the UE performs a scan and provides the produced fingerprint to the positioning server for position determination.

Radiomaps and fingerprints have widely varying performance depending upon the environment and landscape. Especially regarding radiomaps the performance in terms of accuracy is poor in rural areas, where the cell size is large. On the other hand, in cities with smaller cells and high availability of WLAN access points the performance can be anticipated to be excellent. Fingerprint technologies may provide accurate location given that the database is up-to-date. However, the radio landscape changes, whenever there are changes to the floor plans, new buildings are constructed or radio network topology is modified. There may also be seasonal changes depending on, say, whether there are leaves in trees. Hence, keeping a fingerprint database always up-to-date in a large scale is challenging unless there is an efficient self-calibration mechanism. Such mechanisms include utilizing the AGNSS-enabled UEs in the field to provide measurements to the fingerprint database as well as the database having an access to the radio network planning information. Both radiomap and fingerprint technologies can provide much higher

availability than GNSS the drawback being the potentially unpredictable accuracy and integrity. Integrity issues can, however, be mitigated by increasing the RF diversity and using various radio networks in hybrid. For example, combining cellular- and WLAN-based positioning allows for, for instance, first using cellular methods to verify the coarse location and then refining the location with WLAN. This prevents spoofing by, say, intentionally moving WLAN APs from one location to another.

4.2.6 Support for RTK and PPP

Chapters 2.1.4 and 2.1.5 discussed the absolute (PPP) and relative (RTK) high-accuracy GNSS positioning methods. Whereas PPP requires new assistance data types, RTK needs in addition to those also a protocol supporting continuous exchange of measurements between the UE and the reference entity.

For PPP it was recognized that the mitigation of error sources requires accurate information on the SV orbits, clocks, atmosphere, the SV and UE antennas and on the geophysical phenomena. The variety of orbit and clock models in the standards provides a good starting point for providing high accuracy navigation models to the UEs. However, the current models only include an indication, whether the models originate from the GNSS broadcast or are generated artificially. The standards should, therefore, include a possibility for the UE to specifically request for high accuracy orbits and clocks. Moreover, the models should also include an indication of their quality. The models should, therefore, include orbit and clock degradation models in order to provide the UE the possibility to assess the usability of the model in PPP. The degradation model could, for example, be given in the form of a polynomial to also capture the temporal changes instead of the currently used static URA value (User Range Accuracy).

The types of ionosphere and troposphere models discussed in Chapters 4.2.3 and 4.2.4, respectively, can provide data for PPP given that services that can produce the data exist. Regarding the geophysical phenomena, for example, atmospheric loading can be retrieved from local air pressure [57]. The SV antenna information can also be transferred in the assistance. The SV antenna phase center offsets are

readily available from various sources including [103]. The UE antenna is to be accounted for in the UE itself and is of no concern for the standards. However, as shown in [P3] the UE antenna effects can be significant in terms of induced measurement biases.

It should be noted that PPP requires excellent knowledge of the corrections as well as the UE characteristics (antenna) and is thus probably inappropriate for a mobile UE. Therefore, in [P7] the authors propose introducing light-PPP into the scope of standards in order to provide the UE with high-accuracy AGNSS assistance for increased accuracy. Examples of such assistance include high-accuracy orbit/clock and atmosphere models. These two assistance data types already remove the largest residual error sources and bring the accuracy down to sub-meter level.

Whereas PPP requires only additional assistance data types the new features required for RTK are two-fold. Both new measurement element and protocol features are required. In [P2] and [S2] authors outline the required information elements for the AGNSS standards in order to be able to support high-accuracy relative techniques. The measurement element to be exchanged between the UEs (or alternatively between one UE and the AGNSS server) is described in Table 4.12. In addition to the element, messaging to request and deliver such measurements must be defined. The request must include the possibility to define the GNSSs and signals included in the messages as well as the duration of the request.

Although adding a suitable information element to the standards is simple, the actual challenge is in the protocol carrying the data. As discussed in Chapter 2.1.5 RTK requires that the protocol is capable of relaying a continuous stream of periodic measurements between the reference and rover UEs. However, the current AGNSS standards have been designed for one-shot data delivery and, hence, the current protocol features are not sufficient to support high-accuracy methods. This challenge is addressed in Chapter 4.2.7.

Again, the deployment of full-scale multi-frequency RTK may not be feasible for mobile UEs. The reasons include, firstly, the lack of cost-effective dual-frequency receivers in the near future and, secondly, the challenging signal conditions in the environments, in which the UEs are typically used. Therefore, in [P7] the

authors envision mRTK (mobile RTK) using GNSS receivers connected to the UEs via Bluetooth. Although RTK has traditionally been considered as a professional tool for surveying, in [P1] and [S1] the authors show that it is possible to realize RTK using cheap off-the-shelf Bluetooth GNSS receivers by software additions only. Moreover, in [P3] the authors show that the cheap antennas used in the consumer-grade GNSS devices do not affect the ambiguity resolution performance.

RTK is generally known as accurate relative positioning method, but RTK and PPP can also be seen as complementary technologies in a sense that both can be used to provide the high-accuracy absolute UE position. PPP yields this by default and VRR-services (Virtual Reference Receiver, [P7]) enable the same with RTK. In such a case the RTK signaling is between the UE and the AGNSS server. In VRR-based technologies the UE reports its approximate location to the VRR-server, which generates a computational reference receiver at the reported

Table 4.12: RTK measurement message.

Data	Explanation
Occurs once per measurement message	
Time stamp	Time of the measurement - either in UTC or some GST
Position & uncertainty	Reference position for the measurement
State of movement	Stationary or moving - affects the selection of the state model in the filter
Occurs once per GNSS per SV ID per signal	
GNSS - SV ID - signal	Defines from which GNSS, SV and signal the measurement originates
Code phase & uncertainty	Code phase measurement
ADR & uncertainty	ADR measurement
Cycle slip indicator	Indication of the PLL state over the measurement period. A single bit is adequate because the protocol is assumed to guarantee that no data is lost.

location based on the measurement data the VRR-service obtains from the network of physical GNSS receivers. The baseline is, hence, always short and the ambiguity fixing less problematic the drawback being multiple concurrent connections at the server leading to high server and network load. Moreover, an efficient architecture is required to handle the high number of concurrent computational reference points. The use cases for high-accuracy PVT from PPP or RTK include, for instance, giving lane-precise guidance in car navigation.

PPP and RTK both have their advantages and drawbacks. PPP and RTK both require continuous ADR measurements making power save impossible by turning the GNSS receiver off between the epochs. This reduces battery life. Another drawback of high-accuracy methods is the requirement to have good signal conditions making the availability of high accuracy services low. On the other hand, PPP is advantageous from the point-of-view of low bandwidth consumption. PPP needs high-accuracy assistance once after which the PPP PVT calculation can be commenced. Obviously, the data must be refreshed at the end of its validity period. In contrast, RTK requires a constant stream of measurements from the AGNSS server or another UE. This, again, reduces battery life. Finally, regarding convergence times the fixed RTK baseline may be obtained in seconds, when the baseline is short - which it is, when using a VRR. PPP, on the other hand, takes longer to converge.

The choice between RTK and PPP is dependent upon the application, data channel properties and available assistance data. However, eventually the increasing availability of GNSSs and civilian signals in consumer-grade GNSS devices will drive the technologies now in professional use also to the wider audience. The standardized positioning solutions can support and drive this development.

4.2.7 Assistance stream and IP broadcast

Streaming refers to having a data delivery pipe between two entities including two UEs or between one UE and an AGNSS server. The pipe enables continuous periodic exchange of assistance data between the entities. As discussed in Chapter 4.2.6 such capability is needed for delivering RTK services.

In order to realize the functionality the actual positioning protocol must support requesting measurements on a continuous basis. RRLP and RRC do not support this, but it seems possible to define such capabilities in LPP [130]. Moreover, equally important is that the lower level protocols carrying the positioning protocol support keeping such a data pipe open for continuous exchange of measurements. In [P7] the authors outlined the possibility for the SUPL Release 3 to support such a functionality. In fact, such a proposal has been accepted for SUPL Release 3 Requirement Document [134]. The proposal includes the requirement for SUPL Release 3 to support continuous assistance sessions, but also a requirement for the SETs to be able to exchange their positioning capabilities so that the requesting entity can request information on the GNSS signal support in the other UE. Also, the set of requirements includes a privacy requirement that allows for limiting the duration of the measurement exchange. This is important in case the measurements are exchanged between two SETs.

Broadcasting refers to point-to-multipoint delivery of positioning assistance. As discussed in Chapter 3.2.3.2 in the cellular infrastructure a subset of assistance data may be broadcast to the UEs in the cell. Such a possibility does not exist in the user plane, although majority of, for instance, GNSS assistance data is global by nature. Examples of such global data include SV navigation models including almanacs and global ionosphere models. Hence, it may be argued that establishing a dedicated point-to-point connection to fetch such global data is a waste of assistance server resources. Server loads and bandwidth requirements can, therefore, be eased, if GNSS assistance were broadcast in the user plane.

In [P7] the authors outline several use cases for such broadcast. Delivering well-known globally applicable AGNSS data is one example. Moreover, the broadcast may also be targeted geographically, for instance, based on MCC. Regarding high-accuracy methods also the GNSS reference network measurements are suitable for broadcasting. The measurements from the network stations as well as their locations are broadcast and the UE processes those measurements in the same solver as its own measurements in order to solve the UE position accurately with respect to a set of GNSS network stations. Such an approach has been shown

to produce superior results compared to the single baseline with respect to one physical or virtual reference receiver [135]. Moreover, in Chapter 4.2.6 it was noted that the individual per-UE connections for VRRs may lead to issues with the positioning architecture performance. Broadcasting has the advantage of easing these performance requirements in the service provider side. The SUPL Release 3 requirements include the support for the broadcast of GNSS assistance data [136] in the user plane. One option to realize the functionality is to enable the OMA LCS architecture to provide AGNSS data for distribution over the OMA BCAST [137] broadcast enabler. The enabler includes various useful features from the AGNSS data delivery point-of-view including the possibility to geographically target the broadcast.

Interestingly, the European Commission Framework Project 6 for Satcoms in Support of Transport on European Roads project considers similar aspects for automotive use. In the project technologies to provide Europe-wide RTK corrections via satellite communications have been developed. The goal of the work is to collect GNSS measurements from the reference stations in Europe and distribute the parameterized RTK corrections via a satellite channel to the receivers. Upon the reception of correction data, the receiver reconstructs the standard RTK correction message for further processing in the receiver. The standardized assistance broadcast mechanisms, including those envisioned for SUPL Release 3, could provide a complementary distribution channel for the large-scale data generated for the distribution over the satellite channel. [138]

4.2.8 Assistance data push

Assistance data push refers to the UE having a capability to subscribe to a certain set of assistance data and to the assistance server having a capability to create a connection with the UE for the assistance data delivery. Alternatively, the server may have the capability to notify the UE about the changes in the data so that UE knows to fetch the updated assistance data.

Such capabilities are required in case the assistance data contains information with unpredictable life time and/or unpredictable update rate. Examples include

long-term SV navigation models that may become invalid due to the SV being maneuvered to higher altitude. In this case the data has a well-defined validity period initially, but the data becomes invalid during the period due to the SV maneuver. Another example is atmosphere models that may change abruptly due to, say, ionosphere storm. In the storm conditions, in addition to the model parameters changing abruptly, also the update rate is likely to be higher than in the normal quiet conditions. The problem is that in this case the UE does not know the life time of the data and does not know, when to request for the updates. Yet another use case is pushing RTI information in case of a faulty satellite or signal.

Obviously the UE always has an option to poll the assistance server for the changes either always, when initiating a positioning session or periodically. However, these approaches waste bandwidth and server resources. Therefore, an assistance delivery mechanism that is triggered by an event in the assistance data server is needed. Such an event can be, for example, a change in the assistance data - an update or invalidation. Table 4.13 outlines the requirements for the assistance data push.

The push mechanism has been contributed [139] to the SUPL Release 3 requirements. The method approved into the requirements is called an assistance data change notification. The method includes that only a notification on the assistance data change is delivered to the SET, but not the actual updated assistance data. Having received the notification the SET may decide whether to retrieve the changed data from the SLP or not. The SET might decide not to retrieve the changed data, for example, in the conditions that the SET has moved out of the geographical applicability area of the data of which change notification the SET received. Another example on such conditions is that in some cases it may be sufficient for the SET to simply know that the data it has is invalid and, hence, the SET knows not to use the data.

A potential mechanism for the notification delivery in SUPL is a network-initiated session with the SUPL INIT -message containing the notification on the assistance data change. In the network-initiated case the SLP initiates a session with the SET (via, say, an SMS), the SET establishes a connection with the SLP using

Table 4.13: Requirements for the assistance data push.

Requirement	Explanation
Subscription	The UE shall be able to subscribe to changes in specific assistance data. The server shall keep a database of subscriptions and the subscription parameters.
Duration	The UE shall be able to set the duration for the subscription. The subscription may be perpetual. The assistance server may override the duration requested by the UE due to resource shortage, etc.
Rate	The UE shall be able to set the minimum and maximum update rates. Minimum rate controls how often the assistance server shall at least push a new set of subscribed data to the terminal. Maximum rate controls the minimum interval between pushes. The assistance server may override the duration requested by the UE due to resource shortage, etc.
Modifications	The UE shall be able to modify the subscription parameters and remove the subscription at any time.
Method	The UE shall be able to define, whether it wishes to receive assistance data whenever the data changes or only when the previous data has become invalid. For example, in the case of long-term orbit models the server might receive frequent updates to the models, but the old data might still be valid. However, if an SV is maneuvered in the orbit the old data becomes invalid and the data must be updated to the UE. In addition, the UE should be able to define the desired QOS. Exemplary QOS levels may be that the UE wishes to receive an update ¹ only when the data the UE has becomes invalid, or ² only when the data quality gets below a predetermined threshold, or ³ that the UE wishes to have the latest data (best performance) all the time.

the received session ID and the SET receives the assistance data in the response message from the SLP.

4.2.9 Other issues

The purpose of this chapter is to highlight a few potentially interesting items that may affect the positioning assistance standards at some point. The first such emerging technology is the Japanese IMES (Indoor Messaging System). IMES is based on small transmitters that transmit signals that have the same RF (Radio Frequency) characteristics as GPS L1 C/A signal. IMES has been assigned PRN numbers 173-182. The data payload, however, is different from that in GPS. In fact, IMES transmitters broadcast their position data (latitude, longitude, altitude and floor). Positioning with IMES is simply based on decoding the transmitter location information from the broadcast. [16]

Another technology that has been proposed as a solution to the indoor positioning challenge are pseudolites. Pseudolites, or pseudosatellites, are quite similar to IMES in a sense that they also broadcast GPS-type signal that can be received by GNSS receivers. Pseudolites can be installed indoors, urban canyons etc. to provide position information in environments, in which the actual GNSS SVs cannot be heard. However, the difference is that GNSS receiver is also expected to determine pseudoranges to pseudolites for PVT calculation.

Both IMES and pseudolites can function as complementary information to standardized positioning technologies. Moreover, it may be feasible to include in the assistance data, say, pseudolite positions.

The UE position is useful not only in applications such as navigation, but location can also be used to enforce security and authentication. For example, it might be defined that a certain service is accessible only from the UEs that are within a defined geographical area. In such circumstances the UE or the service to be accessed must be able to authenticate the positioning signal in order to reach certainty that the signal is not being spoofed. Moreover, for example Galileo integrity (positioning performance) guarantees are only valid with the presumption of authentic signals [140]. The military signal in GPS have such an authentication

mechanism via the spreading code encryption [141], which is based on multiplexing the P-code with the secret encryption code (W-code). However, neither civil GPS nor GLONASS signals provide any authentication. Certain Galileo signals are anticipated to provide an authentication mechanism.

In [140] the authors present an authentication method that can be realized without modifications to the GNSS infrastructure. The method is based on taking samples of the L1 P(Y) signal, which is multiplexed in phase quadrature with L1 C/A signal in the GPS SV broadcast. Samples of P(Y) are transferred to the authentication service over the data channel. The service correlates the P(Y) samples with the samples from the reference receiver. Here the reference receiver is assumed to be the trusted party in the authentication. Given that there is correlation it is deduced that the signal originates from an authentic GPS SV. Note that the technology does not require the W-code to be known. The method is, in fact, based on the classified W-code that cannot, therefore, be replicated by the spoofers.

The AGNSS standards might address the issue by providing information elements to carry the necessary information for signal authentication. Although in [140] the authors show that the actual amount of data to be transferred is quite small (4800 bits in their example) the challenge is the wide RF front-end bandwidth (>10 MHz) needed to receive the P(Y) signal. In contrast, the L1 C/A reception requires only approximately \mp 1-MHz bandwidth. Increasing the bandwidth not only increases the power consumption, but also makes the device more susceptible to electromagnetic interference. Moreover, it should be noted that in order to have secure authentication the UE must also be able to authenticate the actual authentication service.

4.3 Costs and Benefits

Figure 4.1 conceptually summarizes the costs and benefits of the various existing and proposed positioning methods. The cost is associated with bandwidth, of which usage consumes network resources, energy in the UE as well as induces expenses

to the end user. Benefits, on the other hand, are associated with availability and accuracy of positioning services.

In the figure the abscissa refers to the positioning methods. Going from left to right the GNSS-based methods (blue set of curves) evolve from basic AGNSS (RRLP-type) to light PPP with advanced navigation and atmosphere models finally to RTK. The radio network -based methods (red set of curves) evolve from simple radiomaps (coverage areas) to TA-/RSS-assisted enhanced radiomap methods finally to fingerprinting with modeled/measured database. The y-axis refers to the qualitative attributes that characterize the positioning methods.

Accuracy. The accuracy line (in solid) indicates the general trend that the more complex the positioning method is the more accurate it is. In the case of GNSS the lowest accuracy (in the order of 10 meters) is achieved either in standalone or assisted mode utilizing broadcast navigation and ionosphere models. The PPP-light already utilizes accurate navigation models and local ionosphere models and can, hence, achieve accuracy in the order of meters. Finally, the RTK-based methods generally achieve sub-decimeter accuracy.

In the radio network methods the lowest accuracy is achieved with coverage area -based methods that yield accuracies in the order of kilometers (worst case). When assisting the coverage area -based methods with, say, TA the accuracy improves to the order of hundreds of meters. Finally, having a fingerprint database with sufficient grid density may provide accuracy in the order of one meter, which is the same order-of-magnitude as with the light PPP.

Bandwidth. The bandwidth line (in dashed) indicates the general trend that in the UE-based methods higher accuracy requires more data traffic. Standalone GNSS position can be obtained without any network interaction and the basic AGNSS assistance over RRLP can be delivered with low bandwidth. However, PPP requires frequent updates of atmosphere models and RTK already requires a stream of reference measurements increasing bandwidth requirements significantly.

The radio network -based methods always require some network interaction for example for obtaining coverage area information or reporting NMR to the network. It should be noted that in UE-assisted control plane conditions the

measurements are part of normal communication-related signaling and, hence, no signaling overhead is induced. However, in the user plane the radio network -based methods necessarily result in bandwidth consumption. Moreover, the accuracy of the fingerprint methods that rely on the grid-based fingerprint databases is dependent on the resolution of the grid. Therefore, assuming the shown accuracy behavior the bandwidth requirement of radio network -based methods increases significantly, when going from radiomaps to fingerprinting.

It should be noted that bandwidth should not be considered as an instantaneous quantity, but instead as an average value. Whereas RTK reference measurements cannot be downloaded and cached to the UE for later use, a fingerprint database can be. Hence, with a database there is a significant peak load, whereas with the RTK measurement stream there is a constant small network load.

Availability. Availability (in dotted) comparison is meaningful only, when considering exclusively areas, where a radio network is hearable. In such a case

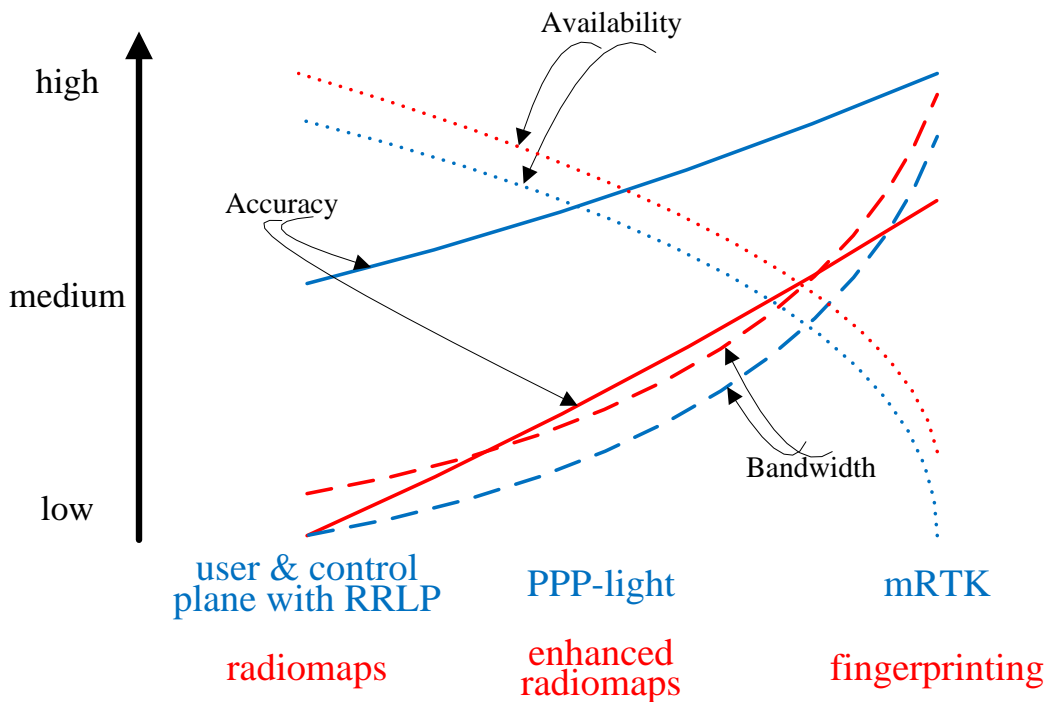


Figure 4.1: Comparison of bandwidth (dashed), accuracy (solid) and availability (dotted) for GNSS-based (blue set of curves) and radio network -based (red set of curves) methods.

assuming the existence of a radiomap, the availability of GNSS-based methods is necessarily lower due to the GNSS having issues indoors and in urban environment. Moreover, the availability of GNSS-based solution further decreases, when going towards methods requiring continuous ADR measurements and, hence, excellent signal conditions.

Finally, the availability of fingerprinting is necessarily lower than that of radiomaps due to fingerprinting requiring diverse RF environment. The diversity is required in order to achieve the improved accuracy over enhanced radiomap (TA-/RSS-assisted) methods as assumed. In the lack of RF diversity fingerprinting approaches enhanced radiomap methods in accuracy. This can be seen by considering a rural environment with a single base station. In such a case fingerprinting cannot outperform the radiomap-based methods. Moreover, fingerprinting will most likely be utilized as a solution for indoors and areas with high population density. The radiomaps can then be used to cover the less densely populated areas at lower cost. This necessarily leads to the radiomap availability being higher.

4.4 Roadmap

The chapter discussing the future needs in location technology standards is concluded by representing a roadmap for various standards and technology features. The evolution is divided into three Volumes, of which two first are quite predictable, but the visibility to the third is already quite low. The Volume One consists of items that practically exist already. These are RRLP Release 8 with GPS and GANSS support complemented by the SUPL Releases 1 and 2 that can carry RRLP in the user plane. The SUPL Release 2 can take full advantage of the RRLP Release 8 features by supporting GANSS. Note that the life cycle of Volume 1 is not expected to end in the scope of this roadmap due to the high number of UEs supporting Volume 1 technologies.

The Volume Two is quite well roadmapped in terms of schedules. The LPP Release 9 will be completed early 2010 and the SUPL Release 3 is scheduled to be

completed by the beginning of 2012 [142]. The LPP extensions discussed in Chapter 3.2.3.4 (possibility by other fora to extend LPP) are roadmapped to mid-2011 [130]. This is feasible with the precondition that the work on LPP extensions in OMA starts early 2010. It should be noted that the actual content of the extension is to be decided. Moreover, there will very likely be several releases that successively add new features. The roadmap shows that LPP is truly becoming the positioning technology protocol for both future RANs and user plane solutions.

The Volume Three includes items still in visionary discussions. One such issue is the signal authentication discussed in Chapter 4.2.9. Authentication aspects might, for instance, be considered as further extensions to LPP. Another intriguing possibility for future development is the peer-to-peer assistance, which refers to distributing positioning assistance in the community or exchanging such data with nearby UEs. However, it should be noted that this is not just a matter of location technology protocols, but also involves aspects including privacy and establishing such peer-to-peer connection.

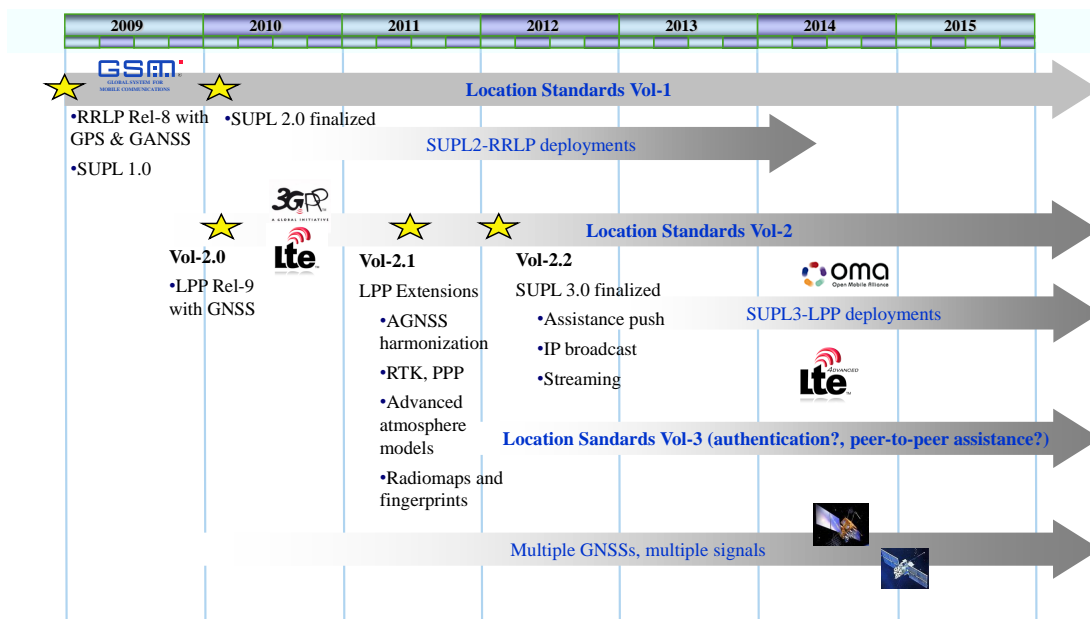


Figure 4.2: Location technology standards roadmap.

Chapter 5

Contribution

5.1 Author's contribution to the publications

This chapter summarizes the publications that are represented as a part of this thesis. The co-authors have seen and approved these descriptions on the publications and on the author's contribution to the publications.

Publication [P1]: The publication introduces the concept of mRTK (mobile Real-Time Kinematic). Whereas previously RTK techniques have only been considered for professional purposes (construction, surveying), mRTK utilizes the carrier phase -measurement capabilities readily available in consumer-grade GNSS devices. mRTK is also characterized by the use of inertial sensors to distinguish between static and dynamic cases. Moreover, positioning standards -related considerations are given.

The key finding is that a cheap, consumer-grade Bluetooth AGPS-device is usable in RTK with software-only additions, when the baselines are fairly short. This is not seen as a drawback, because the concept is envisioned to be used with Virtual Reference Receiver -services that allow for subscribing an artificial reference receiver at the UE location keeping the baseline always short.

In the work presented in [P1] the author has been responsible for implementing the RTK package in an UE. The author has also performed the measurements

using a GPS simulator and done the data analysis used to verify the feasibility of the concept. Finally, the author has contributed to the R&D GNSS assistance protocol implemented during the work.

Publication [P2]: The publication analyzes the feasibility of adding carrier phase -assistance to the cellular standards. The situation is analyzed in the context of RRLP and RRC. Although RRLP and RRC do not include suitable messages for carrying measurements, the feasibility is discussed in the context of protocol capabilities. Whereas RRLP allows for requesting/sending three sets of measurement data, RRC allows for non-continuous periodic reporting.

The key findings include that as expected the probability of the successful ambiguity resolution does increase, when going from the RRLP-type functionality to the RRC-type periodic reporting. It is also shown that the greater availability of signals contributes more than the number of measurement sets. The signals may originate from different GNSSs or from the same GNSS providing various signals. The paper shows that the GPS L1 + Galileo E1 and Galileo E1 + Galileo E5a combinations both have excellent performance in terms of probability of successful ambiguity fix. The high number of SVs in the case of GPS L1 + Galileo E1 further contributes to excellent day-around performance even though only single frequency is utilized.

In the work the author has been responsible for performing the probability simulations using the VISUAL [143] package and analyzing the results. Moreover, the author has also done the analysis of the missing components in RRLP and RRC as well as specified the items that should be added to the cellular standards in order to enable carrier phase -based positioning.

Publication [P3]: This publication studies how the antenna characteristics may hinder the RTK performance in cheap, consumer-grade GNSS devices. One surveying-grade and two consumer-grade antennas are measured using the Satimo [144] antenna measurement system providing a complex (phase and magnitude) antenna response in 3D.

The paper concentrates on the phase response and on its spatial anisotropy. The phase response is of great interest, because carrier phase measurements are the primary measurements in RTK. Measurements show that whereas the surveying-grade antenna has excellent, almost isotropic phase response, the cheap patch antennas in consumer-grade devices may induce up to ∓ 75 -degree biases in the phase measurements depending upon the direction of the signal arrival.

The study proceeds to consider, whether such antenna effects, if not mitigated, hinder the RTK performance. The paper shows that the ambiguity resolution is not widely affected. However, the study leaves open the effects on the baseline accuracy. Moreover, the paper proposes that near-isotropic phase response should be considered as one requirement for antenna performance in case consumer-RTK is commercialized.

In the work the author has been responsible for assisting in the antenna measurements. Moreover, the author has performed the data analysis and simulations to show the negligible effects on the ambiguity resolution.

Publication [P4]: The publication introduces the concept of Assisted GNSS and outlines the large scale structure for GNSS assistance data elements for cellular standards. The paper discusses the future changes in the GNSS landscape and the need for updating the assistance standards. The publication further discusses the opportunities of AGNSS, the benefits to the customers and the requirements for the future assistance standards. The paper then proceeds to outlining the common and generic assistance elements as discussed in Chapter 3.2.3.1.

The upcoming Release 8 of RRLP and RRC are well aligned with the structure outlined in the paper. In the assistance part the RRLP Release 8 is only missing the troposphere model and base station timing models. In the uplink part the differences include the lack of capability to report base station timing measurements from the UE. Further minor differences can also be found, but on the large scale the outlined structure has well been introduced into the cellular assistance standards. Similar structure will at least partly be specified for LPP.

The author has been a part of the team brainstorming the structure and has been active in 3GPP and OMA standardization fora to promote and specify the outlined structure.

Publication [P5]: This paper presents the key ideas behind the multi-mode navigation model now found in RRLP and RRC. The same structure has also been proposed for LPP. The work reviews the orbit and clock correction models in GPS, GLONASS, QZSS, SBAS and Galileo. The special considerations for pseudolites are also given.

The paper acknowledges that the orbit models may be divided into three categories: Keplerian model (GPS, Galileo), high-accuracy Keplerian model (modernized GPS, QZSS) and Cartesian coordinate representation (SBAS, GLONASS). The SV clock models on the hand can be divided into two categories - standard and high-accuracy the difference being that high-accuracy model is used to represent corrections for very stable clocks at high resolution but low dynamic range. These three orbit and two SV clock parameterizations are used as a baseline for outlining a proposal for the multi-mode navigation model for cellular assistance standards.

Because of the bandwidth issues in the control plane the study also seeks for bit saving opportunities. The paper shows that in some cases bandwidth may be saved by taking the most significant bits of certain parameters as common for all the SVs for which assistance is delivered. The paper further discusses the potential of the multi-mode navigation model that decouples orbit and SV clock models as well as decouples navigation model parameterizations from the GNSSs. This allows, in principle, for using any combination of orbit and SV clock model for any GNSS introducing flexibility.

The navigation model structure now found in RRLP and RRC follows the principles outlined in the study. Furthermore, the navigation model in LPP will contain many of the ideas presented in the paper including decoupling the navigation model parameterization from the GNSS.

In the work the author has been responsible for looking for the commonalities in the orbit and SV clock representations as well as defining the multi-mode structure for the cellular standards. The author was active in driving the multi-mode navigation model to RRLP in 3GPP and the author was part of the team writing the change request to RRLP in order to introduce the model into the standards.

Publication [P6]: This study discusses one specific aspect of the multi-mode navigation model - the possibility to utilize non-native parameterizations for orbit and clock models. Specifically, the study looks into the prospects of providing navigation model assistance to GLONASS SVs in the format native to legacy GPS.

The study is based on analyzing the GLONASS orbits over a period of time and fitting the orbit data to the Quasi-Keplerian GPS NAV parameterization in 4-h blocks. The hypothesis is that because GLONASS and GPS orbits resemble each other, the GPS NAV parameterization is suitable also for GLONASS SVs.

The analysis shows that GPS NAV parameterization can be used for GLONASS. Moreover, the SV clock correction model can also be represented with the parameterization found in GPS NAV. The feasibility of the concept potentially simplifies the positioning engine implementation, because the study shows that the navigation model data for at least GPS and GLONASS can be provided in the same format. The finding is highly significant also because it is expected that in the short-term the assistance data originates from the GNSS reference network services predicting orbit and clock behavior several weeks ahead. Using the same orbit and clock parameterization for all the GNSSs simplifies the implementation in the data provider end as well.

In the work the author has been responsible for fitting the GLONASS SV orbit and clock correction data to the desired parameterization and studying the results in order to show the feasibility of the concept.

Publication [P7]: The paper is a discussion about the current status of the positioning protocols in various RANs, shortcomings of the current standards as well as trends in the positioning and LBS domains. This discussion allows for

outlining the requirements for the new location and positioning protocol.

The authors show that the current standards lack both important data content as well as protocol capabilities. Missing data types include atmosphere models, radiomaps and fingerprint databases. In the protocol domain the continuous assistance session for assistance data streaming is seen as an important feature.

In the work the author has been responsible for drafting the requirements for the data content as well as for the protocol. The author has also envisioned an exemplary implementation for the future location and positioning protocol.

5.2 Author's contribution to the positioning standards

The author has been active in 3GPP GERAN WG2 in 2006-07 and in OMA LOC Working Group since the beginning of 2008. The author also contributes to LPP.

In 3GPP the author was especially dealing with the multi-mode navigation model, of which aspects are discussed in [P5] and [P6]. The author contributed the report [S3] to 3GPP GERAN WG2 justifying the need for the multi-mode model. Many of the aspects proposed by the author are now in RRLP, RRC and very likely in LPP as well. In addition, author contributed to the Ephemeris Extension discussion and specification in 3GPP GERAN WG2 [145]. Ephemeris Extension in RRLP is discussed in Chapter 3.2.3.1.

In 3GPP and OMA the author has been actively promoting new positioning methods and assistance data types. Examples include proposing carrier phase-based high-accuracy positioning in 3GPP [146] and discussing the benefits of various new AGNSS assistance data formats in OMA [147]. The author was also driving a work towards a new positioning protocol in OMA in 2008 [148].

The author actively contributes to Secure User Plane Location protocol in OMA LOG WG. Author has contributed several items to the SUPL Release 3 requirements such as the requirement for streaming positioning and location data between entities [134] as well as assistance data push [139]. In addition, the author is also contributing to the LPP extensions work in OMA [130].

Chapter 6

Conclusions

The positioning infrastructure and protocol standardization in wireless telecommunication networks stems from various considerations. The drivers range from legislative reasons, such as emergency call positioning and lawful interception, to commercial demand in the form of location-based services and navigation. Wireless networks provide a variety of options for the UE positioning. These options include both the positioning technologies realized natively in the networks as well as the positioning assistance data distributed over the networks.

The cellular network -native positioning technologies include coverage area -based, signal strength -based and timing-based methods. The accuracy of these methods is shown to be in the order of hundreds of meters. On the other hand, UEs equipped with Assisted GNSS receivers can benefit from the assistance data obtainable in the wireless networks. Receiving navigation models, time, frequency and data bit sequences from the network is shown to improve both TTFF as well as the sensitivity of the AGNSS receiver.

Previously only the legacy GPS L1 C/A signal was supported in the AGPS standards. During the course of the work the positioning standards were successfully updated to accommodate the new GNSSs. The study shows how the new GNSSs were successfully introduced into the AGNSS standards in a generic format that supports straightforward addition of new GNSSs in the future. This is important due to the GNSS landscape still being in constant change. One such

foreseeable change will be the introduction of GLONASS CDMA signals, which will further increase commercial interest towards GLONASS.

Considering the GNSS-based methods the current positioning standards only support the items the GNSSs natively provide in their broadcasts. However, the broadcast items do not allow for taking advantage of the full potential of the GNSS-based methods. From this perspective the standards are shown to lack both in data content and protocol features. The data content additions that should be addressed in the future standards include at least local atmosphere models and improved navigation models that together can bring the AGNSS positioning accuracy to the order of one meter. Such accuracy is enabled by bringing Precise Point Positioning into the scope of wireless positioning standards. The missing protocol features, on the other hand, include continuous periodic delivery of assistance data for high accuracy methods including RTK that can provide centimeter-level accuracy. RTK corrections as well as many other assistance data types are also shown to be suitable for broadcast over IP broadcast enablers.

The recap of the current standards shows that they also lack the support for UE-based radiomap and fingerprint positioning methods. The standards lack both the capability to collect fingerprints and radiomap data as well as to transfer such databases to the UEs. Radiomaps and fingerprinting will drive availability in contrast to the new GNSS assistance data types primarily driving the accuracy aspect. The radiomap and fingerprinting technologies can in principle provide position, whenever a network is available. However, this requires the existence of a database with suitable records on the communication nodes in the network. With appropriate additions to the standards the database maintenance may be simplified by enabling fingerprint collection using the terminals in the field.

The future positioning technology improvements will be implemented in the user plane. There are four reasons to this. Firstly, only a user plane solution can be truly global as required of the present-day solutions. Secondly, the new features are potentially data hungry making them inherently unsuitable for control plane. The discussion shows that, in general, accuracy and bandwidth correlate - higher accuracy requires more assistance data at a higher update rate. An example of

such high-bandwidth technology is RTK requiring near-realtime continuous data exchange between the entities. Thirdly, as the driving use case for the control plane positioning is emergency call positioning, for example, the advanced atmosphere models are unnecessary in the control plane. Finally, radiomaps and fingerprinting as multi-RAN positioning technologies can be considered to be suitable only for the user plane implementations. Therefore, OMA SUPL and OMA-extended 3GPP LPP are seen as the most promising protocols to realize the discussed features for the future positioning standards.

The actual protocol additions required to introduce the new assistance data types are typically straightforward and seem quite innocent most of the time. However, the complexity and challenges are in fact in the data provider side and are two-fold. One challenge is running the architecture required to provide assistance data for potentially hundreds of millions of UEs. Thus the broadcast of assistance data over IP will be of growing importance in the future, even though the uptake of broadcast in the control plane domain has been limited. The other challenge is the availability of the data. The reliability and integrity of wide-area RTK data as well as the availability of both troposphere and ionosphere predictions play key roles in the AGNSS revolution. Finally, the management and scalability of fingerprint and radiomap databases is yet another challenge that the service providers will need to consider. The role of the standardization activity is both to produce key enablers for the wide-scale uptake of new technologies as well as to signal commercial interest and thus encourage R&D investments in the area.

The activities in the various standardization fora, including OMA and 3GPP, show that the envisioned developments are achievable in the foreseeable future driving the location technology convergence between the control and user planes. Such development reduces the needs for proprietary deployments and duplicate implementations due to the control-user plane divergence, brings added value to the consumers via improved accuracy and availability as well as reduces development costs. This convergence is a major achievement in the LCS standardization.

Bibliography

- [1] Canalys: Worldwide navigation market overview Q2 2009, 18th August 2009.
- [2] Canalys: EMEA mobile navigation market overview Q2 2009, 7th August 2009.
- [3] Canalys: EMEA mobile navigation market overview Q3 2009, 10th November 2009.
- [4] The 3rd Generation Partnership Project, <http://www.3gpp.org>.
- [5] Open Mobile Alliance, <http://www.openmobilealliance.org>.
- [6] OMA-TS-ULP-V1_0-20070615-A, User Plane Location Protocol, Release 1.0, <http://www.openmobilealliance.org>.
- [7] OMA-TS-ULP-V2_0-20090630-D, User Plane Location Protocol, Candidate Enabler Release 2.0, <http://www.openmobilealliance.org>.
- [8] 3GPP TS 23.271 Functional Stage 2 Description of Location Services (LCS), <http://www.3gpp.org>.
- [9] Federal Communications Commission, <http://www.fcc.gov>.
- [10] Wireless E911 Location Accuracy Requirements, PS Docket No. 07-114, Federal Communications Commission, October 2001.
- [11] Report and Order, PS Docket No. 07-114, Federal Communications Commission, 20th November 2007.

- [12] Canals: Mobile navigation and location market trends 2009/2010, 17th July 2009.
- [13] IS-GPS-200 Revision C, IRN-200C-003: Navstar GPS Space Segment/Navigation User Interface, 11th October 1999.
- [14] Galileo Open Service, Signal In Space Interface Control Document, Draft 1, 1st February 2008.
- [15] GLONASS interface document version 5.0. Moscow, 2002.
- [16] IS-QZSS ver 1.0, Japan Aerospace Exploration Agency, 30th November 2007.
- [17] Alliance for Telecommunications Industry Solutions, <http://www.atis.org>.
- [18] Telecommunications Industry Association, <http://www.tiaonline.org/>.
- [19] 3GPP TS 04.31, Radio Resource LCS (Location Services) Protocol (RRLP), <http://www.3gpp.org>.
- [20] 3GPP TS 44.031, Radio Resource LCS (Location Services) Protocol (RRLP), <http://www.3gpp.org>.
- [21] Meeting report of TSG GERAN meeting#32, Sophia-Antipolis, France, 13th-17th November 2006, <http://www.3gpp.org>.
- [22] IS-GPS-200 Revision D, IRN-200D-001: Navstar GPS Space Segment/Navigation User Interfaces, 7th March 2006.
- [23] IS-GPS-705 draft Navstar GPS Space Segment/User Segment L5 Interfaces, 22nd September 2005.
- [24] IS-GPS-800 draft Navstar GPS Space Segment/User Segment L1C Interfaces, 19th April 2006.
- [25] Specification for the Wide Area Augmentation System (WAAS), FAA-E-2892b, 13th August 2001.

- [26] Meeting report of 3GPP TSG RAN meeting #36, 29th May - 4th June 2007, Busan, Korea.
- [27] Open Mobile Alliance Location Working Group meeting minutes OMA-LOC-2007-0290-MINUTES_20Aug2007Seoul, Seoul, Korea, 20th-24th August 2007, <http://www.openmobilealliance.org>.
- [28] Internet Engineering Task Force, <http://www.ietf.org>.
- [29] IETF GEOPRIV WG, HTTP Enabled Location Delivery (HELD), Internet-Draft, <http://www.ietf.org>.
- [30] TSG-RAN Working Group 2 document number R2-094121. Draft Report of 3GPP TSG RAN WG2 meeting #66bis, June 29th - July 3rd 2009, Los Angeles, USA.
- [31] E. Kaplan and C. Hegarty, *Understanding GPS Principles and Applications*, 3rd ed. Artech house, Inc. Norwood, MA, USA, 2006.
- [32] "GPS III Passes Preliminary Design Review," *InsideGNSS*, 2009, 21st May.
- [33] "GPS Satellite with L5 Payload Launches Successfully," *InsideGNSS*, 2009, 24th March.
- [34] "Galileo Will No Longer Require Special Manufacturer's License," *GPS World*, 2009, 28th October.
- [35] "Modernization Milestone," *InsideGNSS*, 2009, May/June issue.
- [36] S. Revnivykh, "GLONASS Status Update," in *46th CGSIC Meeting Fort Worth, Texas, USA, September 26th, 2006*.
- [37] "Russia Launches Three More GLONASS-M Space Vehicle," *InsideGNSS*, 2009, 25th December.
- [38] "Satellite system GLONASS Status and Plans," Roskosmos, Tech. Rep., 2009, presented in ICG WG A: Interoperability and compatibility, 30th-31st July, Vienna, Austria.

- [39] “Russia Approves CDMA Signals for GLONASS, Discussing Common Signal Design,” *InsideGNSS*, 2008, 8th April.
- [40] “Russia Building Out GLONASS Monitoring Network, Augmentation System,” *InsideGNSS*, 2009, 19th August.
- [41] “Program Updates,” *InsideGNSS*, 2009, January/February issue.
- [42] “ESA Signs Galileo IOV Launch, FOC Satellite Contracts,” *InsideGNSS*, 2009, 16th June.
- [43] “NOAA, Japan Establish QZSS Ground Station in Guam,” *InsideGNSS*, 2009, 27th August.
- [44] R. Leopold and D. Diekelman, “Geosynchronous satellite communication system and method,” Patent US6 226 493, May, 1996.
- [45] Signal Transmission by MSAS, KOBE Aeronautical Civil Center, 27th September 2007.
- [46] K. Bandyopadhyay, “GSAT-4, a Step Towards Indian Advanced Communications Satellite,” in *Proceedings of International Workshop on Satellite and Space Communications 2008, Toulouse, France, 1st-3rd October, 2008*.
- [47] “EGNOS Fact Sheet 14: SBAS - Interoperability explained - Delivering a global services,” European Space Agency, Tech. Rep., 2005.
- [48] “China Adds Details to Compass (Beidou II) Signal Plans,” *InsideGNSS*, 2008, September/October issue.
- [49] “COMPASS View on Compatibility and Interoperability,” China National Administration of GNSS and Applications, Tech. Rep., 2009, presented in ICG WG A: Interoperability and compatibility, 30th-31st July, Vienna, Austria.

- [50] “China’s Regional Compass System: 12 Satellites in 2012 - ICD Next Year?” *InsideGNSS*, 2009, 5th October.
- [51] “China Launches Second Compass (Beidou-2) Satellite in \$1.46 Billion First Phase,” *InsideGNSS*, 2009, 18th April.
- [52] “China Launches Compass MEO,” *InsideGNSS*, 2007, Spring issue.
- [53] G. Gao, A. Chen, S. Lo, D. Lorenzo, and P. Enge, “GNSS over China - The Compass MEO Satellite Codes,” *InsideGNSS*, pp. 36–43, 2007, July/August issue.
- [54] A. Bhaskaranarayana, “Indian IRNSS and GAGAN,” Indian Space Research Organization, Tech. Rep., 2008, presented in COSPAR meeting, 15th July, Montreal, Canada.
- [55] J. Syrjärinne, “Studies of Modern Techniques for Personal Positioning,” Ph.D. dissertation, Tampere University of Technology, 2001.
- [56] P. Misra and P. Enge, *Global Positioning System - Signals, Measurement and Performance*, 2nd ed. Ganga-Jamuna Press, 2006, iISBN 0-9709544-1-7.
- [57] B. Witchayangkoon, “Elements of GPS precise point positioning,” Ph.D. dissertation, University of Maine, 1997.
- [58] A. Le, “Achieving Decimetre Accuracy with Single Frequency Standalone GPS Positioning,” in *Proceedings of the ION GNSS 17th International Technical Meeting of the Satellite Division, 21st-24th September, Long Beach, CA, USA*, 2004, pp. 1881–1892.
- [59] J. Dow and R. Neilan, “The International GPS Service (IGS): Celebrating the 10th Anniversary and Looking to the Next Decade,” *Advanced in Space Research*, vol. 36, no. 3, pp. 320–326, 2005.
- [60] C. Tiberius and P. Jonge, “Fast Positioning Using the LAMBDA-Method,” in *Proceedings of the 4th International Symposium on Differential Satellite Navigation Systems (DSNS), 24th-28th April, Bergen, Norway*, 1995, pp. 1–8.

- [61] M. Monnerat, "AGNSS Standardization - The Path to Success in Location-Based Services," *InsideGNSS*, pp. 22–33, 2008, July/August issue.
- [62] F. Diggelen, *A-GPS : Assisted GPS, GNSS and SBAS*. Artech House, 2009.
- [63] D. Lundgren and F. Diggelen, "Long-Term Orbit Technology for Cell Phones, PDAs," *GPS World*, pp. 32–36, 2005, October issue.
- [64] 3GPP TS 25.171, Requirements for support of Assisted Global Positioning System (A-GPS), <http://www.3gpp.org>.
- [65] S. Turunen, "Acquisition Performance of Assisted and Unassisted GNSS Receivers with New Satellite Signal," in *Proceedings of Institute of Navigation, 25th-28th September, Forth Worth, USA, 2007*, pp. 211–218.
- [66] Institute of Electrical and Electronics Engineers, <http://www.ieee.org>.
- [67] M. Spirito, S. Pöykkö, and O. Knuutila, "Experimental performance of methods to estimate the location of legacy handsets in GSM," in *Proceedings of IEEE 54th Vehicular technology conference VTC, October 7-11, Atlantic City, NJ, USA, 2001*.
- [68] 3GPP TS 45.010, Radio subsystem synchronization, <http://www.3gpp.org>.
- [69] 3GPP TS 25.133, Requirements for support of radio resource management (FDD), <http://www.3gpp.org>.
- [70] 3GPP TS 25.201, Physical layer - General description, <http://www.3gpp.org>.
- [71] 3GPP TS 25.211, Evolved Universal Terrestrial Radio Access (E-UTRA), Physical channels and modulation, <http://www.3gpp.org>.
- [72] 3GPP TS 36.213, Evolved Universal Terrestrial Radio Access (E-UTRA), Physical layer procedures, <http://www.3gpp.org>.
- [73] 3GPP TS 25.215, Physical layer - Measurements (FDD), <http://www.3gpp.org>.

- [74] S. Poykko and M. Spirito, "Method and system for estimating the position of a mobile device," Patent EP1 552 321, July, 2005.
- [75] Hata, M., "Experimental Performance to Estimate the Location of Legacy Handsets in GSM," *IEEE Transactions on Vehicular Technology*, vol. 29, no. 3, August 1980.
- [76] D. Porcino, "Performance of a OTDOA-IPDL positioning receiver for 3GPP-FDD mode," in *Second International Conference on 3G Mobile Communication Technologies*, 2001.
- [77] P. Hansen and P. Duffett-Smith, "Accurate Time Assistance," *GPS World*, 2005, 1st July.
- [78] 3GPP TS 22.071 Location Services (LCS); Stage 1 Service Description, <http://www.3gpp.org>.
- [79] 3GPP TS 44.071 Mobile radio interface layer 3 LCS specification, <http://www.3gpp.org>.
- [80] 3GPP TS 43.059 Functional Stage 2 Description of Location Services (LCS) in GERAN, <http://www.3gpp.org>.
- [81] 3GPP TS 25.305 Stage 2 Functional Description of User Equipment (UE) Positioning in UTRAN, <http://www.3gpp.org>.
- [82] 3GPP TS 36.305 Stage 2 functional specification of User Equipment (UE) positioning in E-UTRAN, <http://www.3gpp.org>.
- [83] 3GPP TS 49.031 Base Station System Application Part LCS Extension (BSSAP-LE), <http://www.3gpp.org>.
- [84] 3GPP TS 24.080, Mobile radio interface layer 3 supplementary services specification, Formats and coding, <http://www.3gpp.org>.

- [85] “Comparison of OTDOA and RTT Location Methods in LTE,” Polaris Wireless, Tech. Rep. R2-094013, 2009, submitted to 3GPP RAN2 66bis, June 29th - July 3rd, Los Angeles, USA.
- [86] 3GPP TS 25.331, Radio Resource Control (RRC) protocol specification, <http://www.3gpp.org>.
- [87] 3GPP TS 25.225, Physical layer Measurements (TDD), <http://www.3gpp.org>.
- [88] “Reuse of RRLP assistance data IEs in LPP protocol,” Nokia Inc., Nokia Siemens Networks, Tech. Rep. R2-094683, 2009, presented in 3GPP RAN2 meeting, 24th-28th August, Shenzhen, China.
- [89] “Text proposal for TS 36.355 A-GNSS material,” Qualcomm Europe, Tech. Rep. R2-095771, 2009, submitted to 3GPP RAN2 67bis, October 12th - 16th, Miyazaki, Japan.
- [90] 3GPP TS 45.005, Radio transmission and reception, <http://www.3gpp.org>.
- [91] L. Wirola, J. Syrjärinne, and K. Alanen, “Methods and apparatuses for assisted navigation systems,” Patent WO 2007/099196, February, 2006, in application phase.
- [92] 3GPP TS 44.035, Broadcast network assistance for Enhanced Observed Time Difference (E-OTD) and Global Positioning System (GPS) positioning methods, <http://www.3gpp.org>.
- [93] 3GPP TS 23.301, Radio Interface Protocol Architecture, <http://www.3gpp.org>.
- [94] 3GPP TS 25.214, Physical layer procedures (FDD), <http://www.3gpp.org>.
- [95] 3GPP TS 36.355 LTE Positioning Protocol (LPP), <http://www.3gpp.org>.

- [96] L. Wirola, "RRLP shortcomings," Nokia Inc., Tech. Rep. OMA-LOC-2008-0385, 2008, presented in OMA LOC WG meeting, 19th-21st August, Chicago, USA.
- [97] "Initial proposed contents for stage 3 LPP specification," Qualcomm, Tech. Rep. R2-094407, 2009, presented in 3GPP RAN2 meeting, 24th-28th August, Shenzhen, China.
- [98] 3GPP TS 51.010-1, Mobile Station (MS) conformance specification; Part 1: Conformance specification, <http://www.3gpp.org>.
- [99] 3GPP TS 34.171, Terminal conformance specification; Assisted Global Positioning System (A-GPS), <http://www.3gpp.org>.
- [100] 3GPP2-C.S0022-A Position determination service for CDMA2000 Spread Spectrum Systems, <http://www.3gpp2.org>.
- [101] OMA-WID_0183-SUPL3_V1_0-20090114, Work Item Document, Secure User Plane Location Enabler Release (SUPL) V3.0, <http://www.openmobilealliance.org>.
- [102] OMA-RD-SUPL-V3_0-20090702-D, Secure User Plane Location Requirements, Draft Version 3.0, 2nd July 2009, <http://www.openmobilealliance.org>.
- [103] Center for Orbit Determination in Europe.
- [104] 3GPP TS 45.001, Physical layer on the radio path; General description, <http://www.3gpp.org>.
- [105] 3GPP TS 25.211, Physical channels and mapping of transport channels onto physical channels (FDD), <http://www.3gpp.org>.
- [106] 3GPP TS 36.331, Evolved Universal Terrestrial Radio Access (E-UTRA) Radio Resource Control (RRC), <http://www.3gpp.org>.

- [107] S. Hilla, “The Extended Standard Product 3 Orbit Format (SP3-c),” National Geodetic Survey, Tech. Rep., 2007.
- [108] M. Schenewerk, “A brief review of basic GPS orbit interpolation strategies,” *GPS Solutions*, vol. 6, no. 4, pp. 265–267, 2003.
- [109] “Alternative for E-SMLC capability indication,” Nokia Inc., Nokia Siemens Networks, Tech. Rep. R2-094681, 2009, presented in 3GPP RAN2 meeting, 24th-28th August, Shenzhen, China.
- [110] S. Skone, R. Yousuf, and A. Coster, “Performance Evaluation of the Wide Area Augmentation System for Ionospheric Storm Events,” *Journal of Global Positioning Systems*, vol. 3, no. 1-2, pp. 251–258, 2004.
- [111] A. Leick, *GPS Satellite Surveying*, 3rd ed. John Wiley & Sons, 2004, iSBN 0-471-05930-7.
- [112] P. Kinter, T. Humphreys, and J. Hinks, “GNSS and Ionospheric Scintillation,” *InsideGNSS*, pp. 22–30, 2009, July/August issue.
- [113] S. Stankov, I. Kutiev, N. Jakowski, and A. Wehrenpfennig, “A New Method for Total Electron Content Forecasting Using Global Positioning System Measurements,” in *Proceedings of Space Weather Workshop: Looking Towards a European Space Weather Programme, 17th-19th December, Noordwijk, The Netherlands*, 2001.
- [114] N. Jakowski, E. Sardon, and S. Schlüter, “GPS-Based TEC Observations in Comparison with IRI95 and the European TEC model NTCM2,” *Advances in Space Research*, vol. 22, no. 6, pp. 803–806, 1998.
- [115] N. Jakowski, E. Sardon, A. Jungstand, and D. Klähn, “Relationship between GPS-signal propagation errors and EISCAT observations,” *Annales Geophysicae*, vol. 14, pp. 1429–1436, 1996.

- [116] N. Jakowski, S. Stankov, and D. Klähn, “Operational space weather service for GNSS precise positioning,” *Annales Geophysicae*, vol. 23, pp. 3071–3079, 2005.
- [117] C. Stolle, S. Schlüter, C. Jacobi, and N. Jakowski, “3-dimensional ionospheric electron density reconstruction based on GPS measurements,” *Advances in Space Research*, vol. 31, no. 8, pp. 1965–1970, 2003.
- [118] Y. Zheng, Y. Feng, and Z. Bai, “Grid Residual Tropospheric Corrections for Improved Differential GPS Positioning Over the Victoria GPS Network (GPSnet),” in *Proceedings of the International Symposium on GNSS/GPS, 5th-8th December, Sydney, Australia, 2004*.
- [119] M. Noordman, R. Eresmaa, M. Poutanen, H. Järvinen, H. Koivula, and J.-P. Luntama, “Using Numerical Weather Prediction Model Derived Tropospheric Slant Delays in GPS Processing: a Case Study,” *Geophysica*, no. 43, pp. 49–57, 2007.
- [120] J. Boehm and H. Schuh, “Vienna Mapping Functions,” in *Proceedings of the 16th European VLBI Meeting*, 2003, pp. 131–143.
- [121] B. Stoyanov, R. Haas, and L. Gadinarsky, “Calculating Mapping Functions from the HIRLAM Numerical Weather Prediction Model,” in *Proceedings of the IVS General Meeting, 9th-11th February, Ottawa, Canada, 2004*, pp. 471–475.
- [122] K. Alanen and J. Käppi, “Enhanced Assisted Barometric Altimeter A-GPS Hybrid Using the Internet,” in *Proceedings of Institute of Navigation, 13th-16th September, Long Beach, USA, 2005*, pp. 1991–1997.
- [123] J. Parviainen, J. Kantoja, and J. Collin, “Differential Barometry in Personal Navigation,” in *Proceedings of IEEE/ION PLANS 2008, 6th-8th May, Monterey, CA, USA, 2008*, pp. 148–152.
- [124] L. Koski, R. Piché, V. Kaseva, S. Ali-Löytty, and M. Hännikäinen, “Positioning with coverage area estimates generated from location

- fingerprints,” submitted to 7th Workshop on Positioning, Navigation and Communication 2010 (WPNC’10).
- [125] V. Honkavirta, “Location fingerprinting methods in wireless local area networks,” Master’s thesis, Tampere University of Technology, 2008.
- [126] V. Honkavirta, T. Perälä, S. Ali-Löytty, and R. Piche, “A Comparative Survey of WLAN Location Fingerprinting Methods,” in *Proceedings of Workshop on Positioning, Navigation and Communication, 19th March, Hannover, Germany, 2009*.
- [127] N. Sirola, “Mathematical Methods for Personal Positioning and Navigation,” Ph.D. dissertation, Tampere University of Technology, 2007.
- [128] 3GPP TS 23.032, Universal Geographical Area Description (GAD), <http://www.3gpp.org>.
- [129] L. Wirola, “Generic Fingerprinting in SUPL 2.1,” Nokia Inc., Tech. Rep. OMA-LOC-2008-0303, 2008, presented in OMA LOC WG meeting, 19th-21st May, Wollongong, Australia.
- [130] —, “LPP OMA extensions draft,” Nokia Inc., Qualcomm, Nokia Siemens Networks, ZTE, Telcordia, LG Electronics, Andrew Corporation, Polaris Wireless, Tech. Rep. OMA-LOC-2009-0266R02, 2009, presented in OMA LOC WG meeting, 21st-23rd November, Los Angeles, USA.
- [131] Skyhook Wireless, <http://www.skyhookwireless.com>.
- [132] Ekahau, <http://www.ekahau.fi>.
- [133] Polaris Wireless, <http://www.polariswireless.com>.
- [134] L. Wirola, “SUPL3 RD Streaming,” Nokia Inc., ZTE, Telecom Italia, Andrew Corporation, Tech. Rep. OMA-LOC-2009-0178R03, 2009, presented in OMA LOC WG meeting, 22nd-25th June, Boston, USA.

- [135] T. Dao, "Performance evaluation of Multiple Reference Station GPS RTK for Medium Scale Network," Master's thesis, University of Calgary, 2005.
- [136] G. Ghinamo, "GNSS Assistance Broadcasting," Telecom Italia, Nokia Inc., Tech. Rep. OMA-LOC-2009-0081R05, 2009, approved in OMA LOC WG meeting, 22nd-25th June, Boston, USA.
- [137] OMA-TS-BCAST_Distribution-V1_0-20081209-C, File and Stream Distribution for Mobile Broadcast Services, Candidate Enabler 1.0, <http://www.openmobilealliance.org>.
- [138] L. Yang, C. Hill, T. Moore, and X. Meng, "The Communication Satellite Based Network RTK Solution for Road Transportation," in *Proceedings of ENC-GNSS 2009, 3rd-6th May, Naples, Italy*, 2009.
- [139] L. Wirola, "SUPL3 RD Assistance data push," Nokia Inc., ZTE, Telecom Italia, Andrew Corporation, LG Electronics, Tech. Rep. OMA-LOC-2009-0173R05, 2009, approved in OMA LOC WG meeting, 24th-26th August, Singapore.
- [140] S. Lo, D. Lorenzo, P. Enge, D. Akos, and P. Bradley, "Signal Authentication," *InsideGNSS*, pp. 30–39, 2009, September/October issue.
- [141] G. Hein, F. Kneissi, J.-A. Avila-Rodriguez, and S. Wallner, "Authenticating GNSS - Proofs against Spoofs," *InsideGNSS*, pp. 71–78, 2007, September/October issue.
- [142] P. Slaats, "Presentation of WID Information : WID 0183 - Secure User Plane Location v3.0," Vodafone, Tech. Rep. OMA-REQ-NWI-2008-0040R01, 2008, presented in OMA REQ-NWI WG meeting, 15th December, Cancun, Mexico.
- [143] S. Verhagen, *Manual for Matlab User Interface VISUAL*. Delft University of Technology, The Netherlands, 2006.
- [144] Satimo, Courtaboeuf, France, <http://www.satimo.fr>.

-
- [145] “GANSS - Ephemeris Extension,” SiRF Technology, Nokia Corporation, Nokia Siemens Networks, Broadcom, AT&T, Motorola, RIM, Tech. Rep. GP-071985, 2007, approved in 3GPP TSG-GERAN meeting #35, 12th-16th November, Vancouver, Canada.
- [146] “Feasibility of adding carrier phase -assistance,” Nokia Inc., Global Locate, SiRF Technology, Tech. Rep. GP-070835, 2007, presented in 3GPP TSG-GERAN meeting #34, 14th-18th May, Shenzhen, China.
- [147] L. Wirola, “AGNSS for User Plane,” Nokia Inc., Tech. Rep. OMA-LOC-2008-0386, 2008, presented in OMA LOC WG meeting, 19th-21st August, Chicago, USA.
- [148] —, “Socialization of WID 0181 - Generic Location Protocol 1.0,” Nokia Inc., Tech. Rep. OMA-TP-2008-0470R02, 2009, provided to OMA Technical Plenary, 8th January.

Publication 1

L. Wirola and K. Alanen and J. Käppi and J. Syrjärinne. Bringing RTK to Cellular Terminals Using a Low-Cost Single-Frequency AGPS Receiver and Inertial Sensors. In *Proceedings of IEEE ION PLANS 2006*, San Diego, USA, April 25th-27th, pp. 645–652, (2006)

Copyright ©2006 IEEE. Reprinted, with permission, from the proceedings of IEEE ION PLANS 2006.

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the Tampere University of Technology's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this material, you agree to all provisions of the copyright laws protecting it.

Bringing RTK to Cellular Terminals Using a Low-Cost Single-Frequency AGPS Receiver and Inertial Sensors

L. Wirola, K. Alanen, J. Käppi, J. Syrjärinne
Nokia Technology Platforms
Hatanpäänkatu 1 A
FIN-33101 Tampere, Finland

Biographies – Lauri Wirola, M.Sc., received his Master of Science degree from Tampere University of Technology, Finland, in 2005. His major was electrophysics and his minors were technical physics and industrial economics. He did his M.Sc. thesis on the mechanics of the human ear in the Nokia Research Center. Shortly after completing the thesis he started working with GPS in Nokia Technology Platforms. He has now been working with satellite navigation for a year. His present research interests include RTK and A-GNSS standardization. He is currently undertaking postgraduate studies in modern electromagnetism and mathematics.

Kimmo Alanen, M.Sc., received his Master of Science degree from Tampere University of Technology, Finland, in 2003. His major was Software Engineering and his minor was in communication networks and protocols. He did his M.Sc. thesis on the location software architecture in Nokia Mobile Phones. He joined the Nokia Corporation in 1997 and has been working with positioning research for the last 8 years. He is currently undertaking postgraduate studies in research on GNSS assistance protocol enhancements.

Jani Käppi, M.Sc., received his Master of Science degree from Tampere University of Technology in 2001. During 1999-2002, he worked at the Institute of Computer and Digital Systems at Tampere University of Technology doing research in the area of personal positioning. He has been working for the Nokia Corporation since 2002. He is currently researching the development of sensor-enhanced positioning systems.

Jari Syrjärinne, Ph.D., received his M.Sc. in 1996 and his Ph.D. in 2001, both from Tampere University of Technology, Finland, majoring in digital signal processing and applied mathematics. Between 1996 and 1998 he worked for Tampere University of Technology Signal Processing Laboratory in the areas of data fusion and target tracking, and since 1999 he has been with Nokia Inc. He is currently involved in work on A-GNSS and algorithms for hybrid positioning.

Abstract – Today an ever-increasing number of handsets come equipped with a GPS receiver and some even with inertial sensors. Moreover, an even higher number of terminals are already capable of connecting to an add-on device with such capabilities. However, the full potential of these devices is not yet exploited. This paper introduces the mobile RTK (mRTK) solution, which can be included in the wireless standards to enable high-precision double-difference carrier phase positioning in handsets at no extra hardware cost.

mRTK differs from the current OTF/RTK solutions in that it is a software-only solution using the hardware and wireless connections already existing in handsets. Moreover, the mRTK solution can utilize information from on-board inertial sensors. These are the key differentiating factors compared to the previous

solutions. The paper shows that the sensors supplying information on baseline changes during the ambiguity initialization significantly assist the ambiguity resolution.

A new communication protocol and messaging was defined in order to be able to exchange information between mRTK-capable handsets. The protocol includes reservations for additional GPS frequencies as well as for other Global Navigation Satellite Systems (GNSSs), such as Galileo. This protocol can be directly included in the wireless standards.

Challenges in the current implementation include using only the L1 frequency for ambiguity resolution. Utilizing an L1-only receiver necessarily leads to penalties in the baseline accuracy due to inherent problems in the ambiguity resolution and validation. However, this paper shows that the baseline obtained is still better than the plain difference of positions.

This paper shows that the mRTK solution significantly improves A-GPS performance. The mRTK solution also brings near-professional-quality positioning performance to the mass market. It would, therefore, be beneficial to include mRTK in wireless standards in order to expand A-GPS use cases in the short term and A-GNSS use cases in the long term.

I. INTRODUCTION

A plethora of Real-Time Kinematic (RTK) surveying solutions is available on the market today. Generally, they are characterized by the use of two frequencies, L1 and L2, highly stable oscillators as well as by the need for fairly long measurement periods. The use of two frequencies is beneficial in RTK, since it allows for utilizing wide-lane methods in the integer ambiguity resolution as well as for forming ionosphere-free linear combinations of the observed quantities [5]. In turn, this enables long-baseline RTK measurements, since ionosphere and troposphere affect long-baseline surveying. Moreover, high-end RTK devices also benefit from the possibility to use highly accurate post-processed ephemerides from, say, the International GNSS Service (IGS) [19].

The features available in commercial RTK solutions are, no doubt, required in professional work, but these stringent requirements also result in complex receivers and, therefore, in high receiver cost. In fact, the need to have a dedicated set of devices and software has mitigated against the widespread use of RTK techniques in everyday navigation.

This paper demonstrates that RTK functionality may be realized in mobile terminals using cheap single-frequency

A-GPS receivers. However, it should be emphasized that the current paper does not claim to demonstrate similar performance and reliability as high-performance two-frequency receivers with highly stable oscillators and phase locks. However, the paper shows that a decent-quality real-time surveying capability over short baselines may be achieved using off-the-shelf single-frequency A-GPS receivers and software.

II. MOBILE REAL-TIME KINEMATICS OVERVIEW

A. Definition

mRTK stands for mobile Real-Time Kinematics. mRTK is in the first place characterized as an RTK solution running on a mobile terminal utilizing either low-cost integrated or off-the-shelf Bluetooth GPS receivers. The only differentiating requirement is that the receivers must be able to measure and report carrier phase measurements and data polarity.

Secondly, mRTK is also characterized by its users. The design rule is that using mRTK does not require a priori surveying knowledge rather it only requires starting the application and binding it with the other receiver, called the reference. Another requirement set by the users is the speed. The baseline must be solved as quickly as possible. These requirements, however, necessarily mean that a certain percentage of erroneous ambiguity fixings is allowed and that the baseline length is limited to a few kilometers. This follows from using a single-frequency receiver, such as an L1-only receiver.

Thirdly, mRTK is characterized by the use of inertial sensors. The integer ambiguity resolution process benefits from the information that the baseline has been stationary during the resolution period. This will be shown in the study. On the other hand, if the integer ambiguities are initialized in the stationary mode, but the receiver should actually move, the resolution will fail. Now, due to the requirement that any mobile user must be able to use mRTK, the system must be as robust as possible. Hence, the system will also include an accelerometer that supplies the initialization algorithm with information on the baseline state. This function cannot solely be based on GNSS measurements, since even in ideal conditions the position solution drifts due to measurement noise. Since slow true motion cannot easily be detected from the noise-induced drift, an accelerometer is required.

Fourthly, mRTK is characterized by its readiness for inclusion as an optional positioning method in all major cellular standards. By considering the measurement messages exchanged between the two terminals as assistance data, mRTK functionality can fairly easily be added to the GSM and UMTS A-GPS/GNSS standards ([15], [16], [17]). Moreover, future systems, such as Galileo and modernized GPS, have been taken into account in the mRTK concept and the mRTK message format design. Hence, the mRTK solution is ready to utilize the benefits available to RTK by the addition of new signals [14].

Fifthly, mRTK is characterized by the baseline solution generally being better than just the plain difference of positions

and in good conditions (no ionospheric and tropospheric disturbances, a short baseline, and a long enough observation period) excellent. It is known, for instance, that the use of L1 only results in a lower ambiguity fixing success rate [9]. Also, the validation success rate is lower [9] when only L1 is utilized. This necessarily results in an adverse mRTK performance in certain circumstances. However, it should be borne in mind that the feature can be offered at no extra hardware cost and minimal software cost.

Sixthly, mRTK is characterized by a gradually improving baseline estimate. The first estimate for the baseline is given based on a single measurement epoch using code and carrier phase measurements. Alternatively, the first estimate can be given simply as the difference of positions, if the single-epoch ambiguity resolution is judged invalid. As time evolves, the goal is to be able to resolve ambiguities using carrier phase measurements and epochs with a long time span. This will be the best estimate. Between the first and the best estimate, the intermediate resolutions may include using code and carrier phase measurements and epochs with a sufficiently long time span. The choice ultimately depends upon the number of satellites in phase lock.

B. Some implementation aspects

mRTK utilizes code and carrier phase double differences that, ultimately, allow for solving the double-difference integer ambiguities and the fixed baseline solution. The formulation of the single-frequency double-difference ambiguity resolution problem is well documented in the literature and is, hence, not summarized here ([8], [10]).

The integer ambiguity resolution in mRTK is based on the Least-Squares Ambiguity Decorrelation Adjustment, LAMBDA ([4], [6], [10], [11], [12]). The LAMBDA method is well established both theoretically and experimentally, which makes it suitable for the current study. Moreover, a reference implementation is easily available from the developers [18].

The validation of the integer ambiguities is performed by calculating the discrimination ratio [9], which can readily be calculated based on the results produced by the LAMBDA algorithm. The discrimination ratio is a statistical quantity that describes the relative power of the best and the second-best ambiguity candidate vectors [13]. If the discrimination ratio exceeds a certain threshold, K , the best integer ambiguity candidate vector is validated and the fixed baseline solution may be calculated using the ambiguities. The threshold K is commonly set to 2.0 or above [9].

III. DEMONSTRATION PLATFORM OVERVIEW

Fig. 1 describes the mRTK working principle. In the system, there are two Nokia Symbian Series 60 mobile terminals that execute the mRTK software and two Bluetooth A-GPS (BAG) units that are positioned relative to each other. The terminals are connected to A-GPS receivers over Bluetooth using a proprietary low-level GNSS control interface protocol. The A-GPS receivers only report measurements and other relevant data to the calculation engine running in mobile terminals. The

actual calculation engine runs in the mobile terminal on a Symbian platform with mRTK being a module in the calculation engine.

The terminals are connected to an assistance server over GPRS, UMTS or WLAN using a proprietary protocol defined for R&D purposes. This protocol carries position, time, ephemeris, almanac and barometer assistance from the server to the terminal. The same server and protocol also relays mRTK measurements between the two terminals.

A. Bluetooth Assisted GPS and the BAG-terminal protocol

The Bluetooth Assisted GPS, BAG, incorporates the Fastrax iTrax03 12-channel L1-receiver. The A-GPS receiver is used only as a measurement engine and the measurements are transferred over Bluetooth using the proprietary protocol to the mobile terminal. This protocol is as general as possible in order to allow for the smooth addition of modernized GPS, Galileo and even GLONASS signals, if required in the future. The receiver package is shown in Fig. 2.

In addition, the package containing the A-GPS receiver also includes a low-cost MEMS-based 3-axis accelerometer, a 3-axis magnetometer and a barometer. The barometer may be used for generating altitude assistance and to stabilize the position solution in height ([2], [7]). The accelerometer, on the other hand, is used to support navigation should satellites be lost, but also provides the mRTK module with information on the state of motion of the GPS receiver. Since small movements cannot be detected reliably based on the position change, the only option to detect small true movements is by using an accelerometer. The information from the accelerometer is expressed in the form of a single bit, a stationary bit, in the mRTK use.

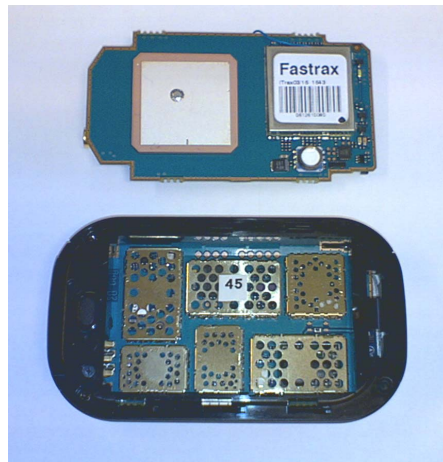


Fig. 2. Bluetooth Assisted GPS (BAG) demonstration platform. BAG has been designed and produced in Nokia for R&D purposes.

B. Assistance server and the server-terminal protocol

The assistance server relays mRTK measurement messages between the terminals. As of now, mRTK requires such a relay server since it is not yet possible to create terminal-to-terminal connections over GPRS or UMTS. The server also has a BAG unit connected to it for collecting time, position, ephemeris, almanac and barometric assistance to server users. These assistance data help the receiver to obtain a position fix more quickly and, therefore, also assist in resolving an accurate baseline in shorter time.

The protocol used in the communication is a proprietary one developed for R&D purposes. A full description of the protocol and the associated messaging is given in [3], but is summarized here for reference. The protocol allows for changing mRTK measurements between the two devices and follows loosely the messages in the RTCM standard [20] in bit counts and resolution, where appropriate. However, the RTCM has certain shortcomings that were remedied in the developed protocol. These include the lack of Galileo support as well as support for the additional GPS and GLONASS civil frequencies. Secondly, the RTCM messages do not allow for transferring Doppler frequency information between the terminals. Doppler frequency is required for the extrapolation of the code and carrier phase data from the two receivers to the common time base.

Thirdly, the RTCM standard [20] states that carrier phase measurements shall not be sent until the data polarity is resolved. The new protocol adds a “polarity unknown” flag to the message. This may be useful when one wants to optimize the performance and does not wish to buffer data at the receiving end.

Fourthly, the new protocol includes a stationary bit and a position uncertainty figure that are not available in RTCM messages. Adding a stationary bit enables relaying information produced by inertial sensors in the messages. Position uncertainty, on the other hand, is useful, if mRTK is used for accurate absolute positioning. In this case, the accurately surveyed reference A-GNSS receiver may report the high position accuracy in terms of low position uncertainty.

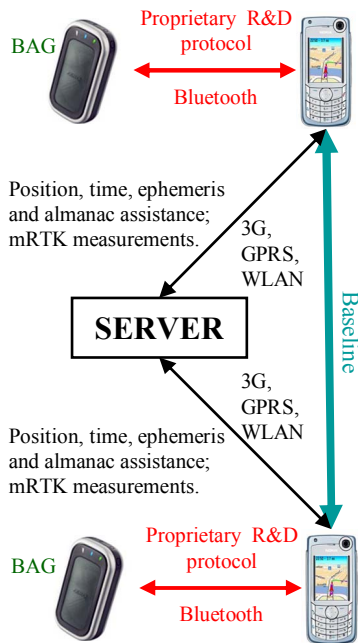


Fig. 1. Diagram of the demonstration platform. Two A-GPS-enabled handsets are positioned with respect to each other. A-GPS is connected to the cellular terminal via Bluetooth. The terminal connects to the assistance server over any given wireless standard. The server also relays mRTK measurements.

The mRTK measurement message contains the receiver time and position information that is, naturally, common to all signals. Moreover, for each signal tracked by the receiver, a set of measurement data is sent. The message fields are defined in Table I. As a general rule, one message is sent each second. However, the actual rate may vary depending upon the capabilities of the receiver, the chosen transfer path, as well as on the desired baseline update frequency.

Finally, it should be noted that the message format resembles the GPS Measurement Information Element that is defined in all the major cellular standards to report GPS measurements to the respective cellular networks. The additional elements include a stationary indicator, carrier phase measurements and a cycle slip indicator. The SS ID is a modification compared to the current realization (GPS PRN). The resemblance is no coincidence, since there has been an attempt to define the message format (fields, bit counts and scale factors) in such a way that it could directly be included in the cellular standards. [15] [16] [17]

III. SIMULATOR RESULTS

A. Stationary initialization

The solid blue and red lines in Figs. 3 and 4 show the mRTK performance in a simulator case in the absence of any non-idealities, such as troposphere, ionosphere, ephemeris or satellite clock disturbances. Eight satellites are used in the ambiguity resolution in each case and the baseline length varies from 100 to 10 000 m between the cases. Moreover, Figs. 3 and 4 represent two different instants, respectively, and, therefore, two different satellite configurations.

TABLE I
mRTK measurement message

Item	Explanation
Time and position information (once for each message)	
Time	UTC time in seconds
Position	Receiver position in the ECEF system
Position uncertainty	CEP50 value for the position uncertainty (meters)
Stationary indicator	Indicates if the receiver has moved between the last and current messages (according to sensors)
Measurement block (once for each signal)	
SS ID	A field expressing the system (such as GPS or Galileo), signal (such as L1 or L2C) and satellite (such as PRN for GPS). Ready for future augmentations.
Code phase	Code phase measurement (in milliseconds)
Code phase standard deviation	Code phase std (in milliseconds)
Carrier phase	Accumulated carrier phase measurement (in meters)
Carrier phase standard deviation	Accumulated carrier phase measurement std (in meters)
Carrier phase polarity	Indicates, whether the data polarity is nominal, inverted or unknown
Cycle slip indicator	RTCM style cumulative loss of continuity indicator [20] for carrier phase measurements
Doppler	Doppler frequency for carrier phase extrapolation/interpolation to common time base between receivers (in m/s)

The baseline is stationary in each case and the initialization is performed in the stationary mode, in which the sensors supply information indicating that the two receivers have remained stationary throughout the initialization period. This information is taken appropriately into account in building the system of equations used in solving the double-difference integer ambiguities. In such a case, the ambiguity resolution may be performed using carrier phase data alone as soon as there are measurement data from both receivers from two epochs having an appropriate time separation.

The data shown in solid lines in Figs. 3 and 4 have been produced by making successive 2-epoch ambiguity resolutions in the stationary mode. The first ambiguity resolution is made after 15 s using measurement data from the 0-s and 15-s epochs. The baseline is tracked with this integer ambiguity set for the next 14 s. The ambiguities are resolved again at 30 s using 0-s and 30-s epochs, then again at 45 s using 0-s and 45-s epochs, and so on. The first aim of the simulation run is to find the minimum time required for convergence and to see if the ambiguity set stabilizes after convergence.

The measure of convergence is the error between the calculated fixed baseline and the true baseline. The true baseline is obtained from the simulator and is the best available estimate for the true baseline. The error is defined as the 2-norm of the 3D error vector. The error vector magnitude is plotted in solid blue as a function of time in Figs. 3 and 4.

The second aim is to test if the discrimination ratio is capable of indicating the baseline convergence. The discrimination ratio is plotted in solid red as a function of time in Figs. 3 and 4. Moreover, a green line representing a validation limit of 2.0 is plotted for reference.

The results from the stationary initialization runs are promising. With the 100-m baseline the integer ambiguities are resolved and validated correctly in 15 s. The same applies to the 1000-m case with the exception that the validation takes 30 s in the data collected in week 340. However, as the baseline gets longer, the time required for the ambiguity resolution constantly gets longer. In the week-340 data, the convergence of the 5-km baseline already takes 105 s and validation 255 s. However, the week-344 data show a significantly better performance with the times being 30 and 60 s, respectively. The reason for this discrepancy is currently unknown.

Finally, the time required for the convergence of the 10-km baseline in week 340 is 570 s, which time might already be considered too long by some users. However, it should be borne in mind that a 10-km baseline might in any case be a too optimistic attempt in the real world due to tropo- and ionosphere effects entering the solution. In fact, the use of an L1-only receiver will most likely restrict feasible baseline lengths to below 10 km in real-world circumstances. This matter will be researched in future field tests.

The increase in the time required for convergence and validation can clearly be seen in Fig. 5. In fact, Fig. 5 seems to suggest that the time required for the ambiguity resolution increases in a logarithmic manner as a function of the baseline length. However, this observation is based only on two sets of data and should not be considered as a definite conclusion.

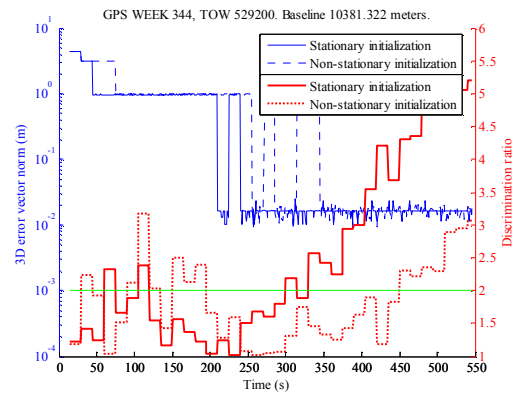
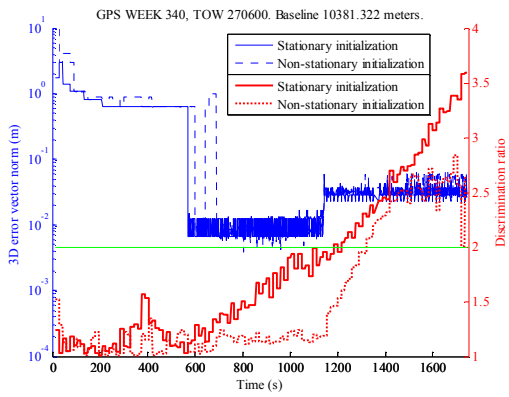
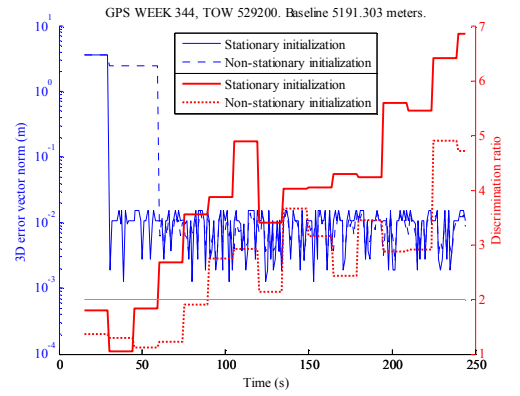
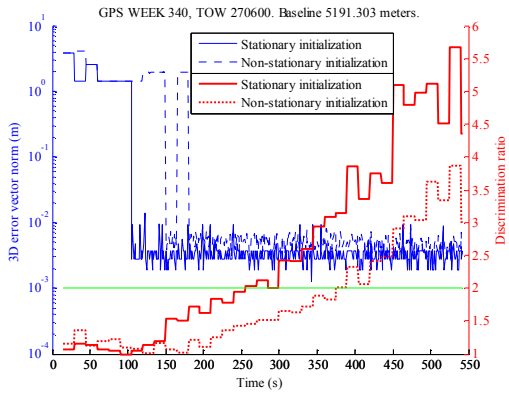
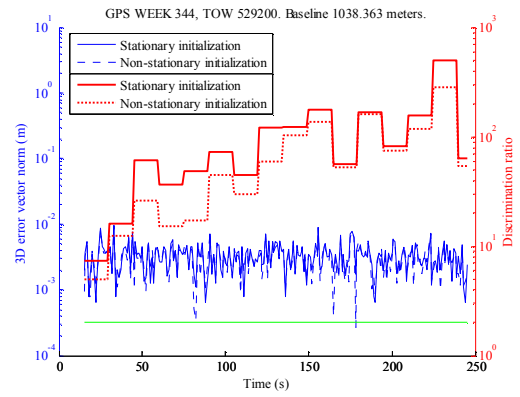
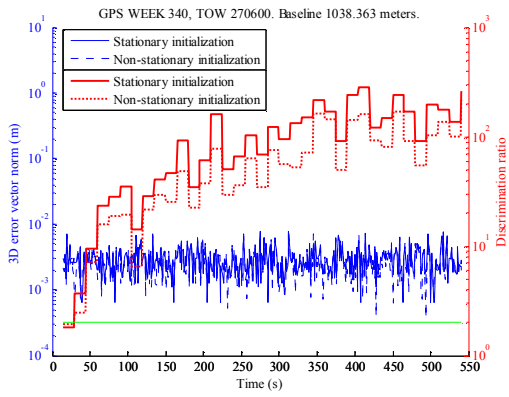
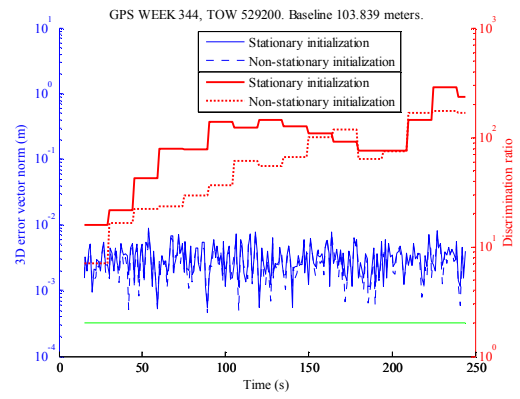
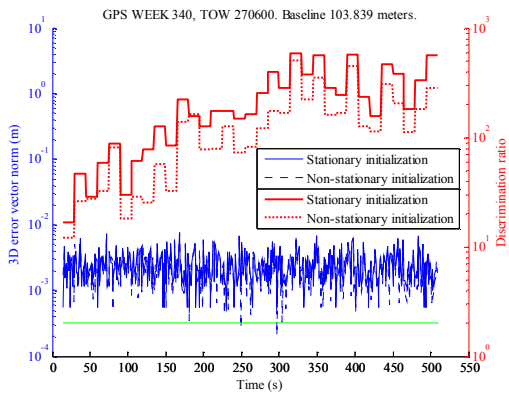


Fig. 3. Performance comparison between cases in which there is stationary information on the baseline available (solid lines) and unavailable (dashed lines).

Fig. 4. Performance comparison between cases in which there is stationary information on the baseline available (solid lines) and unavailable (dashed lines).

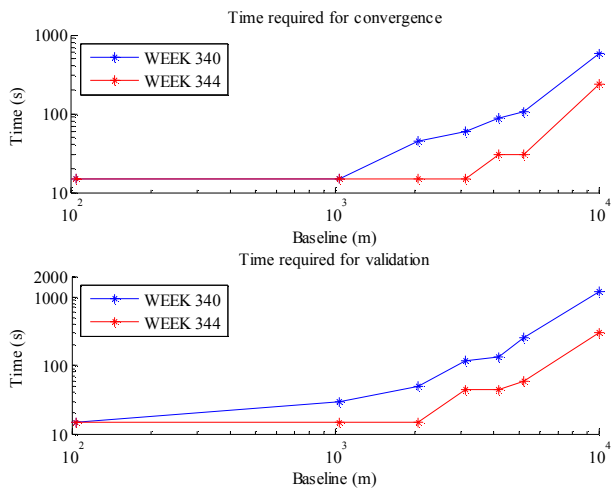


Fig. 5. Time required for convergence and validation as a function of baseline length.

Another promising aspect of the simulator results is that the discrimination ratio is able to indicate the convergence correctly in all but one case (week 344, 10-km baseline). It should still be noted that if the limit for the validation were set to 2.5 no errors in the validation would have been made. However, as one is discussing the validation problem, it should be borne in mind that the validation is essentially an open problem [9]. Moreover, the literature describes a plethora of more sophisticated validation methods than the currently used discrimination method. One such method is presented in [1].

Yet another interesting observation is the increasing trend in the discrimination ratios as a function of time. This means that the resolution becomes more and more certain as data accumulates in the time dimension. This is known from the literature [13], and the current observations support the earlier work.

Finally, considering errors in the converged baseline, it can be observed that the mean errors for the 1-km baseline are 2.5 and 1.5 mm in weeks 340 and 344, respectively. For the 10-km baseline the same figures are 20.8 and 16.2 mm, respectively. This is in agreement with the simplified error formula derived by Leick [8], which takes into account only the receiver and satellite position errors and omits, for instance, the effect of the satellite geometry. The formula yields errors of approximately 1 and 10 mm for 1-km and 10-km baselines, respectively. The order-of-magnitude estimate given by Leick agrees with the current observations.

B. Benefit from the use of inertial sensors

Figs. 3 and 4 show the effect of stationary information from the inertial sensors on the ambiguity resolution. The results shown in solid and dashed lines have been produced by using the same measurement data and the same method, i.e. by performing successive 2-epoch ambiguity resolutions every 15 s with a constantly increasing time span. However, the crucial difference between the two data sets is that solid and dashed lines show the results from stationary and non-stationary ambiguity resolutions, respectively. The non-stationary

ambiguity resolution means that the baseline is assumed to change between the two epochs used in the ambiguity resolution. This affects the equation system formulation. Since the data are the same and the baseline is in reality stationary, the current simulation run enables the evaluation of the significance of the stationary information on the ambiguity resolution.

Figs. 3 and 4 allow for drawing some general conclusions on the non-stationary ambiguity resolution. Firstly, the baselines obtained in different modes eventually converge into the same baseline. This convergence of the results is, naturally, a minimum requirement and an expected result.

Secondly, the discrimination ratios show similar trends between the resolution modes. The ratios are somewhat smaller in the non-stationary case, but are still clearly able to indicate the baseline convergence. An interesting detail is that the validation problem observed with the week-344 10-km stationary baseline is observable in the non-stationary mode as well. In fact, the problem is more emphasized in the non-stationary case, since the discrimination ratio climbs above 3 at approximately 120 s in the non-stationary mode, although the baseline errs by about 1 m.

Thirdly, the ambiguity resolution is fast, sub-15 s, for short baselines (0-1 km) in the non-stationary mode as well. The data also show similar performance degradation with increasing baseline length. This degradation includes the growing error as well as the increase in the time required for the resolution. However, as the baseline length increases, the difference between stationary and non-stationary solutions becomes more profound. For instance, with the 5-km baseline the week-340 data show a 71% and 47% increase in the time required for convergence and validation, respectively, when no stationary information is available.

Fig. 6 shows the increase in the time required for convergence and validation, when the resolution is performed without stationary information. Fig. 6 has been derived directly from the data shown in Figs. 3 and 4. The data clearly show that the ambiguity resolution benefits from the sensors supplying stationary information. Although the sensor integration increases the system cost, this will be offset by the increased usability. This is because the burden of inputting the stationary information cannot be put on the system users, since this would increase error possibilities and degrade user experience. Since the stationary information clearly improves the system performance, the use of an inertial sensor with an mRTK A-GPS/GNSS receiver is recommended.

C. Initialization in a moving case

Fig. 7 shows the increase in the time required for the baseline convergence and validation when the baseline is changing as compared to a situation when the two terminals are stationary and stationary information is also utilized in the equation system formulation. The analysis method is the same as that used before, i.e. successive 2-epoch initializations made every 15 s. Again, the time span of the two epochs used in the ambiguity resolution increases constantly. In the moving case, the reference terminal is moving at 3.0 m/s in a circular path

with a radius corresponding to the baseline length in the stationary cases. The stationary data used as a reference are the same data as shown in solid lines in Figs. 3 and 4. Moreover, the satellite constellations are the same in both the stationary and moving cases. Therefore, the results from the two cases are believed to be sufficiently comparable. Naturally, in the moving case the sensors supply the resolution algorithm with the information that the baseline is changing in time. However, it should be noted that since the speed of the terminal is 3.0 m/s, the change can also be detected from the change in the position. Strictly speaking, therefore, sensors provide no additional information in this case.

Fig. 7 indicates that the mRTK performance in a moving scenario is comparable to the stationary performance for baselines up to 1-2 km. With longer baselines the performance degradation becomes visible. For instance, the week-340 5-km baseline takes about 60% more time to converge in a moving case than in the stationary case. However, it is noteworthy that the moving scenario outperforms the stationary case, in which the stationary information was not utilized in the resolution (Fig. 6). The statement holds true for all the considered baselines in both weeks except for the week-340 10-km baseline. The reason for the observed performance difference might be that in a truly moving case the receiver movement contributes to the satellite geometry change, which assists the ambiguity resolution. However, there is no direct evidence supporting this conclusion and, hence, the issue must be examined more rigorously in the future.

V. FUTURE WORK

The selection logic in mRTK that decides which epochs and satellites to use in the ambiguity resolution must be developed further. The selection logic balances between time and satellite constellation. It is still unclear how the satellite count, constellation geometry and time should be valued with respect to each other, since, obviously, all are important for a successful ambiguity resolution.

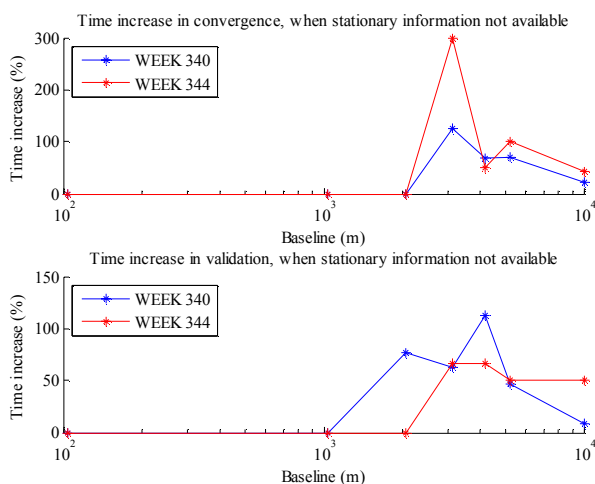


Fig. 6. Increase in the time required for convergence and validation when stationary information on the baseline is unavailable as compared to the situation when this information is available.

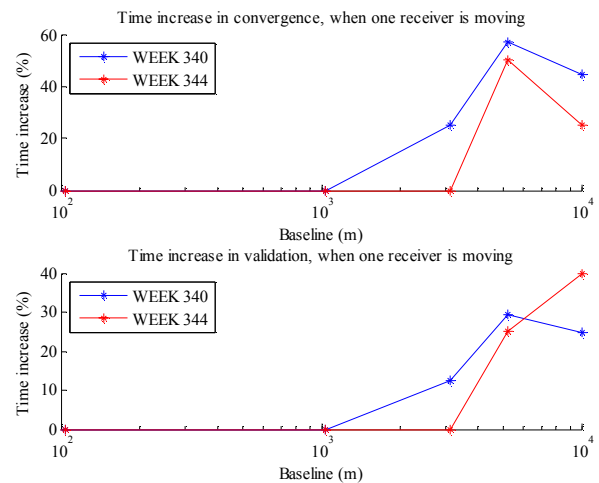


Fig. 7. Increase in the time required for convergence and validation when one receiver is moving at 3.0 m/s as compared to the situation when the baseline is stationary and stationary information is utilized in the resolution.

The other major planned future work item includes thorough field testing. This will be performed with the help of accurately surveyed points in Tampere, Finland.

VI. CONCLUSIONS

This paper introduces the concept of mobile Real-Time Kinematics and shows that the L1 RTK is feasible given certain boundary conditions, such as a fairly short baseline (1-2 km). However, this should not be considered as a too restrictive requirement, since mRTK is able to utilize Virtual Reference Stations (VRSs). A VRS is typically brought very close to the location to be surveyed and, hence, the baseline length limitation is not an issue. Moreover, the mRTK was shown to work with both stationary and changing baselines.

This paper also introduces a new protocol to relay mRTK measurements in wireless networks. The protocol introduces several improvements over the currently widely used RTCM protocol. The new protocol was shown to work in the demonstration system developed. The protocol message format can be included in cellular standards as such.

This paper also shows that integrating inertial sensors in the system can provide valuable information for the ambiguity resolution algorithm. The use of inertial sensors reduces error possibilities and accelerates the baseline convergence simultaneously improving user experience.

The study shows that mRTK is a feasible add-on product for A-GPS-enabled handsets and, therefore, the measurement message format should also be included in the major cellular standards, such as GSM and UMTS. Moreover, mRTK holds a great promise for the future. The concept and results presented herein clearly show the potential the single-frequency mRTK has for everyday navigation. This potential is further enhanced by the launch of new GNSSs, such as Galileo and the Japanese Quasi-Zenith Satellite System, which will together more than double the number of usable satellites. Since mRTK and the associated protocols are completely A-GNSS ready, the mRTK

will be able to exploit all the benefits arising from additional navigation signals as soon as they become available. The same applies also to pseudolites, which current GPS receivers are already able to utilize. This technology may also enable indoor mRTK in the future.

REFERENCES

- [1] M.F. Abdel-Hafez, J.L. Speyer, W.R. Williamson, "Multiple Hypothesis Wald Sequential Probability Ratio Test for GPS Integers Ambiguity Resolution", *Navigation: Journal of The Institute of The Navigation*, volume 51, number 3, pages 231-247.
- [2] K. Alanen, J. Käppi, "Enhanced Assisted Barometric Altimeter A-GPS Hybrid Using the Internet", *Proceedings of ION GNSS 18th International Technical Meeting of the Satellite Division*, Long Beach, CA, 13-16 September 2005, pages 2248-2251.
- [3] K. Alanen, L. Wirola, J. Käppi, J. Syrjärinne, "Inertial Sensor Enhanced Mobile RTK Solution Using Low-Cost Assisted GPS Receivers and Internet Enabled Cellular Phones", *Proceedings of IEEE/ION PLANS 2006*, San Diego, CA, 25-27 April 2006, in press.
- [4] P. de Jonge, C. Tiberius, *The LAMBDA method for integer ambiguity resolution: implementation aspects*, Publications of the Delft Geodetic Computing Center number 12, Delft: Universiteitsdrukkerij TU Delft, 1996.
- [5] E. Kaplan, *Understanding GPS Principles and Applications*, Norwood, MA, Artech House, Inc., 1996.
- [6] D. Kim, R.B. Langley, "GPS Ambiguity Resolution and Validation: Methodologies, Trends and Issues", *7th GNSS Workshop – International Symposium on GPS/GNSS*, Seoul, Korea, November 30 – December 2, 2000.
- [7] J. Käppi, K. Alanen, "Pressure Altitude Enhanced AGNSS Hybrid Receiver for a Mobile Terminal", *Proceedings of ION GNSS 18th International Technical Meeting of the Satellite Division*, Long Beach, CA, 13-16 September 2005, pages 1991-1997.
- [8] A. Leick, *GPS Satellite Surveying*, 3rd edition, Hoboken, NJ: John Wiley & Sons, 2004.
- [9] T. Richert, N. El-Sheimy, "Ionospheric Modeling: The Key to GNSS Ambiguity Resolution", *GPS World*, June 2005, pages 35-40.
- [10] P.J.G. Teunissen, "A new method for Fast Carrier Phase Ambiguity Estimation", *Proceedings of IEEE Position, Location and Navigation Symposium PLANS'94*, Las Vegas, NV, April 11-15 1994, pages 562-573.
- [11] P.J.G. Teunissen, P.J. de Jonge, C.C.J.M Tiberius, "The least-squares ambiguity decorrelation adjustment: its performance on short GPS baselines and short observation spans", *Journal of Geodesy*, number 71, pages 589-602.
- [12] P. Teunissen, P. de Jonge, C. Tiberius, "On the Spectrum of GPS DD-Ambiguities", *Proceedings of ION GPS 1994*, Salt Lake City, Utah, September 20-23 1994, pages 115-124.
- [13] C.C.J.M. Tiberius, P.J. de Jonge, "Fast Positioning Using the LAMBDA-method", *Proceedings of the 4th International Symposium on Differential Satellite Navigation Systems DSNS'95*, Bergen, Norway, April 24-28 1995.
- [14] S. Verhagen, "How will the new frequencies in GPS and Galileo affect carrier phase ambiguity resolution", *InsideGNSS*, March 2006, pages 24-25.
- [15] 3GPP TS 04.31 and 44.035, Location Services (LCS); Mobile Station (MS) – Serving Mobile Location Center (SMLC) Radio Resource LCS Protocol (RRLP), <http://www.3gpp.org>.
- [16] 3GPP TS 25.331, Radio Resource Control (RRC) protocol specification, <http://www.3gpp.org>.
- [17] OMA-TS-ULP-V1-0-20050719-C, User Plane Location Protocol, http://www.openmobilealliance.org/release_program/supl_v1_0.html.
- [18] Delft University of Technology, Netherlands, <http://www.lr.tudelft.nl>.
- [19] International GNSS Service, IGS, <http://igscb.jpl.nasa.gov>.
- [20] Radio Technical Commission For Maritime Services, *RTCM Recommended standards for differential GNSS (Global Navigation Satellite Systems) Service*, version 2.2. RTCM Special Committee No 104, 1998.

Publication 2

L. Wirola and S. Verhagen and I. Halivaara and C. Tiberius. On the feasibility of adding carrier phase assistance to cellular GNSS assistance standards. *Journal of Global Positioning Systems*, Vol. 6, No. 1, pp. 1–12, (2007)

Copyright ©2007 Journal Of Global Positioning Systems. Reprinted, with permission, from Journal Of Global Positioning Systems.

On the feasibility of adding carrier phase –assistance to cellular GNSS assistance standards

Lauri Wirola, Ismo Halivaara

Nokia, Inc., Finland

Sandra Verhagen, Christian Tiberius

Mathematical Geodesy and Positioning, University of Delft, The Netherlands

Abstract. The 3GPP (Third Generation Partnership Project) Release 7 of GSM and UMTS cellular standards as well as SUPL2.0, used in IP networks, include major modifications as to how AGNSS (Assisted GNSS) assistance data is transferred from the network (cellular or IP) to the cellular terminal. Simultaneously position accuracy improvements may be introduced. One potential option is to use carrier phase -based positioning methods. This can be achieved integrally in the cellular network or by the use of Virtual Reference Stations and an IP network. The bulk of AGNSS devices will be single-frequency due to additional cost associated with two RF front-ends. Hence, this study addresses the feasibility of single-frequency carrier phase-based positioning, making comparison with the dual-frequency case. The study shows that single-frequency carrier phase -based positioning is feasible with short baselines (<5 km) given that: 1) real-time ionospheric predictions are available and 2) there are enough satellites available. Namely, this requires hybrid-use of GPS and Galileo.

Keywords. Assisted GNSS, RTK, VRS, Ambiguity Resolution, Success Rate

1 Introduction

The annual sales of AGNSS-enabled (Assisted GNSS) handsets are estimated to rise to 400 million units by 2011 (Strategy Analysts, 2006). Currently the size of the market is approximately 100 million units annually. High growth requires developing constantly more efficient and capable methods to improve user experience in terms of availability, accuracy and short time-to-first-fix. The assistance data available from the network are a

significant factor affecting the user experience. The advantages and benefits of assistance are discussed in (Wirola et al., 2007b).

As GPS/AGPS now becomes commonplace in mobile terminals, the next step in the competition will be the race for accuracy. One option to achieve this is to take advantage of carrier phase -measurements readily available in GNSS receivers integrated in mobile terminals. Methods utilizing carrier phase -measurements include Real-Time Kinematic (RTK) as well as Precise Point Positioning (PPP). The recommendation given in (Nokia, 2006) is that carrier phase -based positioning would be added to the cellular standards in such a manner that the terminal could request for carrier phase-assistance from the SMLC (Serving Mobile Location Center) and calculate the baseline vector between the base station and the terminal.

Carrier phase -based positioning was for the first time introduced in 3GPP (The Third Generation Partnership Project) in GERAN#30 (GSM/EDGE Radio Access Network with GSM being Global System for Mobile communications and EDGE being Enhanced Data rates for Global Evolution) meeting in June 2006 in Lisbon, Portugal (Nokia, 2006). When the baseline implementation for A-Galileo was agreed in GERAN#32, this feature was included in the list of items to be reviewed in the 3GPP Release 7 time frame (Alcatel et al., 2006). However, the feature was not included in the Release 7 due to the identified need to further assess the technical implementation before approving the approach. It is expected that carrier phase -based positioning will be dealt with in the Release 8 of the 3GPP standards.

This paper examines the feasibility of introducing single-frequency carrier phase -based positioning into cellular networks. The use case considered consists of a short baseline (<5 km) and a single-frequency receiver due to the cost reasons. However, the receiver may be a dual-

GNSS (GPS+Galileo) receiver. The paper includes a thorough review of the latest research in the area of carrier phase-based position. The review is complemented by simulations that are performed using a state-of-the-art open-source simulation tool developed for the analysis of carrier phase-based positioning (Verhagen, 2006b).

2 Assisted GNSS

Fig. 1 shows the high-level view of AGNSS architecture. The core of the architecture is the AGNSS server, or more precisely, server centers that are geographically distributed. These centers serve the AGNSS-subscribers in each geographical area. Assuming that the AGNSS-terminal is to receive assistance over the user plane (IP-network) the terminal takes a data connection to the pre-set server and requests for the assistance data. The assistance data is then delivered to the terminal as specified in the associated standards.

The AGNSS server may obtain its data from various sources. These may include physical GNSS-receivers distributed geographically (left hand side in Fig. 1). These receivers can provide integrity information as well as broadcast ephemerides to the AGNSS server for distribution. On the other hand, the orbit and clock models (as well as other data) can originate from an external service providing, for instance, precise ephemerides and orbit/clock predictions (right hand side in Fig. 1). Such services include the International GNSS Service, or IGS (Dow et al., 2006). Should predictions be available, AGNSS-enabled terminals can be provided with extended ephemerides, in which case the terminal does not need to connect to the assistance server in the beginning of each positioning session. This improves user experience due to the time saved in not having to set up a data connection and download the assistance. With long-term ephemerides the assistance is also available, when there is no network coverage (Lundgren et al., 2005).

Currently it is only possible to provide assistance for GPS L1 in GSM and UMTS (Universal Mobile Telecommunications System) networks. In GSM the assistance is specified in the Radio Resource LCS (Location Services) Protocol (RRLP, (3GPP-TS-44.031)) and in UMTS in the Radio Resource Control (RRC, (3GPP-TS-25.331)). Moreover, there are also user plane solutions, such as Open Mobile Alliance (OMA) Secure User Plane Location (SUPL, (OMA-ULP)) protocol.

It should be noted that there are terminological differences depending upon, which standard is in question. For instance, the mobile terminal is MS (Mobile Station) in GSM, UE (User Equipment) in UMTS and SET (SUPL-Enabled Terminal) in SUPL. Moreover, the server sending the assistance to the terminal is an SMLC

in RRLP and RRC, while in SUPL the server is an SLC (SUPL Location Center).

Due the upcoming changes in the GNSS infrastructure (Wirola et al., 2007b), such as modernization of GPS and GLONASS as well as the introduction of Galileo amongst others, the 3GPP standardization body accepted a proposal which opened the way for the addition of new GPS bands as well as other GNSSs to the assistance standard in autumn 2006 (3GPP, 2006). This decision concerned RRLP only, but the same solution was later approved into RRC (3GPP, 2007) as well as SUPL 2.0 (OMA, 2007).

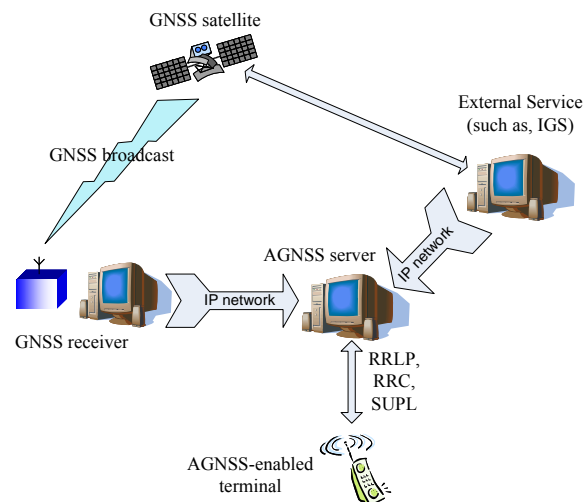


Fig. 1. The AGNSS architecture

AGNSS introduces common and per-GNSS elements into the standards. The superstructure is detailed in (Syrjärinne et al., 2006). The common elements are GNSS-independent and include, for instance, ionosphere model and reference location. In the future, for instance, troposphere models or Earth-Orientation Parameters can be added without obstacles.

The per-GNSS elements, on the other hand, are by definition GNSS-dependent (as well as signal-dependent) and include differential corrections, real-time integrity, GNSS-common time relation, data bit assistance, reference measurements as well as orbit and clock models (ephemerides). The new multi-mode navigation model capable of supporting at least seven GNSSs is discussed in (Wirola et al., 2007a) and (Wirola et al., 2007b). The introduced generic approach significantly reduces the system complexity.

3 Carrier phase -based positioning

Real-Time Kinematic (RTK) techniques utilize carrier phase -measurements that are readily obtained from a GNSS receiver. Carrier phase measurements enable centimeter-level accurate baseline (i.e. distance and

attitude between the receivers) determination between two (or more) GNSS receivers. Also, if the absolute position of one receiver is known at high accuracy, the absolute position of the other receiver can easily be deduced. The addition of carrier phase -based positioning to cellular standards, therefore, potentially enables ubiquitous cm- or dm-level positioning accuracy.

The current commercial solutions typically utilize both GPS L1 and L2 signals for high-precision surveying. Moreover, with the GLONASS modernization (Klimov et al., 2005), the utilization of multi-GNSS is becoming ever more attractive. Also, the recent studies (Wirola et al., 2006; Alanen et al., 2006a; Alanen et al., 2006b) show that single-band single-GNSS RTK is feasible under certain circumstances. In addition, all the Galileo as well as the modernized GPS signals can be utilized in the baseline determination (Eisfeller et al., 2002a; Eisfeller et al., 2002b; Tiberius et al., 2002). The more signals there are the more certain (in statistical sense) the baseline becomes (Wirola et al., 2006).

Carrier phase -based positioning may be introduced either by supporting it in the SMLC or by utilizing an external service. In the case of an SMLC-implementation (control plane solution in the cellular network), the terminal requests for carrier phase -measurements from the SMLC. The SMLC then starts sending the measurements from the LMU (Location Measurement Unit) to the terminal. Another option is to utilize Virtual Reference Stations (VRS) as a service external to the network. In this case the terminal sends the AGNSS assistance server an assistance request that contains the approximate position of the terminal. A VRS is created to this location and measurements are streamed to the terminal most likely over an IP-network. The advantage of this technology is that the baseline is always very short and no additional hardware (LMUs) is required in the network.

The key to the high-accuracy baseline determination is *integer ambiguity resolution*, for which there are many algorithms available. In addition to solving the ambiguities, another key issue is the *validation of ambiguities*. Validation refers to using statistical tools to determine, whether the ambiguities and, hence, the fixed baseline solution can be relied on. If the ambiguities cannot be solved, somewhat less accurate option is to utilize the *float solution*. In this case the ambiguities are not fixed to their integer values, but are considered as real numbers.

This study concentrates on discussing the various factors affecting the ambiguity resolution success rate and how those factors affect the feasibility of adding carrier phase-based positioning to the 3GPP standards.

4 Method and analyses

In the following the performance of the carrier phase -based positioning is analyzed under varying circumstances. Chapter V examines a situation, in which a set of individual measurements is exchanged between two receivers. This corresponds to *Measure Position Response with Multiple Sets* defined in RRLP (3GPP-TS-44.031). Chapter VI studies a situation with periodic reporting of measurements from one receiver to another as defined in RRC (3GPP-TS-23.271).

The performance is characterized in terms of the success rate for fixing the integer ambiguities successfully. Theoretical tools for this analysis are given, for instance, in (Teunissen et al., 2000). This work utilizes an open-source analysis tool called VISUAL (Verhagen, 2006b), which allows for simulating success rates in temporal or spatial dimensions.

In real-time applications ambiguity fixing success rate can be calculated on-the-fly in order to examine, whether ambiguity fixing should be attempted at all. As a general rule, the success rate must be above 99% before fixing should be attempted (Verhagen, 2006b). If the ambiguity solution is not available, the system can provide the user with a float solution. Baseline accuracy obtainable with a float solution is 0.1 - 1.0 meters.

5 Single-shot multiple-sets

The first set of simulations considers a case, in which one receiver makes three measurements with 50-s spacing corresponding to the total measurement time of 100 s. This can be considered as a situation, in which the MS sends multiple sets of carrier phase measurements to the SMLC (3GPP-TS-44.031) allowing the SMLC to calculate the baseline.

Fig. 2 shows the success rates for Galileo E1 (up) and for Galileo E1+E5a (below). The parameters and assumptions of the simulation are

- 5-km stationary baseline
- Date 1st January 2008 00:00:00 UTC
- 15-degree elevation mask
- Fixed ionosphere (i.e. external ionosphere model used to correct the observations)
- Float troposphere (i.e. troposphere delay estimated as state) with Ifadis mapping function
- 3-mm STD for carrier phase observations
- 30-cm STD for code phase observations
- 30-satellite Galileo constellation

Fig. 2 shows that single-band carrier phase -based positioning using only Galileo should be considered too unreliable for implementation. On the other hand, the addition of the second frequency (E5a) improves the

performance significantly. In the dual-band case, the carrier phase -based positioning is enabled and feasible globally.

Consider then temporal changes in the success rates. Fig. 3 shows the success rate as a function of time in Paris for Galileo E1 (up) and Galileo E1+E5a (below). The date and other assumptions are the same as before.

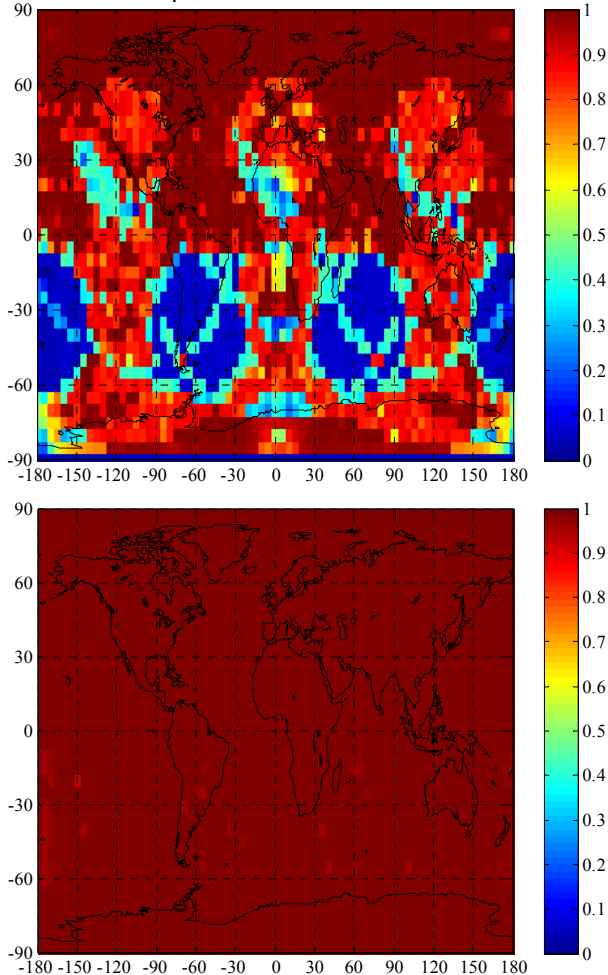


Fig. 2. Ambiguity fixing success rate for single-shot multiple-sets. Up: Galileo E1, Below: Galileo E1+E5a.

The simulation shows that in a single-frequency case the success rate is highly dependent upon the number of satellites available. In general, it seems that carrier phase -based positioning is feasible, when there are at least 10 satellites visible. However, there are only short periods, when this takes place. On the other hand, dual-band positioning does not suffer from the lack of satellites. Only if the number of satellites is below seven the success rate drops below the threshold. The dual-band case clearly outperforms the single-band case.

The literature supports the conclusions drawn from the simulations. Tiberius et al. (Tiberius et al., 1995) report 100% ambiguity fixing rate, when using GPS L1+L2 code and carrier phase measurement and only one set of

measurements (one instant). In the study seven or more satellites were used all the time and the baseline was in the order of one km. However, the authors reported problems with validating the calculated ambiguities.

Finally, if GPS and Galileo are used in hybrid, the situation improves significantly. This is shown in Fig. 4, in which the simulation shown up in Fig. 3 has been rerun adding the GPS L1 signal. The results show that the redundancy from additional satellites (29-satellite GPS constellation) contributes significantly to the success rate. There are only few short periods during which there might be problems with fixing the ambiguities. The finding is also supported by the literature. For instance, Verhagen (Verhagen, 2006a) reports that combined dual-band GPS+Galileo yields a constant success rate of >99.9%. In that case the success rate becomes almost independent of time and location. Increased number of satellites is identified as the single most important factor for high success rate. However, there is no information, how the ambiguity validation success rate behaves in a combined GPS L1 + Galileo E1 situation.

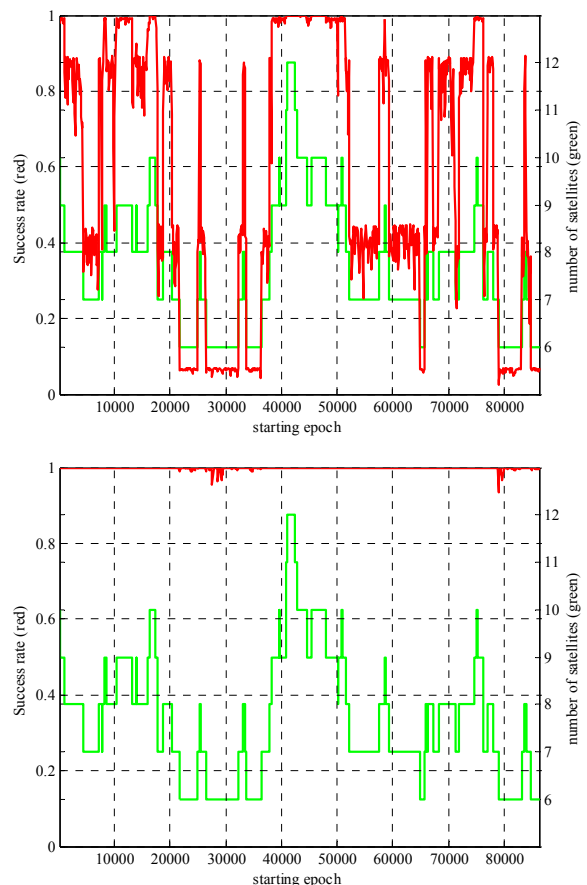


Fig. 3. Ambiguity fixing success rate for single-shot multiple-sets over one day in Paris (48.5° N, 2.2° E). Red denotes success rate and green the number of satellites above the elevation mask. It is assumed that all the satellites above the mask can be used in the ambiguity resolution.

Up: Galileo E1, Below: Galileo E1+E5a.

Single-shot data delivery means that the baseline may be solved once (when the set of measurements arrives), but not updated after that. The receiving terminal/server may extrapolate the measurements for 20-30 s without losing accuracy significantly (Schüler, 2006). However, the baseline is lost after this in the case the receivers (or one of the receivers) are moving. Therefore, the single-shot multiple-set method is useful only for stationary receivers. Moreover, since there is no possibility for rigorous solution quality and integrity monitoring in time, baselines should be limited to short ones. The exact length depends on the bands and GNSSs used as well as on the atmospheric conditions and also on whether ionosphere or troposphere models are available.

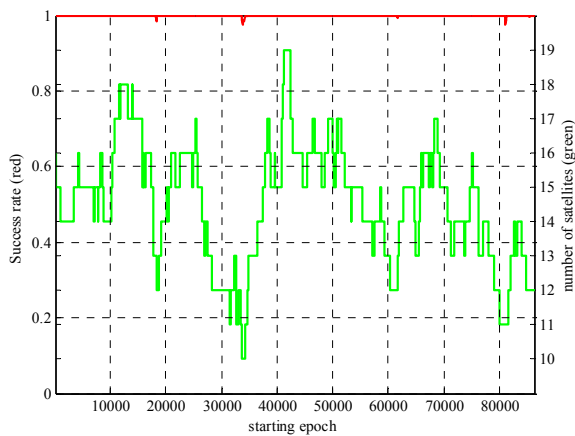


Fig. 4. Ambiguity fixing success rate for single-shot multiple-sets over one day in Paris (48.5° N, 2.2° E), when GPS L1 + Galileo E1 are used.

6 Periodic measurements

Periodic measurements refer to a case, in which one receiver periodically sends its signal measurements to the other receiver. This enables, for example, monitoring the solved parameters in time and, therefore, quality control. Also, with multi-band receivers, filtering of ionosphere advance (as well as troposphere delay) becomes possible. Finally, longer observation periods assist the validation process. Periodic reporting is enabled in UMTS networks over RRC.

Fig. 5 shows the success rates for Galileo E1 (up) and for Galileo E1+E5a (below), when one receiver streams measurements to the other receiver - in this case 1 signal measurement every 10 s for 100 s (in total 11 measurements). Note that by a signal measurement one understands a set of measurements consisting of code and carrier phases for all the observable satellites and signals. The other parameters and assumptions of the simulation are as given in chapter V.

Fig. 5 shows a major improvement in the single-band case. It appears that the single-frequency carrier phase -

based positioning becomes feasible in many locations, when several epochs are utilized in the solution. However, the analysis made for Paris for the same situation running over one day (Fig. 6) shows that although there is an improvement as compared to the results shown in Fig. 3, windows for successful carrier phase -based positioning are still few. The promising periods are now longer (for instance, between epochs 40000 - 50000 s), but it can be assumed that the high variation in the success rate in time makes single-band positioning still very challenging even if more measurements are now available.

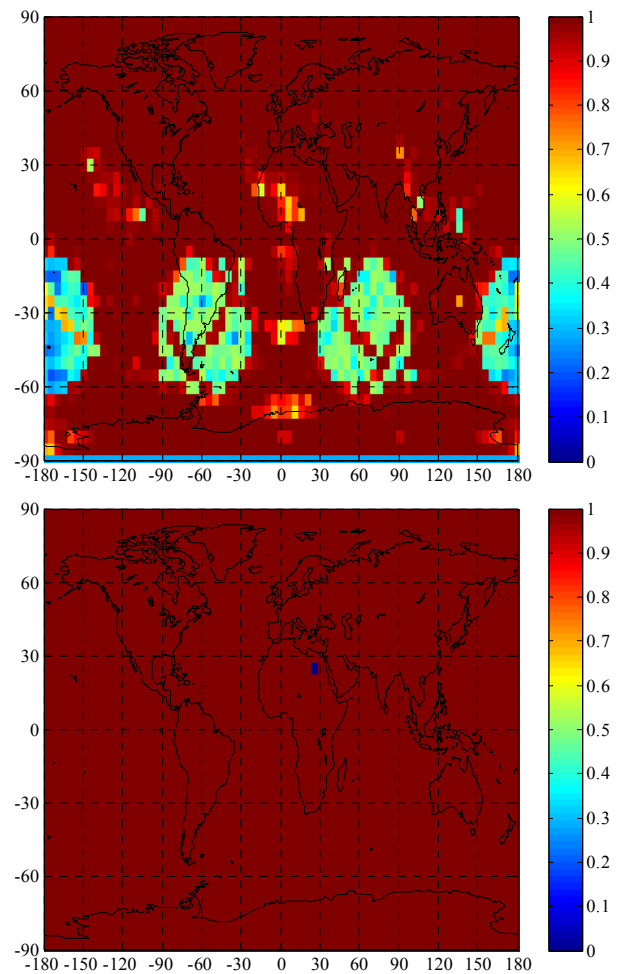


Fig. 5. Periodic reporting. Success rate for Galileo E1 (up) and for Galileo E1+E5a (below).

The dual band case continues to demonstrate excellent performance globally independent of time. This can be verified from the lower graphs in Figs. 5 and 6, respectively.

Finally, in Fig. 4 it was shown that the combined GPS L1 + Galileo E1 shows major improvement over the single-GNSSs case in the single-shot situation. Repeating the same analysis for streaming shows that increasing the

number of available observations yields high success rate (above 99.9%) independent of time. The finding is supported by the literature (Verhagen, 2006b). Once again, the increased availability of signals is identified as the single most important factor.

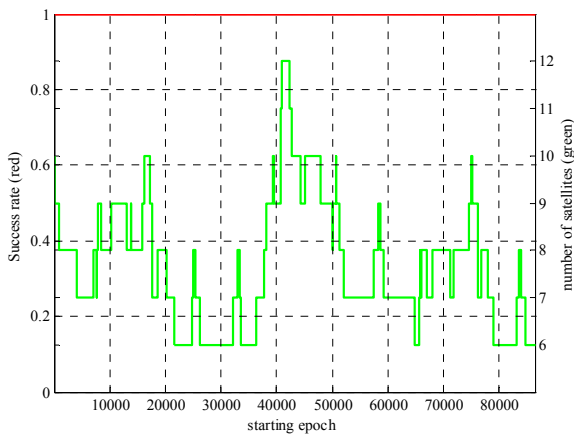
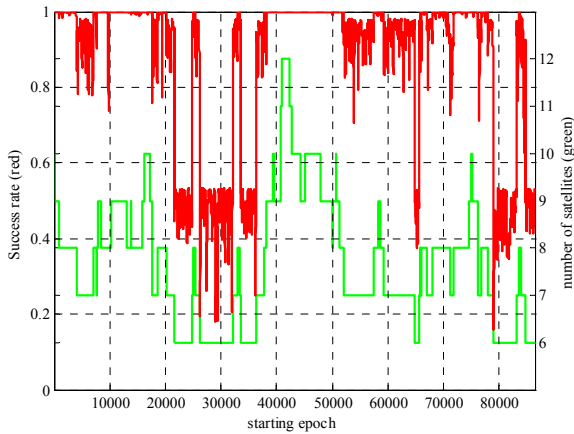


Fig. 6. Periodic reporting. Success rate for Galileo E1 (up) and for Galileo E1+E5a (below).

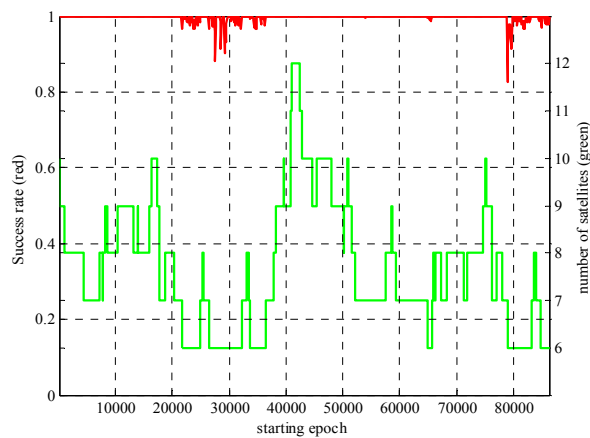


Figure 7a. 100-s spacing between measurements.

7 Measurement update rate

From the bit consumption point of view the most important issue is the measurement update rate, i.e. how often the terminal is required to report the signal measurement to the other receiver or server (or vice versa). This is analyzed by fixing the measurement period to 100 s and varying the measurement interval. The parameters and the assumptions of the analysis are as before, signals used are Galileo E1+E5a and the measurement rates in Fig. 7 a-d are

Fig 7a: a signal measurement every 100 s for 100 s (in total 2 measurements)

Fig 7b: a signal measurement every 50 s for 100 s (in total 3 measurements)

Fig 7c: a signal measurement every 20 s for 100 s (in total 6 measurements)

Fig 7d: a signal measurement every 10 s for 100 s (in total 11 measurements)

The simulations show that the 20-s measurement spacing yields a constant >99% success rate. Therefore, it is deduced that the measurement interval shall not exceed 20 seconds in periodic reporting.

There is also another issue supporting this view. Once the ambiguities have been fixed, the baseline will be tracked using the solved ambiguities. The 20-s measurement spacing requires that in order to be able to update the baseline continuously, the measurements from the sending receiver must be extrapolated for 20 seconds. Note, however, that this is possible only if the sending receiver is stationary. This is the case if the sending receiver is, for example, an LMU.

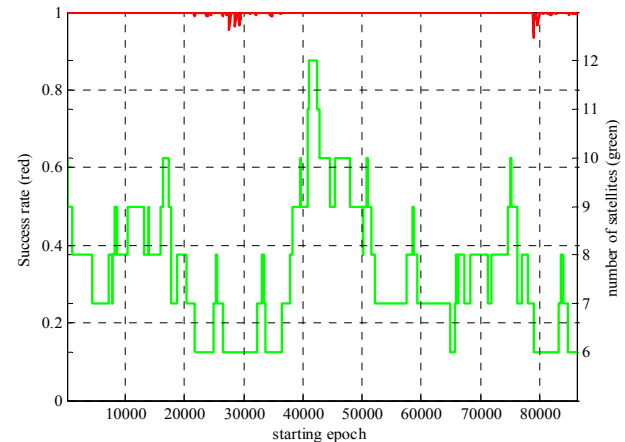


Fig. 7b. 50-s spacing between measurements.

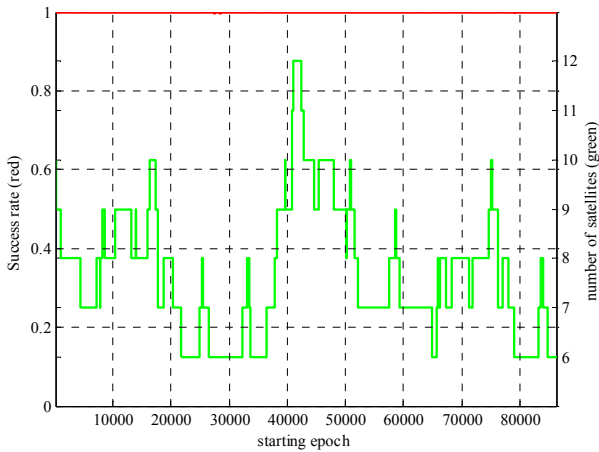


Fig. 7c. 20-s spacing between measurements.

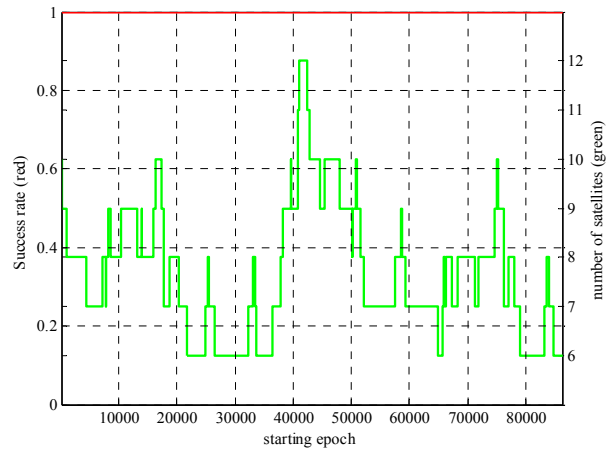


Fig. 7d. 10-s spacing between measurements.

Schüler (Schüler, 2006) reports that 30-s extrapolation leads to 35-mm RMS error in the baseline as compared to a case without extrapolation. However, the article recommends using 5-s - 10-s spacing for the best balance between bandwidth consumption and performance. Accepting errors of few tens of millimeters allows for extending the spacing to 20-s, which was considered maximum interval from the success rate point of view.

longer widelane has on the resolution. This is shown, for instance, in results for Galileo E5a+E5b. Moreover, when using widelane combinations, one must ensure that ¹real advantage can be gained by using them and that ²wide- and narrowlane ambiguities can be decorrelated to such extent that they can be solved. For more discussion see (Teunissen, 1997).

8 Analysis of different systems

Fig. 8 shows an analysis of ambiguity fixing success rates over one day for single-epoch fixing attempts (i.e. only one instant of time used). The height of the bar indicates the span of the success rate over the day and the black dot the average success rate. The blue bars on the left are for GPS, the red bars in the middle for Galileo and the green bars for GPS+Galileo hybrid. The method of analysis is detailed in (Verhagen et al., 2007). The assumptions for baseline, time and other parameters are as before.

Firstly, comparing the blue and red bars in Fig. 8 shows that Galileo outperforms GPS in single- and multi-band cases. This is attributable to a greater number of satellites in the Galileo constellation as well as to higher orbit altitude. Both these contribute to a greater number of visible satellites and, therefore, receivable signals.

In the literature it is often stated that selecting frequencies close to each other yields a longer widelane and, hence, improved ambiguity resolution. This is evident, for instance, in results for GPS L1+L2 and L1+L5, in which L2 is closer to L1 in frequency than L5. Consequently, GPS L1+L2 outperforms L1+L5. However, there is a limit to which this effect can be exploited. In all the widelane combinations noise is amplified by a factor that is dependent upon the frequencies. Now, if the frequency separation becomes sufficiently small, the noise amplification becomes dominant over the effect that a

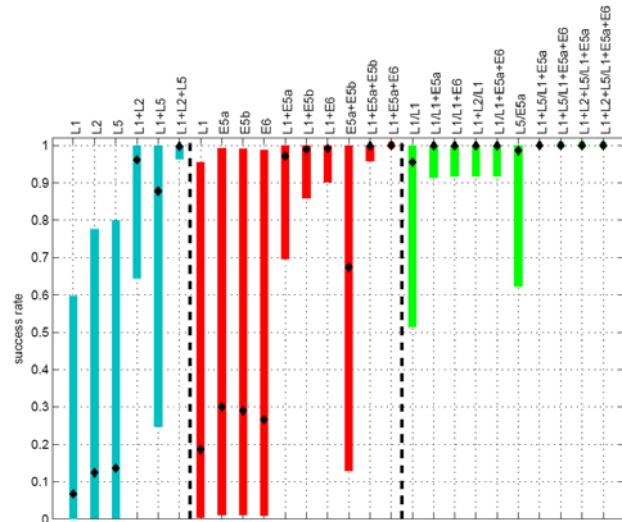


Fig. 8. Single-epoch success rates over one day. Black dot denotes the mean value and the bar the span of success rates over the day. Blue GPS, red Galileo, green hybrid.

Another finding is that the dual-GNSS cases clearly outperform the single-GNSS cases. This is true across all the signal combinations. The main benefit from Galileo is in fact the increase in the number of satellites/signals available for carrier phase -based positioning. However, considering the Galileo-only situation, (Verhagen, 2006a) shows that due to constellation differences, Galileo E1+E5a or E1+E6 performs substantially better at low latitudes than GPS L1+L5 or L1+L2, but at other latitudes no significant differences are observable.

Yet another result visible in Fig. 8 is that adding a third frequency to the solution does not have significant impact on the average success rate, but its span decreases (minimum success rate increases). Hence, a triple-frequency solution has impact on quality-of-service as well as service availability although the average success rate is not affected. Moreover, Richert (Richert et al., 2005) states that the success rate for validation improves significantly as the third frequency is taken into account.

9 Single-frequency field measurement results

Fig. 9 shows field test results for GPS L1 taken 8th January 2007 in Tampere, Finland (61.5° N, 23.7°E) for 300-m and 3600-m baselines, respectively. The number of satellites used varied from 8 to 10.

The code and carrier phase measurements from two GPS measurement engines were double differenced and fed to an extended Kalman filter. Integer ambiguities were solved using the LAMBDA-algorithm using discriminator as the validator with a threshold value of 3 (Tiberius, 1995). Neither ionosphere nor troposphere was modeled and no a-priori model of atmosphere was used.

In the example given the measurement rate was 1 Hz and the time is counted from the beginning of the session. In the beginning of the session the receivers have all the visible satellite stably in track.

It should be noted that if a success rate analysis was made for the current case, the success rate would be very high due to great number of measurements (1 Hz rate). In fact, in the current field tests the ambiguity solution converged relatively quickly, but the solution was validated at 53 and 25 seconds, respectively. As pointed out earlier, the small number of signals (frequencies) makes the validation of the ambiguities challenging (Richert, 2005). This was also confirmed in the reported field tests.

The results show that, when feasible, single-band carrier phase -based positioning is capable of producing cm-level baseline accuracy. On the other hand, the results also show that since with single-frequency measurements it is not possible to compensate for atmosphere without an externally supplied model, there is a cm-level drift in the baseline coordinates. It is assumed that this is due to tropospheric conditions, because the changes are quite slow.

Consider then the accuracy of the baseline, when the integer ambiguities are not or cannot be fixed or validated. In such a case the float solution can be utilized as opposed to the fixed solution. Fig. 10 shows data from the 300-m baseline, which is the same case as in the upper graph in Fig. 9. Only the time span is shorter.

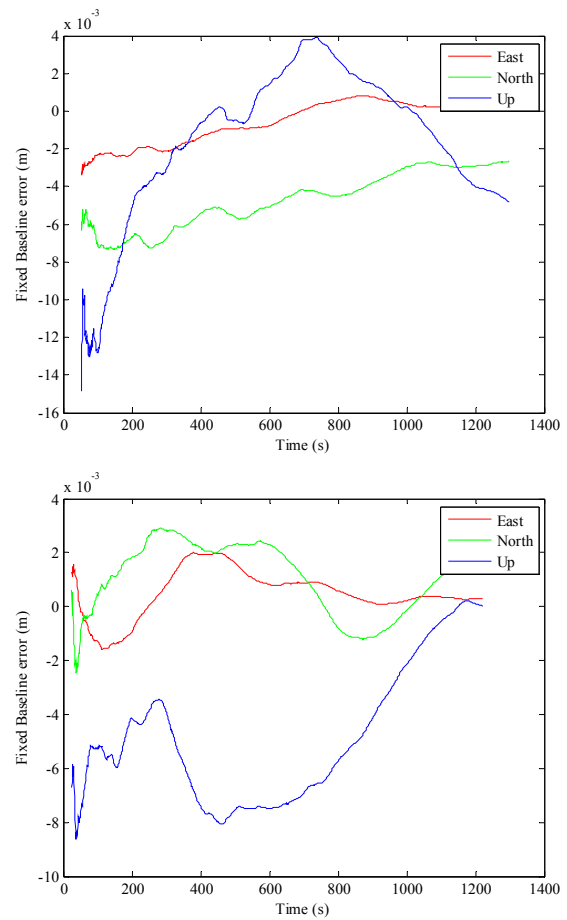


Fig. 9. GPS L1 field test results for 300-m (up) and 3600-m (below) baselines. Time is counted from the beginning of the session. Validation of the solutions took 53 and 25 seconds, respectively.

The upper graph in Fig. 10 represents the baseline obtained by differencing the standalone receiver positions. The error is in the order of several meters in all the baseline coordinates. As expected, the largest error occurs in the up-direction (approximately 5 meters). The lower graph shows the float solution. The float solution is always available (given that there are no cycle slips) and as shown, the error in the float baseline is significantly smaller than in the baseline obtained by differencing the two positions. After 30 seconds from the beginning of the session the errors in the float baseline coordinates are already in the order of 20 cm. Hence, although ambiguity fixing is not nearly always possible in the single-frequency case, the float solution, which is readily available, can improve accuracy significantly.

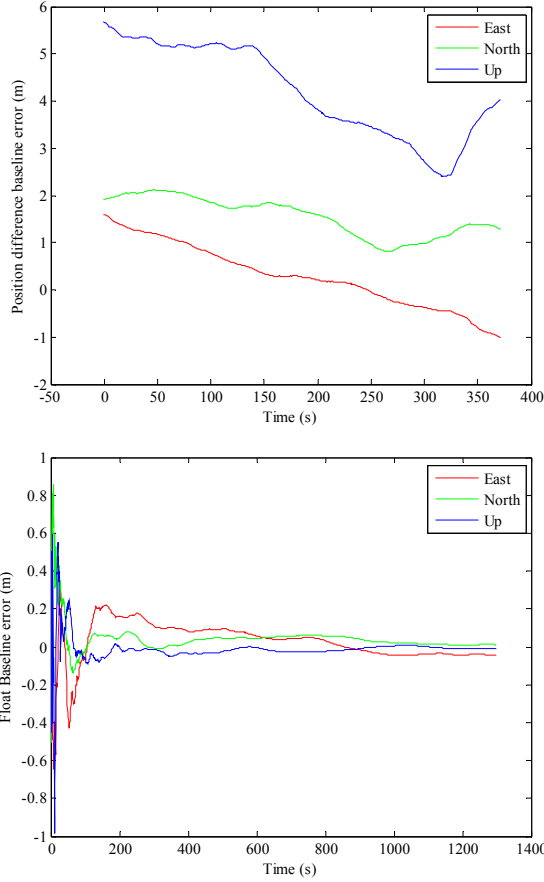


Fig. 10. GPS L1 results for the 300-m baseline. Up: Accuracy obtained using the difference of the receiver positions. Below: Accuracy of the float solution.

10 Bandwidth requirements

The data required for carrier phase -based positioning include

- Time of measurements
- Reference location for the measurements
- Code phase measurements and uncertainties
- ADR measurements, uncertainties and continuities

In the current 3GPP standard releases there are fields for transferring time of measurement, reference location as well as code phase measurements from the AGNSS assistance server to the terminal. The missing fields are ADR (Accumulated Delta Range, or Integrated Doppler), ADR uncertainty and ADR continuity indication.

ADR measurements differ from other measurements in a respect that the range required for the measurement depends upon the reporting interval. This is because of the cumulative nature of the ADR measurement. The

requirement for the range is that it must be greater than four times the maximum increase (or decrease) in ADR over the maximum measurement interval. The condition arises from the need to identify the ADR roll-overs and as the condition is fulfilled, the receiving end is capable of detecting the ADR roll-overs. Therefore, the receiver is capable of reconstructing the original measurement by examining the two upper bits of the previous and current ADR measurements. Hence, the number of bits (b) required for representing the ADR measurement fulfilling the range requirement can be given by

$$4 \cdot \max_t |\partial_t ADR(t)| \cdot T < 2^b \Rightarrow \quad (1),$$

$$b = \left\lceil \frac{\ln(4 \cdot \max_t |\partial_t ADR(t)| \cdot T)}{\ln 2} \right\rceil$$

where $ADR(t)$ the time-varying ADR measurement in meters and T the measurement interval in seconds. Moreover, the resolution of the measurement must be (at least) 1 mm resulting in a requirement to have additional 10 bits ($2^{-10} \text{ m} < 1 \text{ mm}$) for the decimal part.

Now, if the increase (decrease) rate of the ADR would depend solely on the movement of the satellite, one would have for a static GPS-receiver on the surface of the Earth (Parkinson, 1996)

$$\max_t |\partial_t ADR(t)| < 930 \frac{\text{m}}{\text{s}}. \quad (2)$$

Galileo (3000 km higher orbit than GPS - slower orbital velocity) and QZSS (geostationary) have smaller Doppler frequencies than GPS. On the other hand, GLONASS (~1050 km lower orbit than GPS) has 30 m/s greater maximum Doppler than GPS. Hence, 970 m/s is taken as the maximum rate of increase (decrease). However, one must also consider ¹⁾the receiver movement and ²⁾the receiver oscillator frequency error. The receiver movement can be assumed to contribute at maximum 50 m/s. The receiver oscillator stability is assumed to be better than 1 ppm. Hence, the maximum (apparent) Doppler resulting from this is $2 \cdot 1\text{ppm} \cdot c < 600 \text{ m/s}$. Therefore, the maximum absolute ADR rate of increase (decrease) is set to $(970 + 50 + 600) \text{ m/s} < 1620 \text{ m/s}$. The bit consumption based on equation 1 as a function of T taking the decimal part (10 bits) into account is summarized in table I.

In addition to the ADR measurement, carrier phase -based position also requires indication of the measurement continuity as well as on the quality (variance of the measurement). The ADR measurement continuity is defined by 1 bit, which indicates, whether the ADR measurement has been continuous between the current and the previous measurement messages. One bit is sufficient, since the protocols used guarantee that packets arrive in the correct order and that no packets are lost in the transmission channel.

The measurement quality is coded according to the RTCM standard (RTCM, 1998) using a three-bit field and a table mapping the values to ADR measurement uncertainty.

Note that it is also implicitly assumed that the ADR measurement has been corrected for the data bit polarity. Hence, there is no need to transfer the data bit polarity flag between the receivers. Moreover, although there is a field for code phase measurements, it has a resolution of approximately 300 m. This is not sufficient for carrier phase -based positioning. Hence, additional 10 bits are required to increase its resolution down to approximately 0.3 m ($\approx 300 \cdot 2^{-10}$ m).

Therefore, from the bandwidth point of view ADR measurements add some load to the network, but the load can be optimized as shown. The study shows that the reporting interval should be at maximum 20 s, which results in $27+1+3+10=41$ additional bits per each signal. Considering an extreme case of 2 bands, 2 GNSSs and 8 satellites per GNSS (corresponding to 32 signals) the average bit rate is $32 \cdot 41 \text{ b} / 20 \text{ s} = 66 \text{ bps}$.

Table 1. Bits required for a single ADR measurement for different reporting intervals.

T (s)	bits
1	23
5	25
10	26
20	27

11 About ionosphere modelling

Carrier phase -based positioning benefits significantly from ionospheric modelling. Due to the dispersive nature of ionosphere, phase advance may be estimated, if there are measurements on more than one frequency. However, Richert (Richert et al., 2005) reports that even in a multi-band case it is still advantageous to have a-priori estimate for the advance from an external source. If there is no a-priori information available, the solution is potentially unstable. Moreover, Odijk (Odijk, 2000) reports that ionosphere modelling is essential for long-baseline applications, even if using dual-band GPS measurements.

The common element in the new AGNSS standard provides an opportunity to provide the terminal with an ionospheric model (Syrj arinne et al., 2006). Moreover, the architecture shown in Fig. 1 enables such a service by providing an interface to external services generating such ionospheric predictions. Such a source is, for instance, DLR (Deutsches Zentrum f ur Luft- und Raumfahrt), which can provide space weather forecasts (Jakowski et al., 2002). Providing an accurate ionosphere

model contributes significantly on the feasibility of the single-band carrier phase -based positioning.

12 Challenges

The specific challenges to be addressed before carrier phase-based positioning can be added to the cellular standards include, amongst others, the handovers from one serving base station to the other. The carrier phase -measurement need to be continuous over the hand-over, which introduces additional book-keeping exercise to the network. However, if a Virtual Reference Station is used, the terminal can change the VRS without losing the baseline. This can be achieved by subscribing two VRS data streams to the terminal, solving the three baselines (VRS-VRS and 2x VRS-terminal) and discarding the old VRS once the baseline between the new VRS and the terminal has been established. While such an approach is feasible in the user plane, it is difficult to implement in the control plane of the cellular network.

Another concern is the definition of the quality-of-service. The *minimum performance requirements* for Assisted GPS (3GPP-TS-34.171) guide the design and implementation of the terminal. When introducing carrier phase -based positioning to the standards, it must be introduced as a new *positioning method* and similar minimum performance requirements may be required for the new method. Such work requires deep understanding of the use cases as well as the full potential of the technology and extensive field testing. There is currently no work towards such performance requirements.

13 Conclusions

The carrier phase-based positioning has the potential to bring the positioning accuracy down to centimetres. Therefore, it is tempting to consider adding the support for carrier phase-based positioning to the cellular standards.

The analyses presented in this paper show that the most significant problem with single-frequency carrier phase-based positioning is the uncertainty about its performance. The simulations show that during a day there are brief periods during which the carrier phase-based positioning is feasible, but at other times the performance can be expected to be very poor. The lack of measurements (satellites) is the most significant factor contributing to the lack of performance. In conclusion, single-frequency carrier phase-based positioning is not feasible, if there is only one GNSS available and if ambiguities need to be fixed. However, already the float solution, which is always available given that there are no undetected cycle slips, was shown to be a major improvement over traditional point positioning. It was

also shown that the single-frequency case becomes very interesting with the introduction of additional GNSSs (Galileo, GLONASS) to complement GPS.

The study also shows that the full potential of Galileo lies in the use of the various available signals. If future terminals are capable of utilizing, for instance, both GPS L1 + Galileo E1 as well as GPS L5 + Galileo E5a (since they are in the same band, respectively) carrier phase -based positioning is no doubt an attractive addition to the current set of positioning methods. However, this requires that the terminals are capable of multi-GNSS multi-band reception and that the cellular standards/protocols support the periodic reporting of ADR measurements from the network to the terminal and/or vice versa.

It was also shown that the capability can be achieved with small additions to the current standards. The average additional data transfer load was shown to be in the order of 66 bps even when there are several GNSSs and signals available. The resulting accuracy is in the order of centimetres in the best case and, hence, it is believed that the implementation task and additional network load is justified.

References

- 3GPP-TS-23.271 *Functional Stage 2 Description of Location Services (LCS)*, <http://www.3gpp.org>.
- 3GPP-TS-25.331 *Radio Resource Control (RRC) protocol specification*, <http://www.3gpp.org>.
- 3GPP-TS-34.171 *Terminal conformance specification; Assisted Global Positioning System (A-GPS)*, <http://www.3gpp.org>.
- 3GPP-TS-44.031 *Radio Resource LCS (Location services) Protocol (RRLP)*, <http://www.3gpp.org>.
- 3GPP (2006) *Meeting report: Report of TSG GERAN meeting#32*, Sophia-Antipolis, France, 13th-17th November, <http://www.3gpp.org>.
- 3GPP (2007) *Meeting report of the 36th 3GPP TSG RAN meeting*, Busan, Korea, 29th May - 4th June, <http://www.3gpp.org>.
- Alcatel, Ericsson, Nokia, Qualcomm, Siemens Networks, SiRF (2006). *GP-062472 A-GNSS GERAN#32 status*. Presented in 3GPP GERAN2#32, 13th-17th October, Sophia Antipolis, France.
- Alanen K., Wirola L., Käppi J. and Syrjärinne J. (2006a) *Inertial Sensor Enhanced Mobile RTK Solution Using Low-Cost Assisted GPS Receivers and Internet-Enabled Cellular Phones*. In Proceedings of IEEE/ION PLANS 2006, 25th-27th April, San Diego, CA, USA, pages 920–926.
- Alanen K., Wirola L., Käppi J. and Syrjärinne J. (2006b) *Mobile RTK using low-cost GPS and Internet-Enabled Wireless Phones*. InsideGNSS, pages 32–39, May-June issue.
- Dow J.M. and Neilan R.E. (2005) *The International GPS Service (IGS): Celebrating the 10th Anniversary and Looking to the Next Decade*. Advanced in Space Research, 36(3):320–326.
- Eissfeller B., Tiberius C., Pany T., Biberger R. Schueler T. and Heinrichs G. (2002a) *Instantaneous ambiguity resolution for GPS/Galileo RTK positioning*. Journal for Gyroscopy and Navigation, 38(3):71–91.
- Eissfeller B., Tiberius C., Pany T. and Heinrichs G. (2002b) *Real-Time Kinematic in the light of GPS Modernization and Galileo*. Galileo's World, Autumn issue.
- Jakowski N., Heise S., Wehrenpfennig A. and Schlüter S. (2002) and R. Reimer. *GPS/GLONASS-based TEC measurements as a contributor for space weather forecast*. Journal of atmospheric and solar-terrestrial physics, 64:729–735.
- Klimov V., Revnivykh S., Kossenko V., Dvorkin V., Tyulyakov A. and Eltsova O. (2005) *Status and Development of GLONASS*. In Proceedings of GNSS-2005, 19th-22nd July, Munich, Germany.
- Lundgren D. and Diggelen F. (2005) *Long-Term Orbit Technology for Cell Phones, PDAs*. GPSWorld, pages 32–36. October issue.
- Nokia (2006) *GP-061215 Justification for the addition of carrier phase measurements*. Discussion paper, presented in 3GPP TGS-GERAN meeting#30, 26th-30th June, Lisbon, Portugal.
- Odiijk D. (2000) *Weighting Ionospheric Corrections to Improve Fast GPS Positioning Over Medium Distances*. In Proceedings of Institute of Navigation GPS 2000, 19th-22nd September, Salt Lake City, USA.
- OMA-ULP OMA-TS-ULP-VI-0-20050719-C, User Plane Location Protocol, <http://www.openmobilealliance.org>.
- OMA (2007) *Open Mobile Alliance Location Working Group meeting minutes OMA-LOC-2007-0290-MINUTES_20Aug2007Seoul*, Seoul, Korea, 20th-24th August, <http://www.openmobilealliance.org>.
- Parkinson B. and Spilker J. (1996) *Global Positioning System: Theory And Applications Volume I*. American Institute of Aeronautics and Astronautics, Inc. Washington DC, USA.
- Richert T. and El-Sheimy N. (2005) *Ionospheric modeling - The Key to GNSS Ambiguity Resolution*. GPS World, pages 35–40, June issue.
- RTCM (1998) *Recommended Standards for differential GNSS Service*, version 2.2. RTM Special Committee no 104. January 15th, Alexandria, Virginia, USA.
- Schüler T. (2006) *Interpolating Reference Data - Kinematic Positioning Using Public GNSS Networks*. InsideGNSS, pages 46–52, October issue.
- Strategy Analysts (2006) *Global Handsets, GPS/A-GPS Phone Sales*.

- Syrjärinne J. and Wirola L. (2006) *Setting a New Standard - Assisting GNSS Receivers That Use Wireless Networks*. InsideGNSS, pages 26–31.
- Teunissen P. (1997) *On the GPS widelane and its decorrelating property*. Journal of Geodesy, 71:577–587.
- Teunissen P. and Tiberius C. (2000) *Bias Robustness of GPS Ambiguity Resolution*. In Proceedings of Institute of Navigation GPS 2000, 19th-22nd September, Salt Lake City, USA.
- Tiberius C. and Jonge P. (1995) *Fast Positioning Using the LAMBDA-Method*. In Proceedings of the 4th International Symposium on Differential Satellite Navigation Systems (DSNS), 24th-28th April, Bergen, Norway, pages 1–8.
- Tiberius C., Pany T., and Eissfeller B. (2002) *Integral GPS-Galileo ambiguity resolution*. In Proceedings of ENC-GNSS2002, May 17th-30th, Copenhagen, Denmark.
- Verhagen S. (2006a) *How will the new frequencies in GPS and Galileo affect carrier phase ambiguity resolution?*, InsideGNSS, pages 24–25, March issue.
- Verhagen S. (2006b) *Manual for Matlab User Interface VISUAL*. Delft University of Technology, The Netherlands.
- Verhagen S., Teunissen PJG. and Odijk D. (2007) *Carrier-phase Ambiguity Success-rates for Integrated GPS-Galileo Satellite Navigation*. In Proceedings Of Joint workshop WSANE2007, 16th-18th April, Perth, Australia.
- Wirola L., Alanen K., Käppi J. and Syrjärinne, J. (2006) *Bringing RTK to Cellular Terminals Using a Low-Cost Single-Frequency AGPS Receiver and Inertial Sensors*. In Proceedings of IEEE/ION PLANS 2006, 25th-27th April, San Diego, CA, USA, pages 645–652.
- Wirola L. and Syrjärinne, J. (2007a) *Bringing All GNSS into Line*. GPS World, 18(9):40–47.
- Wirola L. and Syrjärinne, J. (2007b) *Bringing the GNSSs on the Same Line in the GNSS Assistance Standards*. In Proceedings of the 63rd ION Annual Meeting2007, 23rd-25th April, Boston, MA, USA, pages 242–252.

Publication 3

L. Wirola and I. Kontola and J. Syrjärinne. The Effect of the Antenna Phase Response on the Ambiguity Resolution. In *Proceedings of IEEE ION PLANS 2008*, Monterey, USA, May 6th-8th, pp. 606–615, (2008)

Copyright ©2008 IEEE. Reprinted, with permission, from the proceedings of IEEE ION PLANS 2008.

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the Tampere University of Technology's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this material, you agree to all provisions of the copyright laws protecting it.

The Effect of the Antenna Phase Response on the Ambiguity Resolution

Lauri Wirola
Nokia Inc.

Ilkka Kontola
Nokia Inc.

Jari Syrjärinne
Nokia Inc.

Abstract—In order to get the best performance from carrier phase -based GNSS positioning methods in terms of accuracy and reliability the factors affecting the signal propagation must be characterized accurately. These carrier phase -based methods include Precise Point Positioning (PPP) as well as Real-Time Kinematic (RTK). While much focus has been put on atmospheric effects, the antenna effects are either ignored (low-end solutions) or handled by utilizing phase center offset and phase center variation (high-end solutions). The latter approach is typical in modern RTK equipment.

Survey-grade antennas are designed to have such fine azimuthal symmetry in the phase response that only elevation-dependent correction must be applied to the observations. This is referred to as the phase center variation. Moreover, the final baseline solution is corrected with the phase center offset in order to map the solution to a physical point in the antenna structure. The approach typically assumes that antennas of the same type have similar spatial response characteristics so that the same correction data can be applied to all the antennas of the same make.

However, carrier phase -based techniques have been proposed for consumer-grade devices, in which the antennas are typically cheap, small and unoptimally positioned in the devices. In such cases the phase response may have high asymmetry both in azimuth and elevation and, hence, the current practices may no longer be sufficient. The unmodelled biases, amongst other, have impact on the probability of successful integer ambiguity fixing in RTK.

This paper characterizes three antennas designed for GPS L1 reception in terms of their magnitude and phase responses as a function of azimuth and elevation of the signal source. Two of the measured antennas were patches mounted in Bluetooth™GPS -receivers and one antenna was Trimble Bullet™III that was measured for reference purposes. The phase responses are analyzed in the context of phase center offset and variation. The phase responses are then utilized in estimating the statistics of ambiguity fixing success rates.

The measured antennas show varying performance in terms of phase response symmetry. The patches mounted in Bluetooth devices show approximately 70- and 49-degree variation in the phase response depending upon the direction of the signal. The lack of azimuthal symmetry prohibits the use of only elevation-dependent phase center variation tables and suggests the need for a full 3D table. The two antennas also show such differing responses that the use of a single PCV table for the antennas is not feasible. The bullet, however, shows only 4-degree variation and, hence, fine symmetry.

Finally, even though the absolute variations in the phase responses are quite significant in antennas mounted in a Bluetooth GPS, the simulations show that these variations do not have a significant effect on the success rates for ambiguity resolution. This is because the probability of having a significant double difference bias turns out to be practically negligible.

I. INTRODUCTION

The integer ambiguity resolution is the most important step in RTK applications. Typically resolution is a two-phase process. Float ambiguities and the corresponding covariance matrix are solved using standard least-squares methods in the first phase. In the second phase the float ambiguities and the corresponding covariance matrix are used as an input in an algorithm for fixing the integer ambiguities. There are several methods available for the resolution, such as the Least-squares Ambiguity Decorrelation Algorithm (LAMBDA) [1].

Unmodelled biases in carrier phase -measurements lead to unaccounted errors in double difference observables. Such biases are readily introduced by ionospheric advance [2] and carrier phase multipath [3]. Consequently, these biases result in float ambiguity biases, which lead to a decrease in the integer ambiguity resolution success rate [4]. The decrease in the success rate is even more profound in cases, in which the biases are not taken into account in the stochastic model. Hence, the biases must be identified and preferably modeled or mitigated.

Biases on carrier phase -measurements, and subsequently on float ambiguities, may also be induced by anisotropic phase response of the receiving antenna. An antenna may be considered as a filter having a complex frequency response that is a function of not only frequency, but also of the direction of the signal source. Hence, a varying phase bias is introduced to the carrier phase -measurements depending upon the position of the satellite with respect to the antenna orientation. In an RTK application this is not an issue if the antennas are identical and identically oriented, because in such a case the biases cancel in double differencing. Alternatively, the biases may be mitigated, if the antenna response characteristics is known. However, should the rover and reference antennas be of different type, antenna orientations not identical or response unknown, biases are introduced in double difference -observables and, ultimately, in float ambiguities.

Antenna-induced phase and group delay errors have widely been documented in the literature. For instance, [5] considers, amongst other, quadrifilar helix and microstrip antennas and their phase responses. Further, [6] presents a compensation method for antenna arrays, in which conventional methods may fail. By conventional the authors mean an approach, where the spatial phase response is measured and used to correct the measurements. With adaptive antenna arrays the

approach fails due to antenna pattern being a function of the antenna weights that are constantly adjusted. Moreover, [7] studies phase errors in different microstrip antenna configurations. Finally, [8] characterize a few widely used geodetic antennas in terms of their spatial phase response.

Typically antennas used in high-accuracy GNSS positioning are massive, expensive and they, for instance, utilize chokes to mitigate multipath. These antennas are also designed to have as even phase responses as possible especially in azimuth, i.e. the phase error is independent of azimuth and, therefore, only a function of the source elevation. The elevation-dependent component can be compensated for using phase center variation (PCV) tables readily available for various geodetic antennas. Moreover, alongside with PCV tables also the phase center offset (PCO), the 3D vector between the phase center and the antenna reference point, are available. Having knowledge of the PCO and PCV allows for obtaining repeatable GNSS positioning and baseline results irrespective of the antenna used. [9]

However, the high-accuracy methods are also being introduced to wider audience via the use of AGNSS-enabled (Assisted GNSS) handsets. The annual sales of AGNSS handsets is estimated to rise to 400 million units by 2011 [10] with the current market size being about 100 million units. The boom in the location-based services market requires developing constantly more efficient and capable methods to improve user experience in terms of, for example, accuracy [11]. For instance, RTK-type solutions are considered as a potential method to introduce increased accuracy in handheld GNSS devices by utilizing VRS services (Virtual Reference Station) [12] [13].

Antennas in either AGNSS-enabled terminals or Bluetooth™GPS receivers are typically small, cheap and placed in electrically as well as magnetically active environment. It is therefore the purpose of this study to examine, if the methods used in the geodetic-community (PCO with elevation-dependent PCV) are sufficient to characterize these consumer-grade antennas. In practice this requires measuring complex spatial responses and studying, how the phase responses behave as a function of elevation and azimuth. In this study three patch-type antennas are measured and their responses analyzed especially in the context of PCV.

Moreover, because this study aims to lay foundation on the use of RTK in low-end devices, phase responses are further used to examine, how asymmetric phase responses, if not compensated for, bias the double-differenced phase observables. Ultimately, the degradation of the ambiguity resolution success rate due to antenna-induced biases is characterized for the three measured antennas. This impact is not sufficiently documented in the literature.

The results are significant in the current context of low-end devices, because in addition to potentially highly asymmetric responses, the antenna characteristics may also have low repeatability between devices even though the antenna type would be the same. Low repeatability makes methods utilizing fixed PCO+PCV useless. Moreover, in mass-production

environment it is impossible to individually characterize each antenna. It is, therefore, important to understand all the effects that uncompensated antenna effects may have in high-accuracy positioning. The degradation of the ambiguity resolution is the first studied such effect.

II. BACKGROUND

The authors of [7] define the phase delay error Φ by

$$\Phi(\theta, \phi, \omega) = \Psi(\|\underline{r}\|, \theta, \phi, \omega) - (-\underline{k}^T \underline{r}), \quad (1)$$

where \underline{r} is the location of the signal source, θ and ϕ azimuth and elevation coordinates respectively, ω frequency, Ψ the total phase at the end of the feed cable and \underline{k} the wave vector. The origin of the coordinate system is at the physical center point of the antenna.

In the GNSS community it is a general practice to refer to the antenna *phase center* \hat{r} as the reference point for high-accuracy measurements. Or more precisely, the measurements corrected with phase center variation tables are referred to the phase center, which is then translated to the antenna reference point by using the phase center offset vector. [9]

In order to fully characterize the phase center one must understand the concept intuitively, be able to formulate it formally as well as to express, how the location of the phase center can be measured. While the formal definition is missing, intuitively the phase center is the point seen from the infinity, from which point the spherical waves originating from the antenna seem to be radiating from. The intuition implicitly assumes that the equiphase surface is a sphere in the infinity.

Another way to understand the concept is to imagine a point around which no rotation changes the phase of the signal given that the distance from the signal source to the point is unchanged. However, an antenna does not need to have a well-behaving phase center and in general such a point cannot be found [8]. Moreover, even in the cases, where such a point can be found, the residual errors may still be large [5]. The residuals are accounted for by the PCV tables, which typically give the phase correction as a function of elevation independent of azimuth, because of high azimuthal symmetry shown by modern geodetic GNSS antennas [9].

From the measurement point-of-view locating the phase center is challenging. This challenge rises from the condition that the equiphase surface should be measured in the infinity. However, since this is not feasible, a technique used in this study is to fit a sphere to the phase response measured close to the antenna. The origin of the sphere is varied and the error norm defined by

$$\int_0^{2\pi} \int_0^\pi |\Psi(\|\underline{r}\|, \theta, \phi, \omega) - (-k_0)\|\underline{r} - \hat{r}\||^2 d\phi d\theta \quad (2)$$

is evaluated at each choice of origin. In essence, in this work the phase center is defined as the point, which minimizes the squared phase error over the sphere. k_0 is the wave number and $\underline{r} = \underline{r}(\|\underline{r}\|, \theta, \phi)$ in (2). Note that should the antenna be a true point source radiating from the origin of the coordinate

system, i.e. $\Psi = (-k_0)\|x\|$, then $\hat{x} = \underline{0}$ and the physical and phase centers coincide. However, in this work the actual location of the phase center is not of interest. Instead, it is the residuals that bias the carrier phase -measurements. Hence, only the residuals are reported and utilized in the analysis.

It should, however, be noted that choice of the 2-norm for the minimization may be questioned. The 2-norm was chosen, because it has previously been used in the literature [5] and due to the ease of implementation. However, future work includes studying flat 2-chain norms, which evaluate the volume between two surfaces, in the optimization. In this special case these two surfaces would be the sphere having an origin in the phase center and the measured equiphase surface.

III. ANTENNAS AND THEIR RESPONSES

In order to characterize the effect that consumer-grade antennas have on the ambiguity resolution, three antennas were measured for their complex spatial responses. Two antennas were mounted in custom-made Bluetooth (BT) AGPS (As-sisted GPS) receivers [14], or BAG, and one was Trimble BulletTMIII. All the antennas considered are patch-type and in the BAG the antenna is fed diagonally. The antennas in the BAGs are 25x25-millimeter rectangular patches with a thickness of 2 mm, dielectric substrate of $\epsilon_r = 20$ and a Q-value of 5000. The PWB of the BT AGPS works as the ground plane. Also, the internal LNA in the Bullet was bypassed. The S11 scattering parameters for all three antennas are given in Figure 1 for reference. Moreover, the antennas and the respective coordinate systems are shown in Figure 2.

The complex RHCP response measurements were performed at GPS L1 frequency using the Satimo SG128 [15] measurement system (shown in Figure 3). The system measures the full 3D complex RHCP frequency response as a function of elevation and azimuth. Measurements were taken at nominal GPS L1 frequency using a three-degree grid in elevation and azimuth coordinates.

Figure 4 shows the measurement results for the three antennas in polar plots. The left and right hand graphs represent

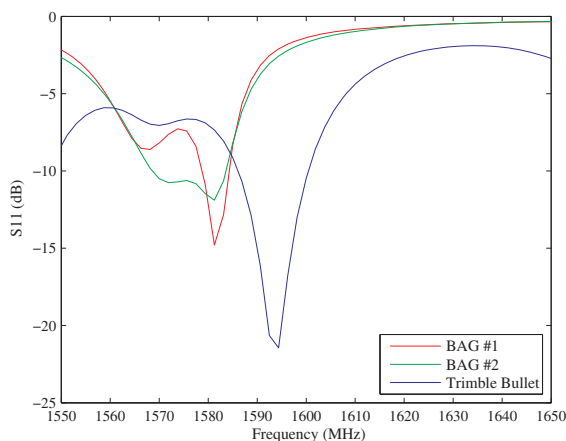


Figure 1. S11 scattering parameters for the measured antennas.

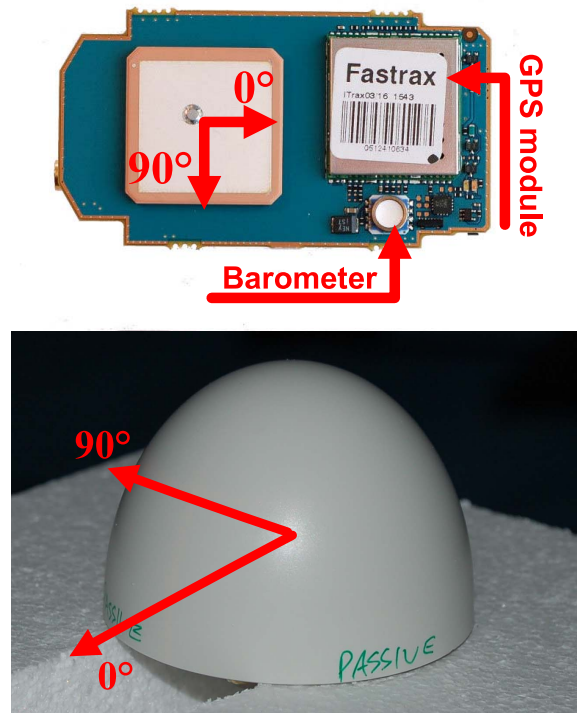


Figure 2. Definition of coordinate systems. Up: Bluetooth AGPS, Down: Trimble Bullet.

amplitude and phase responses, respectively. The coordinate system is defined in Figure 2 and the colors (circular plots) code elevation levels ranging from 15° to 85°. Amplitude responses have been scaled in such a way that the gains are with respect to the lowest gain found in the given antenna at the lowest considered elevation. The absolute gains are not, therefore, comparable between the plots. However, the gain patterns can be compared relative to each other.

The same approach has been applied to the phase responses. Phase responses shown are not absolute, but are biased in such a way that the scale starts from zero for each antenna. Therefore, the phase response variations between different antennas are comparable. Note also that the phase responses are normalized responses, from which the natural change of the phase resulting from the nature of RHCP has been removed [7]. Moreover, it should be noted the phase responses reported in Figure 4 are residuals after finding the phase centers in terms of the minimum squared-error and after the removal of the resulting equiphase spheres from the respective raw measurements.

The results show a clear contrast between the Bullet and cheap patches in BAGs. The approximate directivity of 6 dBs of rectangular patches [16] is, though, reflected in all the responses. The Bullet shows highly symmetric amplitude and phase responses. The phase response shows less than four-degree variation in any given direction. This corresponds to about 1.6 millimeters. The small variation shows that the

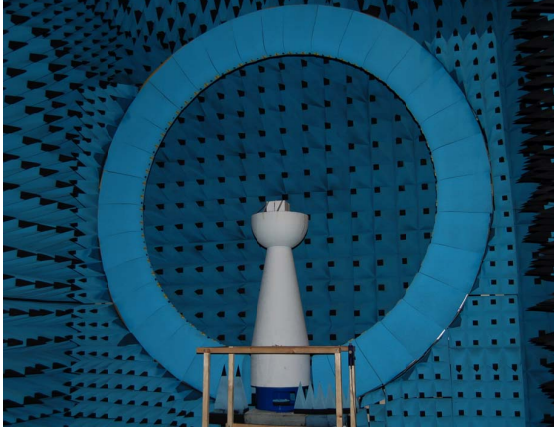


Figure 3. Satimo SG128 3D measurement system.

antenna has accurately been rotated around the virtual phase symmetry point, but also the high quality of the antenna. The standard deviation of the phase response from the best fit sphere is only 0.7° . This shows that a good-quality antenna would perform well with only a PCO correction. With an elevation-dependent PCV table the results could be improved further.

On the other hand, the antennas mounted in BAGs #1 and #2 show 70° and 49° variations in the phase responses at maximum, respectively. Moreover, the standard deviations around the best fit spheres are 8° and 6° , respectively. The magnitude responses for BAGs follow the same behavior as in the Bullet, although the asymmetries are evident in the BAG responses.

The greatest variations in the response occur in low elevations, which is in accordance with measured and theoretical results presented in [8] and [7], respectively. For instance, if with BAG #1 the elevation mask is raised to 30° or 45° , the maximum phase response variation decreases to 32° and 19° , respectively. Furthermore, the associated standard deviations decrease to 4° and 3° , respectively. The behavior is advantageous from the stochastic point-of-view, since the satellites at low elevations are typically assigned higher noise, which may therefore also cover the antenna effects to some extent.

The phase responses for BAG antennas show that the current practice of using a PCO and an elevation-dependent PCV table is not sufficient for BAGs. While the PCO vector can be found, the azimuthal asymmetry is significant. To illustrate the problem with BAGs, consider two distinct azimuths, 50° and 90° . While at 90° azimuth there is essentially no elevation-dependency in the phase response, at 50° azimuth the variation is in the order of 30° (15.9 mm at L1). Consequently, a full 3D PCV table would be needed for a BAG. Moreover, because the two BAGs show low coherence in terms of phase response, the same PCV could not be utilized for both BAGs considered. Therefore, there is a risk that in mass-market solutions the compensation of antenna-effects is not feasible in the short-

term and, hence, the impacts of uncompensated antennas on the performance must be understood.

Although the antennas in BAGs show significantly different phase behavior as compared to the Bullet, the shapes of the responses of the two BAGs still compare reasonably well. In principle, a diagonally-fed patch should show high symmetry (within one degree) with respect to the azimuth angle [7]. However, since this is not the case and the two BAGs show comparable phase patterns, the measurements in fact suggest that the structure of the BAG distorts the spatial phase symmetry in a consistent manner as anticipated. Several factors resulting in this behavior may be identified. First of all, cheap low-quality patch antennas may not be manufactured to high standards in terms of precision and repeatability. Also, the manufacturer of the patch antennas used in BAGs suggests the use of square ground plates. However, the actual ground plate, the PWB, is an elongated rectangle. Finally, there are several metallic parts surrounding the antenna (on the sides as well as below the antenna, see Figure 2). Parasitic resonances of surrounding parts may contribute to asymmetries in the phase response.

IV. DOUBLE DIFFERENCE BIASES

Carrier and code phase double differences, Φ_{km}^{ps} and ρ_{km}^{ps} respectively, between satellites p (base) and s as well as receivers k (rover) and m (reference) are defined by

$$\begin{bmatrix} \Phi_{km}^{ps} \\ \rho_{km}^{ps} \end{bmatrix} = \begin{bmatrix} \Phi_k^p & \Phi_m^p & \Phi_k^s & \Phi_m^s \\ \rho_k^p & \rho_m^p & \rho_k^s & \rho_m^s \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \quad (3)$$

where Φ_k^p and ρ_k^p are carrier and code phase measurements of the satellite p on the receiver k .

It is evident from (3) that if the rover and reference antennas are identical and similarly oriented, the antenna-induced biases are common-mode errors and cancel in double differencing. However, if the antennas have differing phase patterns or have equal phase patterns but are not oriented similarly, a bias will be introduced in the carrier phase double difference unless they are compensated for. This is due to phase errors not being common-mode errors in this case and, therefore, not cancelling in double differencing.

Figure 5 shows on the left hand side the cumulative probability distribution functions (CPDF) for carrier phase double difference biases using the antennas considered. The assumptions in deriving the distributions have been that

- rover and reference antennas are similar
- relative orientation of antennas is random
- antenna planes are parallel, i.e. antennas have no inclination with respect to each other
- minimum satellite elevation is 15°
- minimum base satellite elevation is 70°
- satellites positions in the sky are random

In effect, the simulation is run in such manner that for each relative antenna orientation, the base satellite and the other

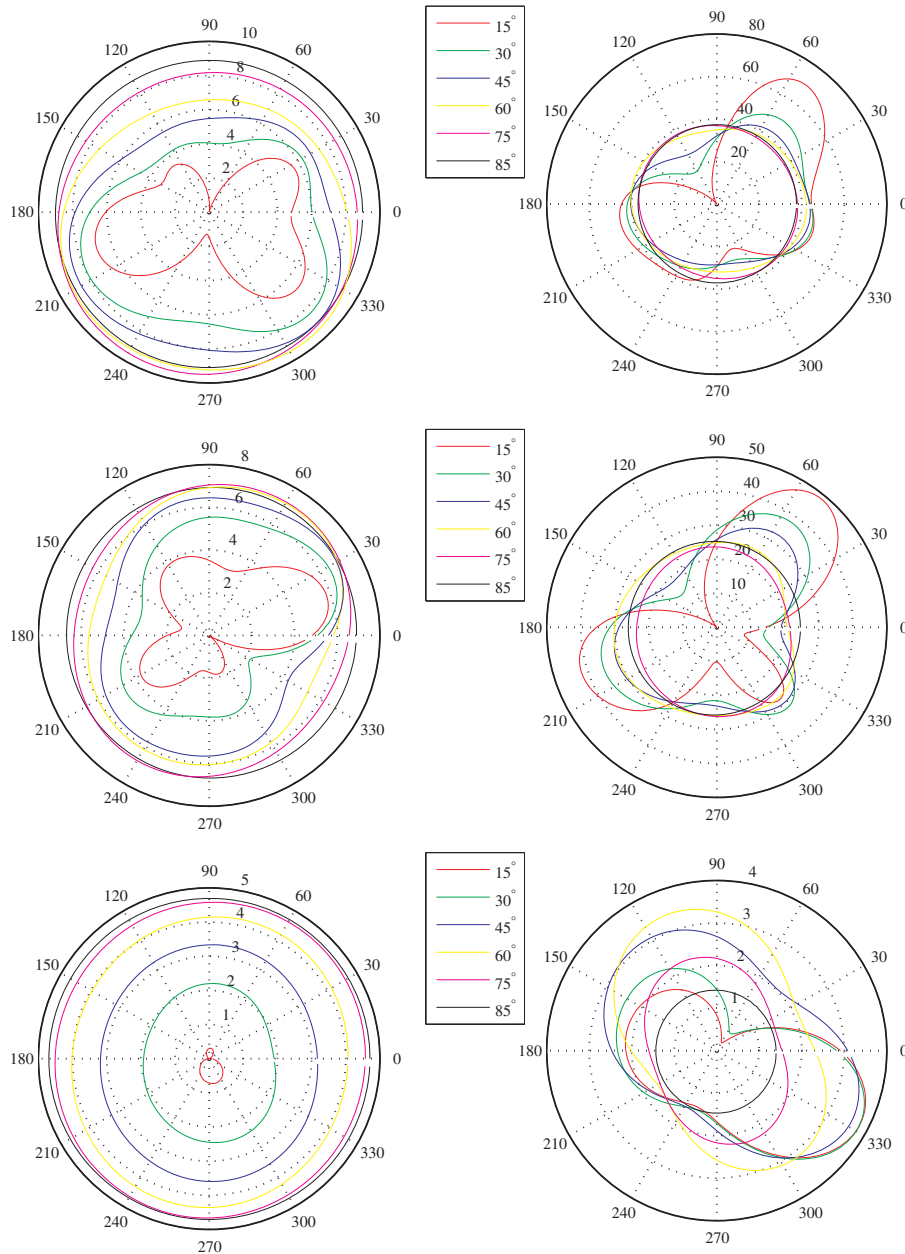


Figure 4. Left: Amplitude responses (in dBs) over fixed (15° , 30° , 45° , 60° , 75° and 85°) elevations. Right: Phase responses (in degrees) over fixed elevations. Up: BAG #1, Middle: BAG #2, Bottom: Trimble Bullet.

satellite are in every allowed location in the sky. Moreover, all the relative antenna orientations are covered. The distributions on left in Figure 5 are, therefore, symmetric due to the nature of the simulation. However, covering all the possible satellite locations and antenna orientations captures the bias expectation values.

Note that distributions for BAGs are cut at $\pm 50^\circ$ in Figure 5, although the DD biases range $\pm 75^\circ$ and $\pm 52^\circ$ for BAGs #1 and #2, respectively. This is because probabilities for high bias

values are low. Note also that in graphs for BAGs #1 and #2 x-axes extend $\pm 50^\circ$, but only $\pm 5^\circ$ in the graph for the Bullet. The CPDFs are read so that that the probability of observing a given or smaller DD bias can be read from the y-axis. The lower limits for the DD biases are -75° , -52° and -5° for BAG #1, #2 and Bullet, respectively. As a numerical example, the CPDF for BAG #1 shows that the DD bias is between -75° and 10° at the probability of 0.863.

The distributions reflect the findings discussed in chapter

III. The DD bias CPDF for the Bullet shows that only in approximately 10% of the cases the absolute bias is greater than 2° corresponding to 1 mm at L1 frequency. Because this is in the order of thermal noise for carrier phase tracking [17], it can be expected that the antenna-induced biases will not have an effect on the ambiguity resolution, when using this antenna.

The BAG antennas, on the other hand, show fairly large biases occasionally. In about 20% of the cases the absolute bias is greater than 10° (5.3 mm). A bias of this magnitude can already be easily noticed in the observables, because such bias is clearly above thermal noise. Moreover, in approximately 50% of the cases the absolute bias is less than 5° (2.6 mm). Although detectable, a bias of this magnitude can be shown not to have an effect on the float ambiguities or ambiguity resolution (see table I). Hence, degradation of the success rate can be expected to occur perhaps in the order of 20% of the time, when using BAGs.

V. FLOAT AMBIGUITY BIASES

The ambiguity resolution success rate calculation requires a float ambiguity covariance matrix $Q_{\hat{a}}$ and bias vector \hat{b} as input. The geometry-free model, though weak, is used for the first approximation [18]. For one satellite pair and one epoch the model is defined by

$$E \left(\begin{bmatrix} \Phi_{km}^{ps} \\ \rho_{km}^{ps} \end{bmatrix} \right) = \begin{bmatrix} 1 & \lambda_{L1} \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \Upsilon_{km}^{ps} \\ a_{km}^{ps} \end{bmatrix}, \quad (4)$$

where Υ_{km}^{ps} is the double difference of geometric ranges between the satellites p and s as well as receivers k and m . a_{km}^{ps} is the double difference integer ambiguity in cycles. Moreover, the covariance of carrier and code phase double difference observables is given by

$$V \left(\begin{bmatrix} \Phi_{km}^{ps} \\ \rho_{km}^{ps} \end{bmatrix} \right) = \begin{bmatrix} 4\sigma_\Phi & 0 \\ 0 & 4\sigma_\rho \end{bmatrix}. \quad (5)$$

The selected approach follows the principle introduced in [18]. The approach mimics a zero-baseline case with ρ_{km}^{ps} set to zero and the *true* value of Φ_{km}^{ps} being zero as well. However, Φ_{km}^{ps} observable is biased in the range $\pm 75^\circ$, $\pm 52^\circ$ and $\pm 5^\circ$ for BAG #1, #2 and Bullet, respectively (see the left hand side in Figure 5), in one-degree steps. Solving the weighted least-squares problem in (4) yields the float ambiguity \hat{a}_{km}^{ps} at each bias level. Note that since without bias $E(\hat{a}_{km}^{ps}) = 0$ the resulting \hat{a}_{km}^{ps} is directly the sought float ambiguity bias \hat{b} .

The variance $V(\hat{a}_{km}^{ps})$ of \hat{b} is also obtained from the least-squares solution. The uncertainties σ_Φ and σ_ρ are set to 0.001 and 0.30 meters, respectively.

The float ambiguity bias and the associated variance is evaluated taking 10, 20, 50 and 100 epochs of data into account. Given that η denotes the number of epochs the measurement equation becomes

$$E \left(\begin{bmatrix} \Phi_{km}^{ps}(t_1) \\ \rho_{km}^{ps}(t_1) \\ \Phi_{km}^{ps}(t_2) \\ \rho_{km}^{ps}(t_2) \\ \vdots \\ \Phi_{km}^{ps}(t_\eta) \\ \rho_{km}^{ps}(t_\eta) \end{bmatrix} \right) = \begin{bmatrix} 1 & 0 & \dots & 0 & \lambda_{L1} \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & \lambda_{L1} \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & \lambda_{L1} \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \Upsilon_{km}^{ps}(t_1) \\ \Upsilon_{km}^{ps}(t_2) \\ \vdots \\ \Upsilon_{km}^{ps}(t_\eta) \\ a_{km}^{ps} \end{bmatrix}. \quad (6)$$

In the simulated zero-baseline case $\rho_{km}^{ps}(t_1) = \rho_{km}^{ps}(t_2) = \dots = \rho_{km}^{ps}(t_\eta) = 0$ and $\Phi_{km}^{ps}(t_1) = \Phi_{km}^{ps}(t_2) = \dots = \Phi_{km}^{ps}(t_\eta)$ is set to each bias level, of which range depends on the antenna type as given above, in turn. Note that in the multiple-epoch simulation the bias level is constant, i.e. the same for each epoch. The simplification can be justified by noting that the satellite-geometry does not change significantly over 100 seconds (assuming 1-second epochs) and, hence, the phase bias due to antenna is constant throughout the simulation.

The float ambiguity bias distributions (not shown) are obtained by combining the results with the double difference bias distributions presented on the left in Figure 5.

VI. AMBIGUITY RESOLUTION SUCCESS RATES

Having now obtained the float ambiguity bias distributions for the antennas measured it is now possible to consider probability distributions for the probability of correct ambiguity fixing. The prerequisite for calculating these probabilities is obtaining the decorrelating transformation matrix Z using LAMBDA [1]. The transformed float ambiguity bias $\underline{\xi}$ and the associated covariance $Q_{\hat{z}}$ are then given by

$$\begin{aligned} Q_{\hat{z}} &= Z^T Q_{\hat{a}} Z \\ \underline{\xi} &= L^{-1} Z^T \hat{b} \end{aligned} \quad (7)$$

The statistics of bias-affected success rate is considered, for instance, in [2]. The probability of correct integer bootstrapping under biased conditions is obtained from

$$p(\hat{z} = z) = \prod_{i=1}^n \left(\Xi \left(\frac{1 - 2\xi_i}{2\sigma_{\hat{z}_{i|I}}} \right) + \Xi \left(\frac{1 + 2\xi_i}{2\sigma_{\hat{z}_{i|I}}} \right) - 1 \right), \quad (8)$$

where ξ_i is the i^{th} element of $\underline{\xi}$, n the number of ambiguities and $\sigma_{\hat{z}_{i|I}}$ the variance of the i^{th} ambiguity conditioned with previous $I = \{1, \dots, (i-1)\}$ ambiguities. $\sigma_{\hat{z}_{i|I}}$ is obtained as the (i, i) -element of D in the LDL-decomposition of $Q_{\hat{z}}$. The resulting L is used in (7) to calculate the transformed bias vector $\underline{\xi}$. Moreover,

$$\Xi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}v^2} dv. \quad (9)$$

Table I shows the minimum and maximum carrier phase double difference biases and the corresponding minimum success rates at the maximum/minimum bias level taking 10, 20, 50 or 100 epochs into account and using one satellite pair.

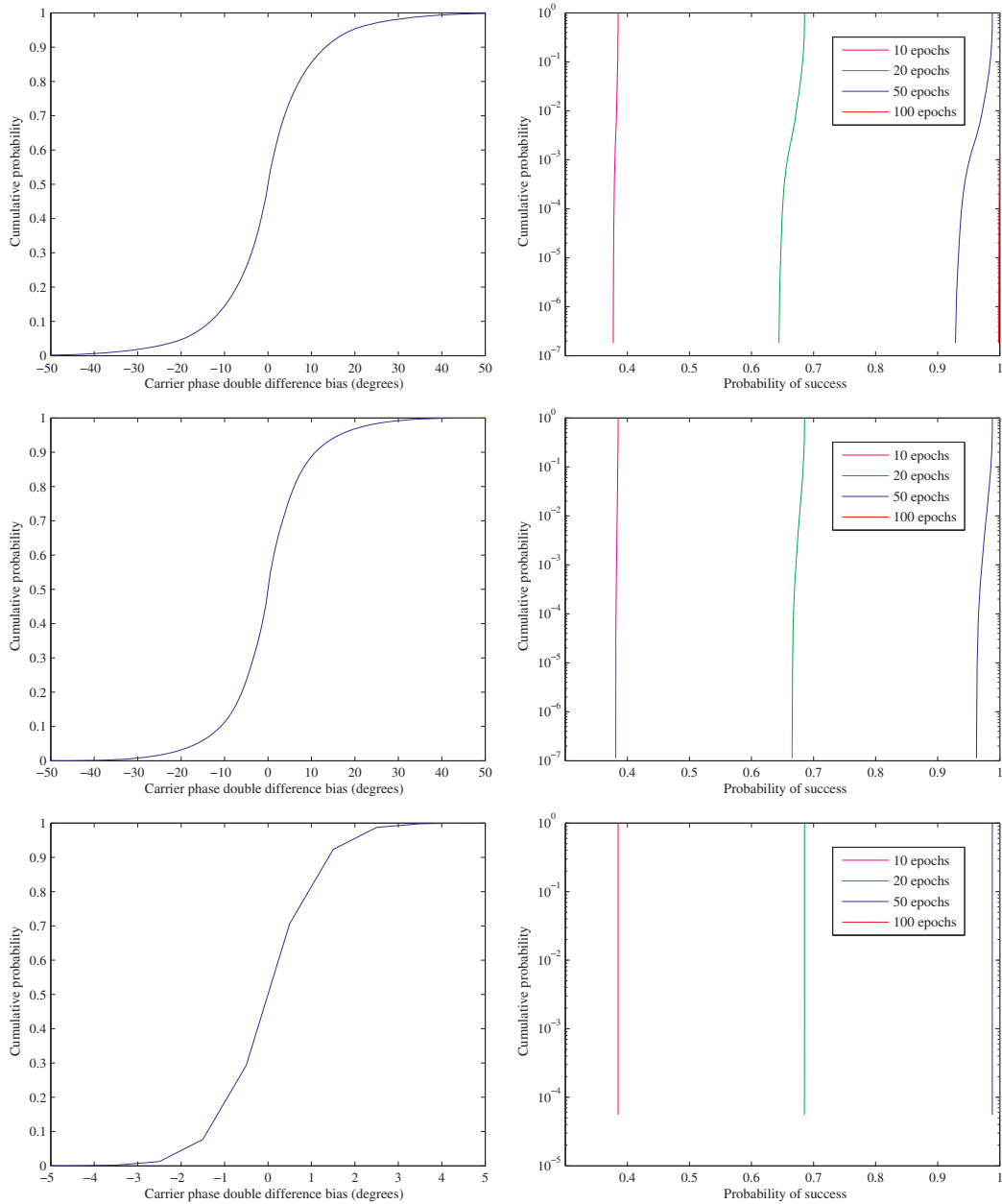


Figure 5. Cumulative PDFs for the double difference biases (left) and for the ambiguity resolution success rates (right) for BAG #1 (up), BAG #2 (middle) and Trimble Bullet (bottom).

Note that as seen from (8) the success rate is symmetric with respect to the sign of the bias. The table, therefore, summarizes the effect in the worst case, i.e. with maximal bias in the observable.

The table, first of all, shows that even in an optimal case the success rate is not unity. This is due to the noise introduced in the observables. However, the maximum success rate increases as a function of number of epochs used in the estimation due to averaging. In 10-, 20-, 50- and 100-epoch cases the maximum

success rates with previously assumed noise are 0.385, 0.686, 0.988 and 1.0, respectively, using an ideal antenna introducing no bias in the observables. Note that an ideal antenna refers to an antenna with a perfectly symmetric response or, alternatively, to an antenna with asymmetric response that is perfectly compensated for.

The right hand side in Figure 5 shows the CPDFs for success rates in the case of one satellite pair. The analysis is repeated for 10-, 20-, 50- and 100-epoch cases. In the

case of a zero-bias with unity probability (an ideal antenna) the cumulative distribution graphs would be points located at (0.385,1.0), (0.686,1.0), (0.988,1.0) and (1.0,1.0) for 10-, 20-, 50- and 100-epoch cases, respectively. However, since the antennas considered show non-zero biases, the CPDFs deviate from a point. Note also that the CPDFs do not start from zero, because biases are bounded and, hence, for each possible bias level there is a non-zero success rate.

The interpretation of the CPDFs in the current case is such that the CPDFs represent the probability at which the success rate is greater than the minimum success rates given in table I, but smaller than the success rate given in the abscissa. For example, for the BAG #1 the CPDF for the success rate (graph up on the right column in Figure 5) shows that considering 20 epochs, the probability of having a success rate less than 0.683 is 10%. However, because the minimum success rate (from table I) in the case of BAG #1 is 0.644, the graph in fact suggests that the probability of observing success rate in the range [0.644,0.683] is 10%. Vice versa, because the maximum success rate is 0.686 considering 20 epochs, the probability of observing success rate in the range [0.683,0.686] in a macro experiment is 90%.

Table I shows that in the case of Bullet the impact on the success rate due to the antenna is negligible. The relative decrease in the success rate is only in the order of 10^{-4} for all the considered solutions. The same conclusion can also be drawn from the CPDFs for the success rate. The graphs are essentially vertical denoting that success rates are distributed in a small range. The behavior is ultimately attributable to the high phase symmetry shown by the antenna.

The consumer-grade antennas, on the other hand, show differing behavior in terms of success rates. The effect due to the biases is observable especially in the results for 20- and 50-epoch solutions. This can be verified firstly from table I, where 20- and 50-epoch solutions show approximately 3-fold decrease in success rate as compared to the 10-epoch solution. Secondly, in Figure 5 on the right the CPDFs for 20- and 50-epoch solutions deviate significantly from those for 10- and 100-epoch solutions in the case of BAG. In the 10-epoch case noise dominates the solution and, hence, the biases do not affect the resolution process. With increased averaging (increasing number of epochs) the contribution of noise decreases and, in contrast, the bias starts to affect the success rates. Hence, the 20- and 50-epoch solutions are affected more than the 10-epoch solution. Moreover, the 50-epoch solution is

affected more, since in the 20-epoch solution 10% and 90% of the probability mass lie in the success rate range [0.644,0.683] and [0.683,0.686], respectively. For the 50-epoch solution the same ranges are [0.923,0.985] and [0.985,0.998], respectively. Hence, 90% of the mass lies approximately in the upmost 7% and 17% of the success rates for the 20- and 50-epoch solutions, respectively. Because in the case of a 50-epoch solution the majority of the probability mass is distributed over a wider success rate range, it can be concluded that the 50-epoch solution is affected more.

However, increasing the number of epochs further seems to dilute the effect of the bias on the success rate. The mechanism is not fully understood and might be subject to further examination. Still, having a large number of epochs leads to low uncertainty in float ambiguities. Now, because DD biases are well below half-a-cycle (in fact, well below $\frac{1}{4}$ cycles) even rounding the float ambiguities should lead to correct integer ambiguities given that the float ambiguities have low noise. Therefore, the combination of relatively low bias levels and averaging seems to dissolve the otherwise expected decrease in the success rate.

Finally, even though the 20- and 50-epoch CPDFs seem to deviate to some extent from the 10- and 100-epoch CPDFs, it can be observed that the differences are relatively small. This is confirmed by examining, where the probability mass is concentrated. As shown above, 90% of the mass covers 7% and 17% of the highest success rates in 20- and 50-epoch cases. Hence, the degradation of the ambiguity resolution success rate due to the antenna effects may be considered to be relatively insignificant in the case of BAGs as well.

VII. FUTURE WORK

Although the current analysis shows that, on average, the examined consumer-grade antennas do not degrade success rate, the current work could and should be extended to examine the spatial and temporal effects. Taking the true constellation into account and running simulations in spatial (referring to different locations on the Earth) as well as in temporal (referring to changes in satellite geometry in time) dimensions would allow for using geometry-based model and looking for, as an example, worst-case situations and their effect on the success rate. Moreover, the analysis shown assumed unrealistically only one satellite-pair. Having more signals degrades the success rate due to the multiplication of probabilities in (8). Hence, it is to be expected that using a true constellation, having more observables and using a geometry-based model would show yet unseen impacts on the success rates.

Although important, the success rate is only one component in the ambiguity resolution. The other equally interesting is the validation of the solution referring to using statistical tools to validate the ambiguities at certain confidence level. The potential increases in either false validations or false alarms should be examined in the future.

Moreover, analyzing the performance using a true constellation would also allow for quantizing the baseline errors due to the antenna-introduced biases. Should either validation

Table I

MAXIMUM AND MINIMUM DD BIASES FOR EACH ANTENNA AND THE CORRESPONDING MINIMUM SUCCESS PROBABILITIES FOR 10-, 20-, 50- AND 100-EPOCH SOLUTIONS WITH RESPECT TO IDEAL CONDITIONS.

Antenna	Bias range	number of epochs			
		10	20	50	100
Ideal	$\mp 0^\circ$	0.385	0.686	0.988	1.0
BAG #1	$\mp 75^\circ$	-2.0%	-6.0%	-6.0%	-0.17%
BAG #2	$\mp 52^\circ$	-0.96%	-2.9%	-2.6%	-0.02%
Bullet	$\mp 5^\circ$	-0.0%	-0.03%	-0.02%	-0.0%

problems or baseline errors prove to be significant, the methods available for antenna-alignment should also be studied further. Such methods include the use of information from a compass and 3D accelerometer to guide the user to align the antenna properly. Such orientation guidance might be transferred within the protocol used in relaying the measurements between the receivers.

Also, the bias simulations could be extended to include pairing of different types of antennas. In the current study it was assumed that the antennas in the ends of the baseline are of the same type. Moreover, it was assumed that the antenna planes were parallel. Therefore, the simulations might include varying all three rotation components yaw, pitch and roll. As shown in Figure 4 the consumer-grade antennas show degrading phase performance as moving towards low elevations. Setting the planes of the antenna in an appropriate manner would result in, say, one antenna observing satellites in negative elevations with respect to the antenna coordinate system. This leads to increasing phase biases and potential degradation in success rates.

More antenna-types, such as inverted-F designs found in certain AGNSS-enabled handsets, should be measured as well. This measurement campaign should include measuring a large set of devices in order to see the repeatability of the responses. The initial results with BAGs showed that the phase responses between two BAGs were too different to allow for compensation with a 3D PCV table. In any case, a simple elevation-dependent PCV table was shown to be inadequate.

Finally, the Satimo-system also supports frequency sweeps and, hence, it is possible to measure complex antenna response not only in spatial coordinates but also in frequency dimension. Such characterization allows for analyzing the distortion in the correlation triangle due to group delay. The analysis could be implemented in the same manner as in [19], which analyzes the effect of the group delay on the Galileo E5 signal due to ionosphere and, subsequently, the impact on the correlation triangle. In fact, the receiver cannot distinguish, whether the group delay is due to ionosphere or antenna and, hence, same method of analysis can be utilized in characterizing the antenna effects. Motivation for this work can be found in [7], which reports high variance in group delay as a function of azimuth for diagonally-fed antennas.

VIII. CONCLUSIONS

The measurements on consumer-grade antennas and Trimble Bullet antenna show that there are significant differences in performance with respect to phase responses. While Trimble Bullet shows phase variation in the order of few degrees, the maximum double difference biases due to consumer-grade antennas analyzed may be up to $\mp 75^\circ$.

Moreover, the consumer-grade antennas show high azimuthal asymmetry, which indicates that the method of using elevation-dependent PCV tables to compensate for phase errors is inadequate, at least, for this design. Therefore, a full 3D PCV table would be needed to characterize these

antennas. However, the responses in the two measured antennas mounted in Bluetooth GPS receivers show very different phase characteristics. Hence, using a single PCV table for all the receivers of this design may not be sufficient in order to achieve the best accuracy. Therefore, should carrier phase-based high-accuracy positioning become more commonplace in mass-market solutions, antenna designers must be given requirements with respect to the symmetry of the phase response.

Although high carrier phase bias levels can occasionally be observed because of antennas, the ambiguity resolution success rates in the geometry-free solution are still essentially unaffected in one satellite pair circumstances. This can be attributed to the low probability of observing high carrier phase bias values. It was also observed that long observation periods dilute the effect of the antenna-induced biases. However, even though the ambiguity resolution success rate is largely unaffected, the impact of the biases on the ambiguity validation or on the baseline accuracy was not analyzed. These effects are subjects to future studies.

ACKNOWLEDGMENTS

The authors would like to thank Petri Sinisalo from Nokia-OCTO (Helsinki, Finland) for performing the antenna measurements. In addition, the authors would like to acknowledge Sandra Verhagen and professor Christian Tiberius from TU Delft (Netherlands) for assisting with success rate calculations and providing helpful comments, respectively.

Finally, the authors would like to acknowledge professor Lauri Kettunen from Tampere University of Technology for discussions on the nature of the phase center.

REFERENCES

- [1] C. Tiberius and P. Jonge, "Fast Positioning Using the LAMBDA-Method," in *Proceedings of the 4th International Symposium on Differential Satellite Navigation Systems (DSNS), 24th-28th April, Bergen, Norway, 1995*, pp. 1-8.
- [2] P. Teunissen and C. Tiberius, "Bias Robustness of GPS Ambiguity Resolution," in *Proceedings of Institute of Navigation GPS 2000, 19th-22nd September, Salt Lake City, USA, 2000*, pp. 104-112.
- [3] P. Joosten and P. teunissen, "The impact of unmodelled multipath on ambiguity resolution," in *Proceedings of ION GPS 2002, Portland, OR, USA, 2002*, pp. 953-961.
- [4] P. Joosten and P. Teunissen, "On the error sensitivity of the GPS ambiguity success rate," in *International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation, 2001*, pp. 317-320.
- [5] J. Tranquilla and B. Colpitts, "GPS Antenna Design Characteristics for High-Precision Applications," *Journal of Surveying Engineering*, vol. 115, no. 1, pp. 2-14, 1989.
- [6] C. Church, I. Gupta, and A. O'Brien, "Adaptive Antenna Induced Biases in GNSS Receivers," in *Proceedings of the 63rd ION Annual Meeting 2007, 23rd-25th April 2007, Boston, MA, USA, 2007*, pp. 204-212.
- [7] W. Dong, J. Williams, D. Jackson, and L. Basilio, "Phase and Group Delays for Circularly Polarized GPS Microstrip Antennas," in *Proceedings of the 63rd ION Annual Meeting 2007, 23rd-25th April 2007, Boston, MA, USA, 2007*, pp. 545-554.
- [8] B. Schupler and R. Allshouse, "Signal Characteristics of GPS User Antennas," *Navigation*, vol. 41, no. 3, pp. 277-295, 1994.
- [9] G. Mader, "GPS Antenna Calibration at the National Geodetic Survey," *GPS Solutions*, no. 3, pp. 50-58, 1999.
- [10] Analyst report: Global Handsets, GPS/A-GPS Phone Sales. Strategy Analysts, November 2006.
- [11] J. Syrjärinne and L. Wirola, "Setting a New Standard - Assisting GNSS Receivers That Use Wireless Networks," *InsideGNSS*, pp. 26-31, 2006.

- [12] Nokia. GP-061215 Justification for the addition of carrier phase measurements. Discussion paper, presented in 3GPP TGS-GERAN Meeting#30, 26th-30th June 2006, Lisbon, Portugal.
- [13] L. Wirola, S. Verhagen, I. Halivaara, and C. Tiberius, "On the feasibility of adding carrier phase assistance to cellular GNSS assistance standards," *Journal of Global Positioning Systems*, vol. 6, no. 1, pp. 1–12, 2007.
- [14] L. Wirola, K. Alanen, J. Käppi, and J. Syrjärinne, "Bringing RTK to Cellular Terminals Using a Low-Cost Single-Frequency AGPS Receiver and Inertial Sensors," in *Proceedings of IEEE/ION PLANS 2006, 25th-27th April, San Diego, CA, USA*, 2006, pp. 645–652.
- [15] Satimo, Courtaboeuf, France, <http://www.satimo.fr>.
- [16] J. Kraus and R. Marhefka, *Antennas For All Applications*, 3rd ed. McGraw-Hill, 2002, ISBN 0-07-232103-2.
- [17] E. Kaplan and C. Hegarty, *Understanding GPS Principles and Applications*, 3rd ed. Artech house, Inc. Norwood, MA, USA, 2006.
- [18] S. Verhagen, "On the reliability of integer ambiguity resolution," *Navigation*, vol. 52, no. 2, pp. 99–110, 2005.
- [19] X. Gao, S. Datta-Barua, T. Walter, and P. Enge, "Ionosphere Effects for Wideband GNSS Signals," in *Proceedings of the 63rd ION Annual Meeting 2007, 23rd-25th April 2007, Boston, MA, USA*, 2007, pp. 147–155.

Publication 4

J. Syrjärinne and L. Wirola. Setting a New Standard: Assisting GNSS Receivers That Use Wireless Networks. *InsideGNSS*, October issue, Vol. 1, No. 7, pp. 26–31, (2006)

Copyright ©2006 Gibbons Media & Research LLC. Reprinted, with permission, from Inside GNSS magazine.

Setting a New Standard

Assisting GNSS Receivers That Use Wireless Networks



Receiver manufacturers and mobile phone designers face a plethora of wireless and GNSS standards in their efforts to build user equipment that employs telecommunications networks to improve positioning accuracy and speed. As a result, telecom engineers are proposing a single, common standard for A-GNSS.

istockphoto/Joë Bertagnoli

JARI SYRJÄRINNE, LAURI WIROLA
NOKIA TECHNOLOGY PLATFORMS

An evolution in GNSS is making new satellite systems and signals available for open-service users. This evolution offers new opportunities to improve the performance of location-based services in mobile terminals by using the increased availability and accuracy of the positioning services.

As the Global Positioning System adds signals and GPS satellites get more company in space, the wireless/cellular standards currently supporting only L1 GPS (assisted-GPS or A-GPS) need to be adapted to reflect changes in the satellite

constellations as well as recent innovations and advances in receiver and wireless infrastructure technologies. Instead of assisting only L1 GPS-receivers over wireless networks, the assistance data service must be extended to a variety of GNSSes. This means working with A-GNSS (Assisted GNSS) instead of A-GPS in the future.

The need for A-GNSS augmentations is steadily approaching as GLONASS and GPS modernizations are proceeding at a fast pace and Galileo deployment starts in the coming years. Together, these developments will multiply the number of satellites and signals available for open-service positioning in the near future.

One should also not forget the deployment of Japan's Quasi Zenith Satellite System (QZSS), India's GPS-Aided Geo-Augmented Navigation (GAGAN) system, and various other satellite-based augmentation systems planned towards the end of this decade. Moreover, the recent development in local area augmentation systems (LAAS) could bring GNSS even indoors in the form of GNSS-like pseudolite signals.

If the schedules and plans for the GNSS evolution as illustrated in **Figure 1** do not significantly change in near future, L1 A-GPS alone clearly will no longer be sufficient from 2009–2010 onwards.

Development of A-GNSS also enables a face-lift of A-GPS technology

by incorporating the latest advances in GNSS receiver and wireless infrastructure technologies. This will allow for a totally new class of applications providing high accuracy, superior availability, and seamless hybrid use of GNSSes and/or terrestrial wireless networks on a global scale.

It seems to us that copy-pasting GNSS Interface Control Documents (ICDs) into cellular standards, similar to the A-GPS concept, may not be the best way to introduce A-GNSS. Instead, the full potential of GNSS could be introduced by novel approaches leaning on increased bandwidths of the near-future radio interfaces and external GNSS monitoring and tracking networks such as the International GNSS Service (IGS). This article explores these possibilities and advances in the context of wireless networks and mobile terminals.

Current Work Towards A-GNSS

During the past three years new work items to add A-GNSS functionality to cellular assistance data have been approved and launched in Third Generation Partnership Project (3GPP) standardization bodies.

These work items have concerned modifying Radio Resource LCS Protocol (RRLP) and Radio Resource Control (RRC) defined for the Global System for Mobile Communications (GSM) and Universal Mobile Telecommunications System (UMTS) A-GPS protocols, respectively. Moreover, the Open Mobile Alliance (OMA) forum has work items to modify the Secure User Plane Location (SUPL) Service to add the support for A-GNSS. **Table 1** lists the cellular standards/protocols for GNSS assistance.

The most initiative and activity have come in 3GPP GSM/EDGE Radio-Access Network (GERAN) meetings, where several proposals towards A-GNSS have been presented and discussed. Currently, 3GPP GERAN aims to extend the scope of the work from A-Galileo-only additions towards a more general A-GNSS concept. The group has recognized that A-Galileo-only addi-

tions will no longer suffice as other new GNSS signals and services will become available along with the full deployment of the Galileo constellation.

Exploiting the Full Capability of A-GNSS

The main benefit of A-GNSS should not merely be an increased number of satellites available to GNSS-capable wireless terminals. Instead, A-GNSS should be an enabler for technologies and services that will make it possible to exploit the full potential of GNSS. A well-formulated A-GNSS standard could help extend GNSS service into new applications and operating environments, especially, to the applications and use scenarios that were not seriously considered 10 years ago for A-GPS.

Extending the lifespan of assistance data. The typical environment for A-GNSS terminals is an urban or indoor area where positioning and navigation needs to be carried out under signal blocking, high signal attenuation, and multipath conditions. High-sensitivity and hybrid uses of GNSS satellites are, therefore, important aspects to enable navigation, time determination, and

RAIM from a rather scarce number of satellites and signals.

This means that, in order to operate and maximize the performance under these harsh conditions, the terminals should either have a continuous access to assistance data service or the terminals should have the assistance data already in memory. As the former might sometimes be either limited or unavailable, extension of the lifespan or *persistence* of the assistance data in the terminals becomes an important consideration, especially for navigation.

Various ways exist to extend the usable lifetime of the assistance data. One of the most promising is long-term orbit (LTO) data for satellite orbit and clock models that could be provided to terminals for full constellations even for several days ahead. As the GNSS evolution brings along better and more stable satellite clocks, the performance of LTO data will become better in the near future as the satellite clock drifts can be predicted more reliably. (For further discussion of this subject, see the article by David Lundgren and Frank van Diggelen cited in the "Additional Resources" section at the end of this article.)

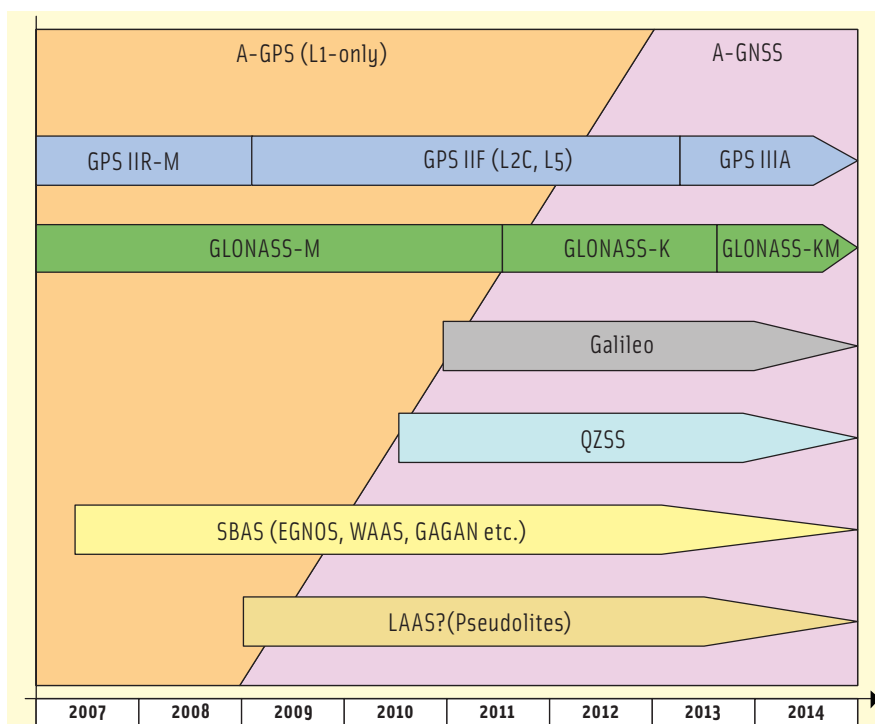


FIGURE 1 Parallel evolution in GNSS systems and A-GNSS

Protocol/Standard	System	GNSS Support	Comment
RRLP*	GSM	GPS	Additional features such as request of additional assistance introduced as Supplementary Service
Broadcast Assistance	GSM	GPS	No acquisition assistance
RRC*	UMTS	GPS	
IS-801.1	CDMA	GPS	No DGPS
IS-801.A	CDMA	GPS + some support GPS modernization and WAAS	Support for L1 WAAS, L2C C/A, L2C LM, and L5 measurements. No modernized ephemeris.
SUPL1.0	IP based	GPS	Wrapper for RRLP*, RRC*, IS-801.1
SAMPS	TDMA	GPS	TDMA market reducing
IS-817	AMPS	GPS	AMPS market reducing
X.P0024	IP based for 3GPP2	GPS	

* RRC and RRLP have messages to report the relation between GPS system time and cellular framing time from the terminal to network.

TABLE 1. Cellular standards/protocols for GNSS assistance.

Maximizing sensitivity in asynchronous networks. Sensitivity depends directly upon the accuracy of the reference time in the terminals. Naturally, sensitivity is also a function of reference frequency, initial position, and ephemeris, but these elements are typically available either from the terminal itself or from network assistance. For example, coarse location based on the cell ID and ephemeris data from a reference receiver are typical elements of any GPS assistance data protocol.

Accurate time assistance requires either a synchronized network (such as CDMA cellular telephone systems) or deployment of network time-measuring elements to calculate the time differences between the cellular base stations and GNSS (read GPS) system time. The latter is specifically for asynchronous networks such as GSM and UMTS.

Assuming accurate time assistance, the signal search window in the code phase domain can be reduced even down to a few GPS chips (fewer than 10 chips) in typical urban conditions. The sensitivity is improved not only by the possibility of performing coherent integration over the full GPS bit (20 milliseconds) but also by minimizing the probability of false alarms. Naturally, time to first fix (TTFF) and power consumption will also be minimized, as the fix can be calculated as quickly as possible without

the need to carry out exhaustive, full code domain signal searches.

GNSS evolution will open the door for even higher levels of sensitivity by bringing a wide range of pilot signals for open service users. Coherent signal integration can be prolonged to well more than 20 milliseconds, extending the coverage of A-GNSS positioning services beyond that of A-GPS service.

However, accurate time is still needed. The maximum benefit of the new pilot signals will be gained by having reference time accurate within a few microseconds. Nonetheless, reference time accurate within few hundred microseconds will still prove useful for predicting phases of possible secondary codes while keeping the size and cost of the search engine hardware within reasonable limits.

If accurate reference time is not directly available from the network, indirect methods can make use of the network to ensure precise timing. Even though a network is not synchronized, the cellular signals (cellular base stations) typically have very good frequency stabilities that can be employed in the terminals to *maintain* the relation between GPS/GNSS and cellular system times.

A-GNSS can also come to aid the terminals in this area, for example, by enabling transmission and delivery of GNSS-cellular system time differences from the serving base stations and even

from neighboring stations in the form of observed time difference measurements. Further, time relations and measurements from multiple mobile terminals can be gathered in network servers to improve the accuracy and quality of the measurements transmitted back to terminals. This data helps the terminals to maintain timing relationships accurately during handovers and sleep periods.

Face-lifts. Modern GPS receivers not only track the code, but also the carrier phase. However, carrier phase measurements are not included as such in any A-GPS.

The applications of carrier phase measurements are naturally in the RTK area (see the article by K. Alanen et al in the May/June issue of *Inside GNSS* cited in Additional Resources). They also appear in accurate terminal velocity calculations and in precise point positioning (PPP). PPP would improve the stand-alone positioning accuracy to less than one meter assuming proper assistance data. However, the introduction of PPP evidently requires more than just carrier phase measurements to be available in the network assistance.

Additional assistance elements include Earth orientation parameters (EOP) as well as accurate ionosphere and satellite orbit models. All of these elements are pieces of information that either exist today or are included in the near-future GNSS evolution. Moreover, although there are bandwidth limitations in today's radio interfaces, the coming Orthogonal Frequency Division Multiplexing (OFDM) radio interfaces (WiMAX, 3.9G) are capable of delivering high-accuracy satellite orbit data on a frequent basis.

Seamless Use of Assistance Data

Arguably an optimal and future-proof A-GNSS solution can only be achieved by creating a generic, scalable, and flexible assistance data format that not only fits all the existing or soon-to-be-deployed GNSSes but also has reservations for coming systems. Moreover, the planned tangible performance improve-

ments compared to A-GPS ensure fast deployment.

In order to achieve this, the format, content, quality and applicability of assistance data needs to be the same regardless of the carrier medium. In fact, one of the greatest current risks is the divergence of A-GNSS implementations. The variety of A-GPS standards is already becoming a challenge for multi-mode terminals needing to support multiple A-GPS implementations, not to mention the issues with interworking and interoperability.

Last, but not the least aspect is the question of backwards compatibility, which has to be maintained so as not to jeopardize the functioning of existing implementations in the future. This almost inevitably leads to a conclusion that the current A-GPS implementation should not be touched and that the A-GNSS is best accomplished as a totally new concept as illustrated in **Figure 2**.

This road could also lead to a convergence of A-GNSS standards instead of increasing the complexity and number of assistance data standards by upgrading individually the existing A-GPS standards in different systems at different times into the A-GNSS standard.

A-GNSS Assistance Data Structure

The assistance data elements may be divided in two categories based upon

This road could also lead to a convergence of A-GNSS standards instead of increasing the complexity and number of assistance data standards by upgrading individually the existing A-GPS standards in different systems at different times.

whether they are GNSS-independent or GNSS-specific. GNSS-independent elements are called common elements, which are summarized below.

Common assistance data elements. The information elements that are the same regardless of the GNSS include:

- reference time – common system time from wireless network
- reference location – initial location of the receivers

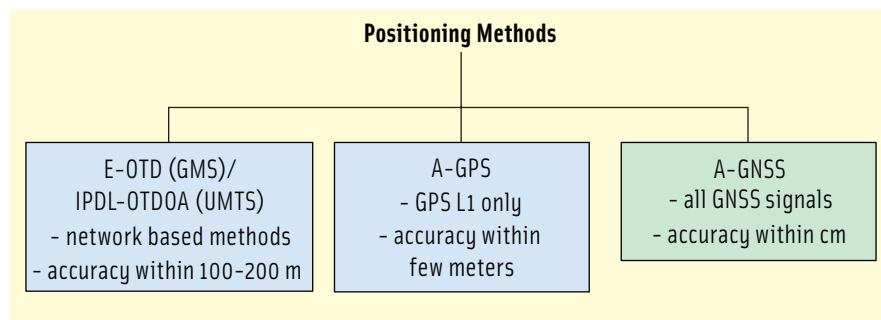


FIGURE 2 Location Services (LCS) methods

- atmosphere models – troposphere and ionosphere models for atmosphere error correction
- base station/access point timing models
- Earth orientation parameters

Any of these data elements are needed whether one or several GNSSes are included in the assistance data, and the data can be applied on any GNSS. An example of this is the troposphere model that can be applied to any GNSS signal. Notably, these elements can be derived or obtained from a variety of sources, including GNSS broadcasts as well as external (commercial) services.

Common system time. One of the challenges in A-GNSS is the reference time. Surely, one possibility would be to include all GNSS-specific system times into A-GNSS. However, this approach not only leads to complexity in implementation due to differences in the terminal and network capabilities sup-

(read GPS)-specific timing to a Universal Coordinated Time (UTC) base that acts as a virtual time reference convertible to any GNSS-specific time using the UTC-GNSS time relations provided in the reference time–assistance element. Natural benefits of a common system time include:

- 1) only one time base in assistance data and response messages
- 2) PPS generation from any GNSS or combination of GNSSes, which is important to maintain reference time in the terminal and to synchronize wireless terminals with a common time
- 3) seamless hybrid use of GNSS as per the article by Moudrak et al
- 4) Ease in adding new GNSS and time bases
- 5) UTC will be available for terminal resident LBS and time-based applications such as validation and timing of financial transactions.

Use of a common UTC-based time has one drawback, namely leap seconds. Therefore, in order to keep the system time continuous over the leap second occurrences the following approach could be taken:

- GNSS-specific UTC leap second counts are frozen to the values at January 1, 2006.
- The reference time information element in the assistance data has two parameters for the leap seconds that occur after January 1, 2006: one that indicates the current number of leap seconds since January 1, 2006 and another that indicates the next occurrence of the UTC leap second. (Terminals would increment the number of leap seconds

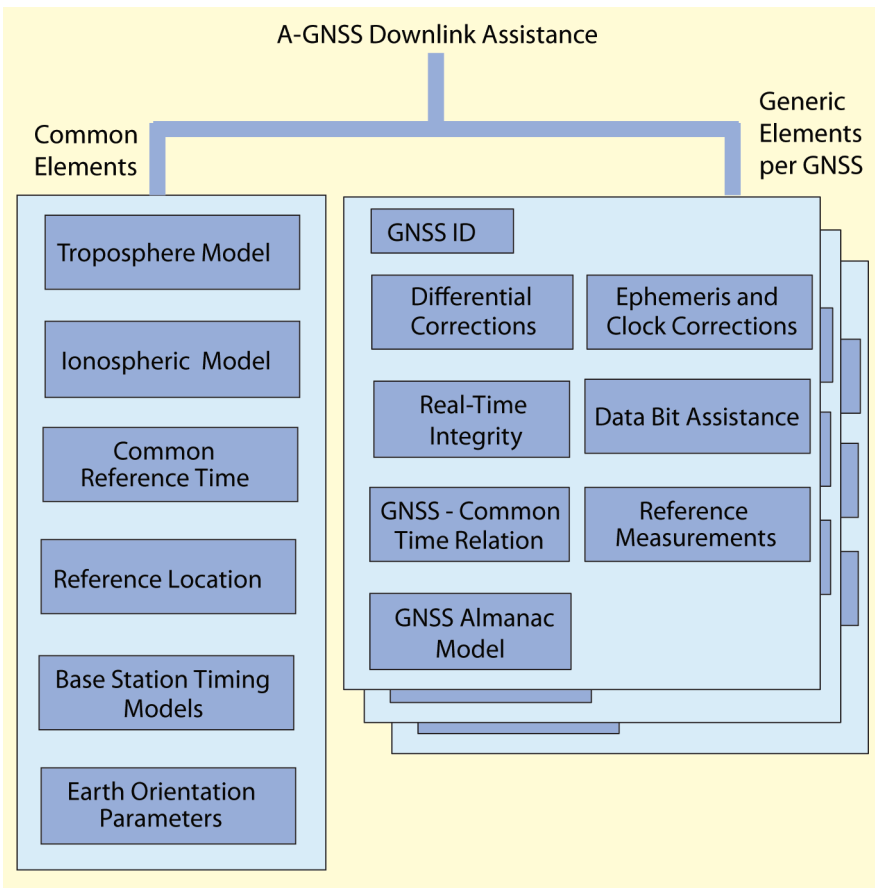


FIGURE 3 A-GNSS downlink

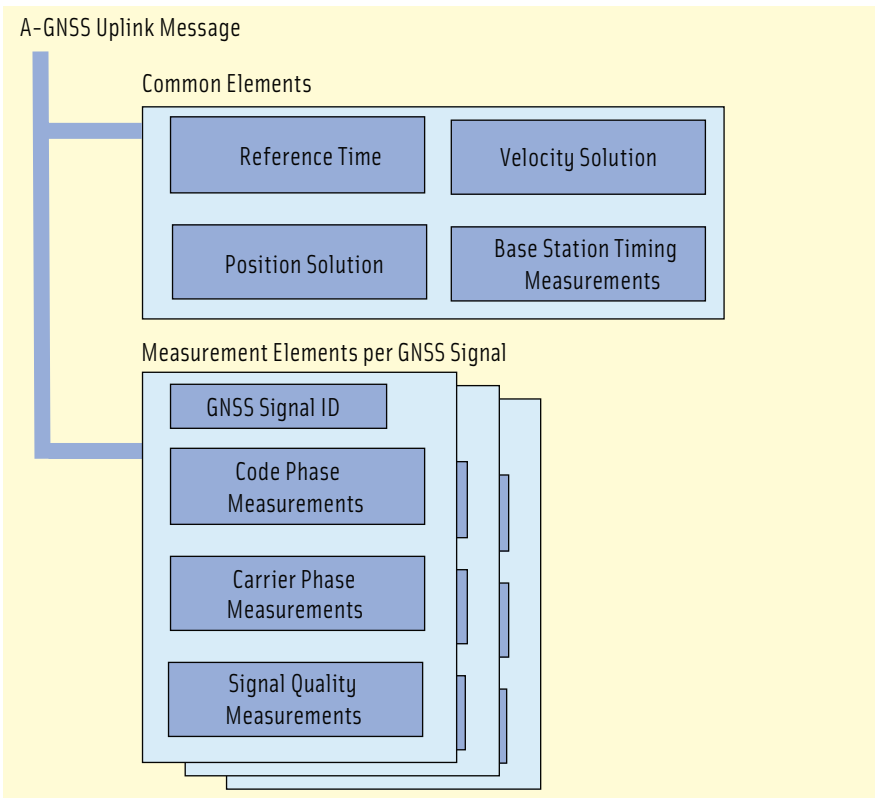


FIGURE 4 A-GNSS uplink

by one, when the next leap second occurs.)

In this manner terminals would be capable of maintaining accurate track of the UTC time based on any GNSS time for PPS generation and, for instance, for NMEA messages.

Generic assistance data elements.

An A-GNSS standard will also need to carry GNSS-specific elements. Due to the nature of the technology, however, GNSSes are very similar in some respects. For instance, measurements such as code and carrier phase data available in different systems are the same. This characteristic enables the introduction of *generic* assistance data elements.

Generic data formats can be applied from system to system and, therefore, reduce the implementation complexity. These elements should include at least the following:

- differential corrections
- real-time integrity
- data bit assistance
 - subframes and so forth to enable receiver processing techniques such as data wipe-off
- reference measurements — code and carrier phase measurements from a reference station for high-accuracy positioning
- GNSS – common system time relationship
- almanac models
- ephemeris and clock models (Note that the utilization of LTO models makes almanacs unnecessary.)
- satellite clock and orbit data in a suitable format.

The most interesting element here is the ephemeris and clock models. GNSSes share a lot of commonalities here (see, for instance, the GPS and Galileo ICDs), which automatically suggests a generic format that could be applied to any GNSS.

The generic ephemeris and clock model format brings along a very interesting and tempting possibility of using non-native parameter formats for GNSSes. For example, the Keplerian parameter representation typically employed by GPS could be used for GLONASS when

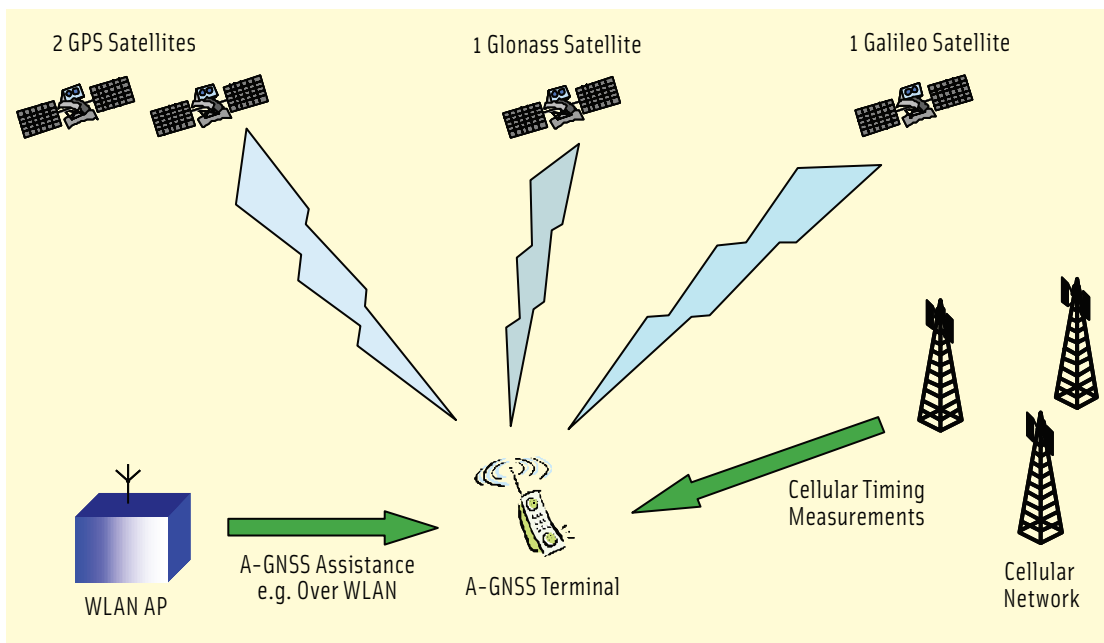


FIGURE 5 A nominal A-GNSS architecture

ing the hybrid use of more than one GNSS, because the assistance data quality and the performance of the GNSS signals can now be truly balanced for the first time.

Additional Resources

- [1] Averin, Sergey V., "GLONASS System: Present Day and Prospective Status and Performance," European Navigation Conference GNSS-2006, Manchester, UK, May 7-10, 2006.
- [2] Lundgren, David, and Frank van Diggelen, "Assistance When There's

providing LTO data to GLONASS-capable terminals.

The possibility of mixing the formats would allow an A-GNSS standard to equalize the satellite position information across the different GNSSes. It also would make the lifetime of the data elements the same, further simplifying the implementation. Figure 3 illustrates the structure of the proposed A-GNSS approach for the downlink of assistance data and Figure 4, a generic structure for the measurement and position information response suitable for any specific or combination of GNSSes.

Native data support. Despite the generic approach, the ephemeris and clock models can be constructed so that the native data parameters can be carried in the messages without any loss of data or precision. This is an important aspect for A-GNSS server implementations, where the assistance data may be derived directly from the satellite broadcast without any LTO processing.

Scalability. The limitations of radio channels in terms of bandwidth and latency vary greatly from network to network, which needs to be taken into account in data formats. For example, the GERAN control plane channel cannot be used to deliver large amounts of ephemeris data in a reasonable amount

of time due to bandwidth limitations. On the other hand, a broadband 3G high speed data packet access (HSDPA) connection can carry ephemeris data for full GNSS constellations in virtually no time. Regardless of the data volume, the most important aspect is to have the same format in all cases and to design scalable assistance data information elements (IEs).

The Way Forward

In formulating an A-GNSS standard, we need to have an open-minded approach to ensure that the principles and features selected will exploit the full potential of the future GNSS signals, services, and wireless networks, not forgetting the growing terminals capabilities.

The natural way forward is a solution supporting seamless hybrid use of any GNSS and wireless network as illustrated in Figure 5. The solution must also be as future-proof as possible, scalable and addressing all the GNSSes equally. This means that the GNSSes have to be addressed as a whole, not as single, separate systems.

As the data elements are made common and generic, it is equally important to also make them, as much as possible, system- and carrier-independent. This particularly applies when consider-

No Assistance," *GPS World*, October 2005, pages 32-36.

[3] Moudrak, A., and A. Konovaltsev, J. Furthner, J. Hammesfahr, P. Defraigne, A. Bauch, S. Bedrich, and Arno Schroth, "Interoperability on Time," *GPS World*, March 2005.


[4] "Background for starting a new Study Item "LCS Evolution," GP-061463, 3GPP TSG-GERAN Meeting #30, Lisbon, Portugal, 26-30 June, 2006.

[5] ICD-GPS-200, Revision D, Navstar GPS Space Segment/Navigation User Interfaces, December 7, 2004.

[6] ICD-GPS-800, Draft, Navstar GPS Space Segment/User Segment L1C Interfaces, April 19, 2006.

Authors

Jari Syrjälä received his M.Sc. and his Ph.D. from Tampere University of Technology, Finland, majoring in digital signal processing and applied mathematics. Between 1996 and 1998 he worked for the Tampere University of Technology Signal Processing Laboratory in the areas of data fusion and target tracking, and since 1999 he has been with Nokia Inc. He is currently involved in work on A-GNSS and hybrid positioning.

Lauri Wirola received his M.Sc. degree from Tampere University of Technology, Finland, with a major in electrophysics. Since 2005 he has been working on navigation issues with Nokia Technology Platforms. His present research interests include RTK and A-GNSS standardization. He is currently undertaking postgraduate studies in modern electromagnetism and mathematics. 

Publication 5

L. Wirola and J. Syrjärinne. Bringing the GNSSs on the Same Line in the GNSS Assistance Standards. In *Proceedings of ION Annual Meeting*, Boston, USA, April 23rd-25th, pp. 242–252, (2007)

Copyright ©2007 ION. Reprinted, with permission, from the proceedings of ION Annual Meeting 2007.

Bringing the GNSSs on the Same Line in the GNSS Assistance Standards

Lauri Wirola, *Nokia Technology Platforms*
Jari Syrjärinne, *Nokia Technology Platforms*

BIOGRAPHY

Lauri Wirola, M.Sc., received his Master of Science degree from Tampere University of Technology, Finland, in 2005. He majored in electrophysics and minored in technical physics and industrial economics. In 2004-05 he worked on acoustics at the Nokia Research Center. In 2005 he joined the positioning group at Nokia Technology Platforms. His present research interests include multi-GNSS positioning, RTK and A-GNSS standardization. He is currently undertaking postgraduate studies in modern electromagnetism and mathematics.

Jari Syrjärinne, Ph.D., received his M.Sc. in 1996 and his Ph.D. in 2001, both from Tampere University of Technology, Finland, majoring in digital signal processing and applied mathematics. Between 1996 and 1998 he worked for the Tampere University of Technology Signal Processing Laboratory in the areas of data fusion and target tracking, and since 1999 has been with Nokia Inc. He is currently involved in work on A-GNSS and algorithms for hybrid positioning.

ABSTRACT

The cellular networks, such as GSM (Global System for Mobile Communications) and UMTS (Universal Mobile Telecommunications Systems), have an in-built capability to assist integrated GPS receivers in cellular terminals. This assistance includes, among other things, the navigation model (orbit and clock parameters). In an assisted situation, the receiver does not need to obtain a copy of the navigation model from the satellites, but receives it over the cellular network.

Until now there has been no need to assist any GNSS (Global Navigation Satellite System) other than GPS and, in fact, currently there is support only for L1 AGPS (Assisted GPS). Now, however, the introduction of new GNSSs, such as Galileo and GLONASS, demands updating the outdated AGPS-only solution. The main driver of the revision process is to add support for new GPS bands, new GNSSs, and to enhance the performance

of positioning technologies in terms of accuracy, sensitivity and availability.

This article discusses the various systems from the point of view of orbit and clock models. Ultimately, based on this discussion the future AGNSS assistance format for orbit and clock models is synthesized. The format is called the multi-mode navigation model and it has been approved by the 3GPP (3rd Generation Partnership Project) standardization body to be included in the GSM standard. The benefits of the new navigation model include that it can be used for all the current, known and future GNSSs. In addition, the multi-mode navigation model allows using non-native navigation models for GNSSs. This results in harmonized performance between the different GNSSs.

Keywords: AGPS, Assisted GNSS, multi-mode navigation model, 3GPP

INTRODUCTION

The AGNSS (Assisted Global Navigation Satellite System) is a concept that significantly improves the performance of a GNSS receiver integrated in a wireless terminal. The assistance data contains, among other things, orbit and clock models that are necessary for position calculation. For example, for a GPS receiver it takes at least 18 seconds to receive this data from the satellites. However, in AGPS (Assisted GPS) receivers the cellular network sends the receiver a copy of the navigation message. Time-To-First-Fix (TTFF) can then be reduced to sub-18 seconds. This reduction in TTFF is essential, for instance, when positioning an emergency call. This also improves user experience.

Moreover, in adverse signal conditions an AGNSS receiver may be the only option for navigation and positioning. This is because low signal power levels make it impossible for the GNSS receiver to obtain a copy of the navigation message. However, when the navigation data is provided to the receiver from an external source (such as communications network), navigation is enabled

again. This feature is important in indoor conditions as well as in urban areas, where signal levels may vary significantly due to buildings and other obstacles.

It should be noted that the assistance data may also include accurate time assistance and reference location, the use of which potentially significantly improves the sensitivity of the receiver.

The AGPS and navigation in cellular and other wireless terminals have only recently become interesting to the mass market. Until now the MS-Assisted (Mobile Assisted) version of AGPS has primarily been used to fulfill the FCC (Federal Communications Commission) E911 (Enhanced 911) ruling. The ruling sets requirements for mobile-originated emergency call positioning in the US so that in the case of an emergency call, the terminal must be positioned with an accuracy of 50-300 meters depending on the technology used. MS-Assisted GPS does not require the network to send navigation model assistance to the terminals, but positioning is carried out in a network server using the measurements provided by the terminal. The only assistance data needed for the measurements are predicted code phases and Doppler frequencies at the estimated location of the terminal. This assistance is called acquisition assistance.

However, the demand for Location Based Services (LBS) as well as for navigation-capable terminals is growing fast. Altogether 30% of the cellular terminals sold in 2011 are expected to be navigation capable [23]. The boom in LBS and navigation applications will also mean a rapidly growing use of MS-Based GPS and the related assistance data information elements such as navigation models, reference location and time.

Until recently it was possible to provide assistance only for GPS L1 in GSM (Global System for Mobile Telecommunications) and UMTS (Universal Mobile Telephone System) networks. In GSM the assistance is specified in the Radio Resource LCS (Location Services) Protocol (RRLP, [17]) and in UMTS in the Radio Resource Control (RRC, [18]). Moreover, there are also user plane solutions, such as Open Mobile Alliance (OMA) Secure User Plane Location (SUPL, [22]), that effectively carry the same information elements over packet networks that are specified for circuit-switched networks in RRLP and RRC.

However, in autumn 2006 the 3GPP standardization body [16] accepted a proposal which opened the way for the addition of new GPS bands as well as other GNSSs to the assistance standard [20]. The work towards this was initiated due to the prospects seen in the Russian GLONASS (ГЛОНАСС, [12]), European Galileo [11], Japanese Quasi-Zenith Satellite System (QZSS, [13]) and various Space-Based Augmentation Systems, such as the

Wide Area Augmentation System (WAAS, [14]) and European Geostationary Navigation Overlay System (EGNOS, [15]).

The two proposals that competed in the 3GPP differed in how the new systems should be brought into the standards. One proposal was that the systems would be added one-by-one. The other suggested a more generalized approach to enable as smooth and bit-efficient an addition of currently known and future systems as possible. The proposition also included the possibility for performance improvements, such as carrier-phase measurements, orbit extensions and use of non-native navigation models. This generalized approach was chosen over the proposal to add the systems individually. [20]

In order to be able to provide assistance for various GNSSs, a navigation model supporting a variety of GNSSs in a generic manner is required. The model should also make it possible to equalize the performance of various GNSSs, support seamless hybrid use of GNSSs and allow performance improvements by allowing the use of non-native navigation models.

This paper introduces a 3GPP-approved multi-mode navigation model. The navigation model is defined to consist of the orbit and clock models including the group delay parameters. Naturally, there are also a plethora of other essential parameters, such as fit interval, time of ephemeris, health and integrity data. Although all these parameters are also transferred in the assistance data flow, no stand on their implementation is taken in this paper.

GNSS TIMELINE

Currently there are two global satellite navigation systems available for public use: GPS and GLONASS. Moreover, the modernization of GLONASS is well underway with, for instance, the launch of three new GLONASS-M satellites in late 2006. The first GLONASS-M satellite was launched in 2003. The Russians have announced that GLONASS will achieve minimal operation capability (18 satellites) by the end of 2007 and full operation capability (24 satellites) by the end of 2009. Unlike the current GPS, GLONASS provides civilians with ranging data on two bands. With the introduction of GLONASS-K 2008 onwards, there will also be a third civilian frequency. The respective timelines of GLONASS-M and GLONASS-K are given in figure 1. [6] [7]

GPS modernization is also well in progress and currently there are three modernized GPS-IIR-M satellites in orbit [24]. In a later phase, the GPS-IIF-block will bring additional frequencies (L2C, L5) to civilian use.

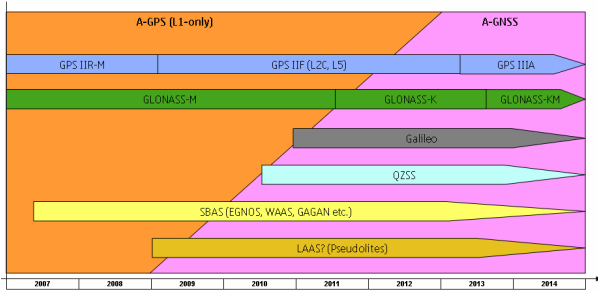


Figure 1. Timelines for various GNSS launches and modernization programs. Reproduced from [1].

The European Galileo system is currently being ramped up and will, according to the current best knowledge, achieve full operational capability by 2011 [28]. The draft version of the SIS ICD [11] is, however, available. The various Galileo-related items in this article are based on the draft ICD.

Space Based Augmentation Systems (SBAS) are currently being deployed to distribute correction data to GPS receivers. The two functioning SBASs are the WAAS in North America (3 satellites) and EGNOS in Europe (3 satellites). In addition, the Japanese GPS augmentation system QZSS is expected to be ramped up at the beginning of 2009 with the launch of the first QZSS satellite. The QZSS consists of at least three satellites in highly inclined and highly elliptical geostationary orbits. The Japanese intend to have at least one QZSS satellite above the 60°-elevation mask at all times over the whole of Japan, augmenting GPS. The QZSS signal specification closely follows that of modernized GPS. [13]

Finally, yet another interesting development is introduced by the deployment of Local Area Augmentation Systems (LAAS), also known as Pseudolites. LAASs are currently used to support air traffic near airports, but due to the technical challenges their use has not become general elsewhere. Specific challenges in the use of LAASs include RF interference in the GPS band, near-far problem in power levels, lack of available PRN space and problems in transferring the LAAS transmitter location within the native GPS ICD navigation data. Despite these problems, it is still advantageous to support LAAS in the assistance standards due to completeness and consistency.

OVERVIEW OF SYSTEMS USING KEPLERIAN ORBIT REPRESENTATIONS

The currently known GNSSs may be divided into two categories based on the orbit parameterization used in their native navigation model broadcasts. The first group is those utilizing Keplerian parameters while the other group utilizes position, velocity and acceleration coordinates at the given instants.

Table I. Parameters required in representing GPS and Galileo satellite orbits, their bit counts and scale factors. (u) denotes an unsigned parameter and sc a semi-circle.

Parameters	Bits	Scale factor
ω , Argument of perigee	32	2^{-31} sc
M_0 , Mean anomaly	32	2^{-31} sc
e , Eccentricity	32(u)	2^{-33}
$A^{1/2}$, Square root of semi-major axis	32(u)	2^{-19} m ^{1/2}
i_0 , Inclination	32	2^{-31} sc
Ω_0 , Longitude of the ascending node	32	2^{-31} sc
Δn , Mean motion correction	16	2^{-43} sc/s
\dot{i} , Rate of change of inclination	14	2^{-43} sc/s
$\dot{\Omega}$, Rate of change of Ω_0	24	2^{-43} sc/s
C_{rs} , Correction to radius	16	2^{-5} m
C_{is} , Correction to inclination	16	2^{-29} rad
C_{us} , Correction to latitude	16	2^{-29} rad
C_{rc} , Correction to radius	16	2^{-5} m
C_{ic} , Correction to inclination	16	2^{-29} rad
C_{uc} , Correction to latitude	16	2^{-29} rad

The systems utilizing Keplerian parameters may further be subdivided into two sub-groups. The first includes the well-known format utilized by GPS [8] and Galileo [11]. This format is summarized in tables I and II, which show the parameters, bit counts and the respective scale factors for orbit and clock models. The other sub-group includes modernized GPS ([9], [10]) and QZSS [13], which both utilize High-Accuracy Keplerian parameterization (later HA Keplerian). The term High-Accuracy Keplerian model is introduced in order to be able to differentiate between the two formats in GPS and modernized GPS.

The orbit model in modernized GPS utilizes parameters with greater resolution and also the first time derivatives of certain parameters in order to capture their temporal changes during the fit interval. The parameters and the respective bit counts and scale factors for orbit and clock models are summarized in tables IV and III.

Although the Japanese QZSS follows the message format of modernized GPS, there are a few minor differences that must be accounted for in the implementation. These include, firstly, a different interpretation of the group delay parameter T_{GD} . In the QZSS this parameter is either the delay between L1 C/A and L2C signals or between L1 C/A and L5. Secondly, the range of eccentricity is not limited to 0.03 as in GPS, but the QZSS messages utilize the full potential range up to 0.5. Thirdly, the reference value of the semi-major axis is 42 164 200 m instead of 26 559 710 m utilized in GPS. [13]

Table II. Parameters required in representing GPS and Galileo satellite clock models, their bit counts as well as scale factors.

This model is called the standard clock model as opposed to High-Accuracy model. The upper bit figures are for GPS and the lower for Galileo.

Parameters	Bits	Scale factor
a ₀ , bias	22	2 ⁻³¹ s
	28	2 ⁻³³ s
a ₁ , 1 st order term	16	2 ⁻⁴³ s/s
	18	2 ⁻⁴⁵ s/s
a ₂ , 2 nd order term	8	2 ⁻⁵⁵ s/s ²
	12	2 ⁻⁶⁵ s/s ²
T _{GD} , group delay	8	2 ⁻³¹ s
	10	2 ⁻³² s

The orbit data represented in a Keplerian model can be represented in a HA Keplerian format, since although the resolution is higher in the HA model, the bit counts have respectively been increased in order to maintain the original range. As an example of this, the argument of perigee is represented in GPS format with 32 bits and the scale factor of 2⁻³¹ semi-circles. On the other hand, in modernized GPS the same parameter is expressed with 33 bits but with the scale factor of 2⁻³² semi-circles. However, the best way to see why the Keplerian model is a proper subset of the HA Keplerian model is that they are used to represent the same orbits, because GPS and modernized GPS satellites have similar orbits.

Moreover, the HA model contains time-dependent correction terms and, hence, is arguably capable of representing orbits more accurately. It should still be noted that the mapping from the Keplerian model to the HA model may not be as straightforward as it may first seem, because of, for instance, inclination parameters. The HA model uses higher resolution for inclination and for the rate-of-change of inclination. This has apparently made it possible to increase the resolution of the harmonic correction terms for inclination without increasing the bit count. Therefore the range used for harmonic inclination correction in GPS/Galileo is greater than the range in the HA Keplerian model. This must be taken into account in the mapping.

The clock models in tables II and III allow for drawing two conclusions. Firstly, the model utilized in Galileo has sufficient range and resolution for use with GPS as well. Secondly, unlike with the orbit models, the standard clock model (table II) is not a subset of the HA clock model (table III). The HA clock model can represent neither the GPS nor Galileo clock model, collectively denoted as the standard clock model. This is because the HA clock model has been designed to model the more stable clocks onboard modernized GPS satellites.

Table III. Parameters required in representing modernized GPS and QZSS satellite clock models, their bit counts as well as scale factors.

Parameters	Bits	Scale
a ₀ (bias)	26	2 ⁻³⁵ s
a ₁ (1 st order term)	20	2 ⁻⁴⁸ s
a ₂ (2 nd order term)	10	2 ⁻⁶⁰ s/s ²
T _{GD}	13	2 ⁻³⁵ s
ISC _{L1CA}	13	2 ⁻³⁵ s
ISC _{L2C}	13	2 ⁻³⁵ s
ISC _{L5I5}	13	2 ⁻³⁵ s
ISC _{L5Q5}	13	2 ⁻³⁵ s
ISC _{L1CP}	13	2 ⁻³⁵ s
ISC _{L1CD}	13	2 ⁻³⁵ s

Because the Keplerian orbit model can arguably be mapped to the HA orbit model, it would be tempting to specify that the orbits parameterized in terms of Keplerians will be re-parameterized in terms of the HA Keplerian model in the assistance standards. In such a case, the terminal implementation would become simpler and only one data structure would need to be defined for all the GNSSs that use Keplerian or HA Keplerian parameterization. However, this is unacceptable due to the high bit consumption. Comparing the bit counts in tables I and IV, it can be noticed that the HA model consumes 20% more bits (410 b versus 342 b). Moreover, standard and HA clock models could be represented by a new model taking both models into account. This could be done by increasing the bit counts appropriately. However, this would also lead to an unacceptable loss of bit-efficiency. Therefore since the models cannot be combined, they must both be included in the assistance standard. Using both models also maintains the support for native model formats in the assistance standards.

OVERVIEW OF SYSTEMS USING CARTESIAN COORDINATE SYSTEMS FOR ORBITS

GLONASS, SBASs and LAASs use orbit representations that are based on giving the position, velocity and acceleration of the satellite at a certain instant. In the case of GLONASS, these coordinates are initial values for the integration of the equations of motion [12]. Tables V and VI summarize the orbit and clock models used in GLONASS.

With GLONASS one must also consider coordinate systems. GLONASS uses natively the Russian PZ-90 coordinate system [12] that can, however, be easily mapped to the WGS-84 system through a linear transformation [4]. This transformation is essential should the GLONASS and GPS be used in hybrid [5]. The purpose of the discussion is to remind that the coordinate system must also be specified in the assistance standard.

Table IV. Parameters required in representing modernized GPS and QZSS satellite orbits, their bit counts as well as scale factors. (u) denotes an unsigned parameter.

Parameters	Bits	Scale factor
ω , Argument of perigee	33	2^{-32} sc
M_0 , Mean anomaly	33	2^{-32} sc
e , Eccentricity	33(u)	2^{-34} sc
Δn , Mean motion correction	17	2^{-44} sc/s
$\dot{\Delta n}$, Rate of change of mean motion correction	23	2^{-57} sc/s ²
Ω_0 , Longitude of the ascending node	33	2^{-32} sc
$\dot{\Delta \Omega}$, Correction to the rate of change of longitude of the ascending node	17	2^{-44} sc/s
i_0 , inclination	33	2^{-32} sc
\dot{i} , Rate of change of inclination	15	2^{-44} sc/s
ΔA , Semi-major axis correction	26	2^{-9} m
\dot{A} , Rate of change of semi-major axis	25	2^{-21} m/s
C_{rs} , Correction to radius	24	2^{-8} m
C_{is} , Correction to inclination	16	2^{-30} rad
C_{us} , Correction to latitude	21	2^{-30} rad
C_{re} , Correction to radius	24	2^{-8} m
C_{ic} , Correction to inclination	16	2^{-30} rad
C_{uc} , Correction to latitude	21	2^{-30} rad

Another point concerning the GLONASS navigation model is that the clock model can be described by the standard GPS/Galileo clock model. It can easily be seen that the terms in the standard clock model have suitable range and resolution for also representing the GLONASS satellite clock model. However, in GLONASS the 2nd order clock acceleration term (see table II) is not required, which leads to non-efficient bit consumption should the GLONASS clock model be described using the standard GPS clock model. Hence, the 2nd order term must be made optional (i.e. its existence must depend upon the GNSS in question) in the standard clock model.

WAAS and EGNOS satellite positions are determined by extrapolating the satellite position using the first and second time derivatives for position. The satellite positions are directly given in the GPS-compatible WGS-84 ECEF-system. Tables VII and VIII summarize the orbit and clock models for WAAS and EGNOS. The clock models used in WAAS and EGNOS are easily represented by the standard model. However, it should be noted that as with GLONASS, the 2nd order term is not used. SBASs do not use the group delay parameter either.

Table V. Initial values for the initial value problem arising from the GLONASS satellite position calculation. Note that the coordinates are natively given in the PZ90 frame

Parameters	Bits	Scale factor
\underline{x} , SV coordinates	27	2^{-11} km
$\dot{\underline{x}}$, SV velocities	24	2^{-20} km/s
$\ddot{\underline{x}}$, Luni-Solar accelerations	5	2^{-30} km/s ²

Table VI. GLONASS satellite clock model.

Parameters	Bits	Scale factor
τ_n , bias	22	2^{-30} s
γ_n , 1 st order term	11	2^{-40} s/s
$\Delta \tau_n$, group delay	5	2^{-30} s

Comparing the GLONASS and SBAS navigation models shows that the models have only a small intersection. For instance, the SBAS model could be used to express the GLONASS satellite position in x and y coordinates, but not in the z coordinate. This is because the SBAS model has been optimized for the geostationary orbit. Moreover, in velocity coordinates the SBAS model does not have the required range and, on the other, in acceleration coordinates the resolution in the SBAS model is too coarse for GLONASS.

Finally, in the case of LAAS the only information required is the pseudolite coordinates and the clock model. Hence, an orbit representation using the Cartesian coordinate system, such as that utilized by GLONASS or SBAS, is preferred. However, neither of these has the required resolution, since one should be able to represent the pseudolite position at centimeter level [26].

Table VII. Position, velocity and acceleration for the SBAS satellite at a given instant. The satellite position is extrapolated using these parameters.

Parameters	Bits	Scale factor
x and y position	30	0.08 m
z position	25	0.40 m
x and y velocity	17	0.000625 m/s
z velocity	18	0.004 m/s
x and y acceleration	10	0.0000125 m/s ²
z acceleration	10	0.0000625 m/s ²

Table VIII. The SBAS satellite clock model.

Parameters	Bits	Scale
a_0 , bias	12	2^{-31} s
a_1 , 1 st order term	8	2^{-40} s/s

In principle, it is possible to express the position of a stationary object using the Keplerian orbit model. If the pseudolite position is given in the ECEF frame by $\underline{x} = [x_0 \ y_0 \ z_0]$, the position may be expressed with the GPS/Galileo Keplerian orbit model by the set of equations in (1).

$$\left\{ \begin{array}{l} t_{oe} = e = M_0 = \Omega_0 = \dot{i} = 0 \\ A^{1/2} = \sqrt{\|\underline{x}\|} \\ i_0 = \frac{1}{\pi} \cos^{-1} \left(\frac{y_0}{\sqrt{y_0^2 + z_0^2}} \right) \\ \omega = \frac{1}{\pi} \cos^{-1} \left(\frac{x_0}{\|\underline{x}\|} \right) \\ \Delta n = -\frac{1}{\pi} \sqrt{\frac{\mu}{\|\underline{x}\|^3}} \\ \dot{\Omega} = \frac{\dot{\Omega}_E}{\pi} \\ C_{rs} = C_{rc} = C_{is} = C_{ic} = C_{us} = C_{uc} = 0 \end{array} \right. , \quad (1)$$

where $\dot{\Omega}_E$ is the Earth rotation rate.

Although [2] shows that the resolution obtainable with the bits given in table I is sufficient for representing the pseudolite position at centimeter-level accuracy, the Keplerian model is still inherently unsuitable for describing the LAAS position. For instance, Δn has too few bits for keeping the pseudolite stationary. This is because Δn should be -0.000396 here, but the smallest number presentable with 16 bits and a scale factor of 2^{-43} is $-3.73e-9$. Therefore, more bits would be needed in order to have sufficient range – an unacceptable solution from the bandwidth point of view. The same applies to the HA Keplerian model as well. Moreover, in the case of the HA model, for instance, the reference semi-major axis would need to be modified.

The LAAS clock model can be represented using the standard clock model [26].

INTRODUCING THE MULTI-MODE NAVIGATION MODEL

The navigation models used by GPS, modernized GPS, GLONASS, Galileo, QZSS, SBAS and LAAS can be summarized by figure 2. There are three types of orbit models and two types of clock models.

The multi-mode navigation model can now be defined based on the presented considerations. One of the important findings is that the orbit and clock models must be decoupled. In the current AGPS implementation there is one orbit model (Keplerian) and one clock model (standard). However, once GLONASS is introduced in addition to GPS in the assistance standards, there will be two different orbit models (Keplerian and Cartesian), but only one clock model (standard). Hence, in the multimode navigation model the orbit model does not fix the clock model, but it can be selected independently. Therefore the superstructure of the multimode navigation model can be defined as in figure 3.

Figure 3 is interpreted so that for each GNSS for which assistance is given there is a message that contains a common element and navigation models for an arbitrary number of satellites belonging to the GNSS. In addition to navigation model, there may also be per satellite clock and orbit model degradation models, for instance. [19]

The common per GNSS element enables transferring some data fields or parts of the data fields from the navigation model to the common part. This induces bit savings, since redundant data is not duplicated in the navigation models. It turns out that bit consumption savings can be achieved in the Keplerian model in two parameters, in eccentricity and semi-major axis. Neither the Cartesian coordinate representation nor High-Accuracy Keplerian model lend themselves to further reduction in bit consumption.

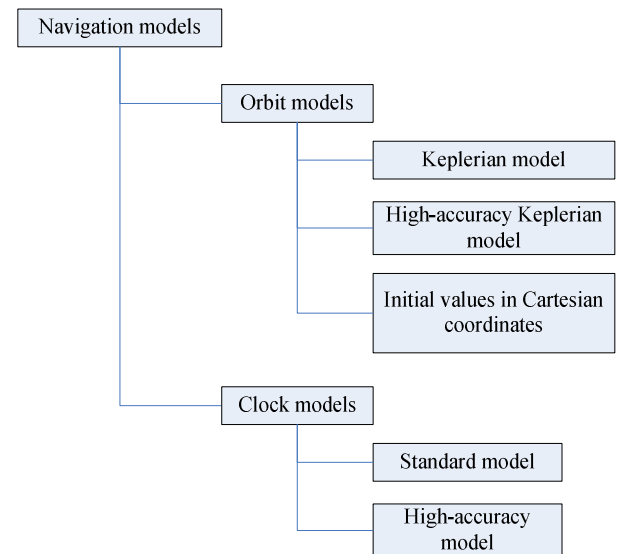


Figure 2. Different model types

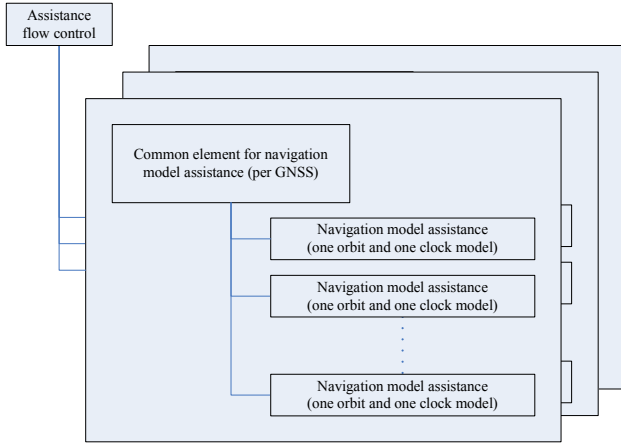


Figure 3. Navigation model assistance superstructure

Eccentricity in the Keplerian model

The European Space Agency (ESA) specifies [27] that the Galileo satellite orbits will have eccentricity of 0.002. Moreover, it is specified [8] that the GPS satellite eccentricity is 0.003 at maximum. Despite this, the eccentricity is represented using 32 bits and the scale factor of 2^{-33} , which allows for representing the range [0, 0.5). Since only a small part of the range is needed, bit saving may be achieved by dividing the eccentricity parameter into two parts: Most Significant Bits (MSB) and Least Significant Bits (LSB). This can be done by defining

e MSB: 7 bits , scale of 2^{-8} and range of [0, 0.496)
e LSB: 25 bits, scale of 2^{-33} and range of [0, 0.00391)

Combining these yields the original range [0, 0.499) and resolution. Hence, now by transferring the 7 MSBs to the common element, each satellite added after the first saves 7 bits.

Semi-major axis in Keplerian model

The square-root of semi-major is expressed with 32 bits (unsigned) using the scale factor of $2^{-19} \text{ m}^{1/2}$ resulting in the range [0, 8191.999) $\text{m}^{1/2}$. Now, Galileo [27] and GPS [8] have semi-major axes of 29 601 000 and 26 559 710 m, respectively. Moreover, it is stated [9] that the GPS satellite semi-major axis varies ± 65 km around the nominal (ΔA). The same can also be assumed of the Galileo satellite orbits. Therefore in the square-root of the semi-major axis there is a large redundant part. Hence, bit savings may be achieved by dividing the semi-major axis into two parts. Now,

$$\begin{aligned} \left\lceil \sqrt{A_0^{GPS} + \Delta A} \right\rceil &= 5160_{10} = 1010000101000_2 \\ \left\lfloor \sqrt{A_0^{GPS} - \Delta A} \right\rfloor &= 5147_{10} = 1010000011011_2 \\ \left\lceil \sqrt{A_0^{GAL} + \Delta A} \right\rceil &= 5447_{10} = 1010101000111_2 \\ \left\lfloor \sqrt{A_0^{GAL} - \Delta A} \right\rfloor &= 5434_{10} = 1010100111010_2 \end{aligned} \quad (2)$$

There are seven common MSBs in the case of GPS and six in the case of Galileo. Since the number of MSBs is fixed, it can be at most six (Galileo as the restricting case). Therefore, the semi-major axis can be represented by

$A^{1/2}$ MSB: 6 bits, scale of $2^7 \text{ m}^{1/2}$ and range [0, 8064.0) $\text{m}^{1/2}$
 $A^{1/2}$ LSB: 26 bits, scale of $2^{-19} \text{ m}^{1/2}$ and range [0, 127.999)

Note that even though it was assumed that the Galileo satellite semi-major axis changes only ± 65 km around the nominal, the above examination would hold as long as the Galileo delta term is smaller than 600 km. Moreover, it is noteworthy that if only GPS had been considered, 7 bits could have been moved to the common part.

In conclusion, if the model is Keplerian, the transfer of eccentricity and semi-major axis MSBs to the part common to all the satellites results in saving (7+6) bits/satellite after the first satellite.

PROPERTIES OF THE PER GNSS COMMON ELEMENT

The common element shown in figure 3 is specified in table IX. Should the navigation models included in the message be Keplerian models, the common part includes the MSBs for eccentricity and semi-major axis square root. In the case of the navigation models being other than Keplerian models, the MSBs are not included. This is indicated by the inclusion column, in which flag C (Conditional) is given for eccentricity and semi-major axis square-root. Conditionality denotes that whether or not MSBs are included depends on the navigation models in the message.

The common part also includes other information that is common to all the satellites in the message. This information may contain, for instance, MSBs of time-of-ephemeris [19]. However, this is beyond the scope of this article.

Table IX. The common element of navigation model assistance

Parameter	Bits	Scale Factor	Inclusion
<i>Navigation Model Flow Control (once per message)</i>			
Other common information	---	---	---
e MSB (if Keplerian model)	7	2^{-8}	C
$A^{1/2}$ MSB (if Keplerian model)	6	$2^7 \text{ m}^{1/2}$	C
<i>The following fields occur once per satellite</i>			
Other satellite-specific information	---	---	---
Orbit Model			M
Clock Model			M
Other possible models			

PROPERTIES OF THE PER SATELLITE ELEMENT

The part that is individual for each satellite consists of satellite-specific information (such as satellite identification [2] or in the case of GLONASS the frequency channel) and information on the orbit as well as clock models. Moreover, there may also be other satellite-specific models, such as orbit and clock model degradation models [19].

The different orbit and clock models are defined in Tables XII and XIII, respectively. The multi-mode structure is now revealed, because the different modes refer to different models. For instance, when describing satellite orbits, there are currently three different modes (or models), which are specified in table X.

Note that natively mode 1 can be used for GPS and Galileo, mode 2 for GLONASS, SBAS and LAAS and mode 3 for modernized GPS and QZSS.

The examination of mode 1 (Keplerian model) in table XIII shows that the model has been taken directly from GPS ICD [8]. However, as discussed earlier, the MSBs of eccentricity and semi-major axis square-root have been moved to the common part of the message. The satellite-specific LSBs are given in mode 1. The inclusion flag C (Conditional) expresses that the fields in the mode are conditional depending upon the selected model mode: if mode 1 is used, then all the fields defined in mode 1 must be present.

Table X. Different modes for orbit models

Mode	Orbit model
1	GPS/Galileo-style Keplerian orbit model
2	Cartesian coordinate presentation
3	Modernized GPS-style High-Accuracy Keplerian orbit model

In mode 2 (Cartesian model) there are two types of inclusion flags: Conditional and Optional (O). This is because mode 2 can be used with, for instance, LAAS, SBAS and GLONASS. They all need position coordinates and, hence, only position MSBs are conditional to mode 2. Other fields are conditional or optional to mode 2 depending upon the system.

If mode 2 is used to provide assistance for GLONASS, position coordinate MSBs as well as velocity and acceleration coordinates are required. Hence, with GLONASS all the fields except for position LSBs are conditional. Moreover, LAAS requires neither velocity nor acceleration, but only position coordinate MSBs and LSBs in order to fulfill strict resolution requirements (sub-dm). Velocity and acceleration coordinates are not required for LAAS and are, therefore, optional to mode 2, when using LAAS. Finally, SBASs use all the fields in mode 2. Position coordinate MSBs and LSBs are both needed in order to comply with the SBAS ephemeris. Moreover, velocity and acceleration components are needed for orbit extrapolation in time.

In mode 3 all the fields are conditional to the model. This is because mode 3 is the High-Accuracy Keplerian model from modernized GPS. There are no fields that could be left out and, hence, there are no optional fields.

Further, table XIII shows two additional modes, 4 and 5, that can be utilized for future orbit formats. These orbit formats might include, for instance, long-term orbit models. Having two additional modes in table XIII is exemplary. There is no upper limit for the number of modes.

The clock models are also specified in terms of modes specified in table XI. Natively, mode 1 can be used for GPS and Galileo and mode 2 for modernized GPS and QZSS. However, since the clock model can be freely chosen independent of the system, the same clock models (mode 1) can be used for GLONASS, SBAS and LAAS. This shows the strength of the current multimode approach, because the need to have only two different clock models reduces complexity. Note also that the clock model can occur multiple times. This is required for Galileo [11].

Table XI. Different modes for clock models

Mode	Clock model
1	Standard clock model
2	High-Accuracy clock model

Table XIII shows the two possible modes for the clock models. The first mode is the standard clock model usable for GPS, Galileo, GLONASS, SBAS and LAAS. Note that only the 0th and 1st order terms are conditional to mode 1 and the 2nd order and group delay terms are optional for the mode. This is because, for instance, SBAS requires neither the 2nd order nor group delay parameters.

Mode 2 is the high-accuracy clock model usable with modernized GPS and QZSS. The clock model terms as well as the group delay parameters are conditional to mode 2, since these parameters are required in every case. However, the need for the group delay terms between different signals depends on the receiver capabilities. Therefore these parameters are optional.

Table XII. The three orbit models for the multimode navigation model.

Parameter	Bits	Scale Factor	Inclusion
<i>Navigation Model shall include only one of the following presentations for the Satellite Navigation Model</i>			
<i>Satellite Navigation Model Using Keplerian Parameters (mode 1)</i>			
Other model-specific information	---	---	---
ω	32	2^{-31} sc	C
Δn	16	2^{-43} sc/s	C
M_0	32	2^{-31} sc	C
$\dot{\Omega}$	24	2^{-43} sc/s	C
e LSB	25(u)	2^{-33}	C
\dot{i}	14	2^{-43} sc/s	C
$A^{1/2}$ LSB	26(u)	2^{-19} m ^{1/2}	C
i_0	32	2^{-31} sc	C
Ω_0	32	2^{-31} sc	C
C_{rs}	16	2^{-5} m	C
C_{is}	16	2^{-29} rad	C
C_{us}	16	2^{-29} rad	C
C_{rc}	16	2^{-5} m	C
C_{ic}	16	2^{-29} rad	C
C_{uc}	16	2^{-29} rad	C

<i>Satellite Navigation Model Using Cartesian Coordinates (mode 2)</i>			
Other model-specific information			
X MSB	28	2^{-1} m	C
Y MSB	28	2^{-1} m	C
Z MSB	28	2^{-1} m	C
X LSB	7(u)	2^{-8} m	C/O
Y LSB	7(u)	2^{-8} m	C/O
Z LSB	7(u)	2^{-8} m	C/O
X'	26	2^{-12} m/s	C/O
Y'	26	2^{-12} m/s	C/O
Z'	26	2^{-12} m/s	C/O
X''	19	2^{-22} m/s ²	C/O
Y''	19	2^{-22} m/s ²	C/O
Z''	19	2^{-22} m/s ²	C/O
<i>Satellite Navigation Model Using High-Accuracy Keplerian Parameters (mode 3)</i>			
Other model-specific information			
ω	33	2^{-32} sc	C
Δn	17	2^{-44} sc/s	C
$\dot{\Delta n}$	23	2^{-57} sc/s ²	C
M_0	33	2^{-32} sc	C
$\dot{\Delta \Omega}$	17	2^{-44} sc/s	C
e	33(u)	2^{-34}	C
\dot{i}	15	2^{-44} sc/s	C
ΔA	26	2^{-9} m	C
\dot{A}	25	2^{-21} m/s	C
i_0	33	2^{-32} sc	C
Ω_0	33	2^{-32} sc	C
C_{rs}	24	2^{-8} m	C
C_{is}	16	2^{-30} rad	C
C_{us}	21	2^{-30} rad	C
C_{rc}	24	2^{-8} m	C
C_{ic}	16	2^{-30} rad	C
C_{uc}	21	2^{-30} rad	C
<i>Future Satellite Navigation Model (mode 4)</i>			
To be specified			
<i>Future Satellite Navigation Model (mode 5)</i>			
To be specified			

BENEFITS OF THE MULTI-MODE MODEL

The first benefit of the introduced navigation model assistance is the decoupling of the system from the choice

of the orbit and clock model formats. A straightforward implementation would have been to define the assistance in terms of the GNSS native orbit/clock model format individually for each new system [3]. However, the current solution allows for reusing the defined data structures for all the possible GNSSs. This is well shown, for example, with mode 2 in the orbit model. The same data structure is used for GLONASS, SBAS as well as for pseudolites. The arrangement reduces the system complexity significantly.

Secondly, the structure allows for an easy addition of any new system in the form of a new mode in either orbit models, clock models or both. This might be needed, for instance, if orbit extensions are added to the assistance standards. These might be implemented as a new mode supplying correction data to ephemerides. It should also be noted that the multi-mode navigation model can also be given multiple times for the same satellite. This allows sending multiple sets of, for example, ephemeris correction data to the terminal. This extends the validity period of the assistance up to several days [21].

Thirdly, although the currently introduced multi-mode navigation model is capable of supporting at least seven GNSSs with only three orbit and two clock modes, the navigation model is still compatible with the native broadcast ephemeris. In certain cases simple scaling operations are required, when mapping broadcast ephemeris to assistance, but essentially only copying is needed. This further reduces the complexity of the implementation.

Finally, non-broadcast formats are also supported. Mode 2 using Cartesian coordinate representation is suitable, for instance, for transferring orbit data from the International GNSS Service (IGS, [25]).

CONCLUSIONS

The paper reviewed the orbit and clock models, collectively called navigation models, for GPS, modernized GPS, GLONASS, Galileo, QZSS, WAAS, EGNOS and pseudolites. Based on these considerations, a 3GPP-compliant multi-mode navigation model was defined, which allows for providing the terminal with the navigation model assistance for various GNSSs.

The multi-mode navigation model decouples orbit and clock models from each other. Therefore it is, for example, possible to describe the GLONASS navigation model in terms of the GLONASS native orbit model and GPS native clock model. This permits great flexibility in using the assistance data. Moreover, the format also allows for easy addition of new modes (or models) to the assistance standard, if the need arises.

Table XIII. The two clock models of the multimode navigation model

Parameter	Bits	Scale Factor	Inclusion
<i>Standard Satellite Clock Model (mode 1)</i>			
Other model-specific information			
af_2	12	2^{-65} s/s^2 , if Galileo 2^{-55} s/s^2 , otherwise	C/O
af_1	18	2^{-45} s/s	C
af_0	28	2^{-33} s	C
T_{GD}	10	2^{-32} s	O
<i>High-Accuracy Satellite Clock Model (mode 2)</i>			
Other model-specific information			
af_2	10	2^{-60} s/s^2	C
af_1	20	2^{-48} s/s	C
af_0	26	2^{-35} s	C
T_{GD}	13	2^{-31} s	C
ISC_{L1CA}	13	2^{-35} s	O
ISC_{L2C}	13	2^{-35} s	O
ISC_{L5I5}	13	2^{-35} s	O
ISC_{L5Q5}	13	2^{-35} s	O
ISC_{L1CP}	13	2^{-35} s	O
ISC_{L1CD}	13	2^{-35} s	O

ACKNOWLEDGMENTS

The authors would like to acknowledge their colleagues Ismo Halivaara, Jani Käppi and Kimmo Alanen from Nokia Technology Platforms.

REFERENCES

- [1] Syrjärinne, J. & Wirola, L. 2006. Setting a New Standard – Assisting GNSS Receivers That Use Wireless Networks. October 2006 issue of InsideGNSS
- [2] Wirola, L. & Syrjärinne, J. 2006. Introduction to Multi-Mode Navigation Model. Presented in 3GPP TSG GERAN2#29bis, 22nd-24th May, Sophia-Antipolis, France, <http://www.3gpp.org>
- [3] Alanen, K. & Syrjärinne, J. 2007. From Divergence to Convergence, GNSS Assistance in Network Carrier Independent format. Proceedings of Institute of Navigation National Technical Meeting, 22nd-24th January, San Diego, USA
- [4] Rossbach, U. 2000. Positioning and Navigation using the Russian Satellite System GLONASS. Dissertation, der Universität der Bundeswehr München

- [5] Dodson, A. & Moore, T. & Baker, D. & Swann, J. 1999. Hybrid GPS+GLONASS. GPS Solutions volume 3, number 1, pages 32-41
- [6] Klimov, V. & Revnivykh, S. & Kossenko, V. & Dvorkin, V. & Tyulyakov, A. & Eltsova, O. 2005. Status and Development of GLONASS. Proceedings of GNSS-2005, 19th-22nd July, Munich, Germany
- [7] Averin, S. 2006. GLONASS system: Present Day and Prospective Status and Performance. Proceedings of GNSS-2006, 7th-10th May, Manchester, United Kingdom
- [8] IS-GPS- 200 Revision C, IRN-200C-003: NAVSTAR GLOBAL POSITIONING SYSTEM Navstar GPS Space Segment/Navigation User Interface, 11th October 1999
- [9] IS-GPS-705, IRN-705-003: NAVSTAR GLOBAL POSITIONING SYSTEM Navstar GPS Space Segment/User Segment L5 Interfaces, 22nd September 2005
- [10] IS-GPS-800 draft Navstar GPS Space Segment/User Segment L1C Interfaces, 19th April 2006
- [11] Galileo OS SIS ICD draft 0
- [12] GLONASS interface document version 5.0. 2002 Moscow
- [13] IS-QZSS ver 0.0. 2007. Japan Aerospace Exploration Agency
- [14] FAA-2892b: Specification for Wide Area Augmentation System (WAAS), August 31st 2001
- [15] E-TN-ITF-E31-0008-ESA, ESTB SIS User Interface Description, 20th June 2006
- [16] 3rd Generation Partnership Project, <http://www.3gpp.org>
- [17] 3GPP TS 04.31 and 44.035, Location Services (LCS); Mobile Station (MS) – Serving Mobile Location Center (SMLC) Radio Resource LCS Protocol (RRLP), <http://www.3gpp.org>
- [18] 3GPP TS 25.331, Radio Resource Control (RRC) protocol specification, <http://www.3gpp.org>
- [19] Nokia. 2006. Comparisons between Different A-GNSS Options G2-060276, presented in 3GPP TSG-GERAN Meeting #31bis, Turin, Italy, 16th-20th October, <http://www.3gpp.org>
- [20] Meeting report: Report of TSG GERAN meeting #32, Sophia-Antipolis, France, 13th-17th November 2006 <http://www.3gpp.org>
- [21] SiRF Technology, 2007. A-GNSS, Orbit Extension. Presented in 3GPP TSG GERAN meeting #33, Seoul, Korea, 12th-16th February. <http://www.3gpp.org>
- [22] OMA-TS-ULP-V1-0-20050719-C, User Plane Location Protocol, http://www.openmobilealliance.org/release_program/supl_v1_0.html
- [23] Analyst report: Global Handsets, GPS/A-GPS Phone Sales. Strategy Analysts, November 2006
- [24] News article: Now we are three, Third block IIR-M satellite launched. GPS World, November 19th 2006
- [25] International GNSS Service, IGS, <http://igsceb.jpl.nasa.gov>
- [26] Private communications with Mårten Ström, Space Systems Finland, February 2006
- [27] Galileo Project Office, ESA-EUING-TN/10206 Specification of Galileo and Giove Space Segment Relevant for Satellite Laser Ranging, 3rd July 2006
- [28] Press release: Next Step in the Galileo Program. Galileo Joint Undertaking , Brussels 30th November 2006

Publication 6

L. Wirola and J. Syrjärinne. GLONASS Orbits in GPS/Galileo-style Ephemerides for Assisted GNSS. In *Proceedings of ION National Technical Meeting*, San Diego, USA, January 28th-30th, pp. 1032–1039, (2008)

Copyright ©2008 ION. Reprinted, with permission, from the proceedings of ION National Technical Meeting 2008.

GLONASS Orbits in GPS/Galileo-style Ephemerides for Assisted GNSS

Lauri Wirola, *Nokia Inc., Finland*
Jari Syrjärinne, *Nokia Inc., Finland*

BIOGRAPHY

Lauri Wirola, M.Sc., received his Master of Science degree from Tampere University of Technology, Finland, in 2005. He majored in electrophysics and minored in technical physics and industrial economics. In 2004-05 he worked on acoustics at the Nokia Research Center. In 2005 he joined the positioning group at Nokia. His present research interests include multi-GNSS positioning, RTK and PPP as well as A-GNSS standardization. He is currently undertaking postgraduate studies in modern electromagnetism and mathematics.

Jari Syrjärinne, Ph.D., received his M.Sc. in 1996 and his Ph.D. in 2001, both from Tampere University of Technology, Finland, majoring in digital signal processing and applied mathematics. Between 1996 and 1998 he worked for the Tampere University of Technology Signal Processing Laboratory in the areas of data fusion and target tracking, and since 1999 has been with Nokia Inc. He is currently involved in work on A-GNSS and algorithms for hybrid positioning.

ABSTRACT

The 3GPP (Third Generation Partnership Project) Release 7 of GSM (Global System for Mobile Communications) and UMTS (Universal Mobile Telecommunications System) cellular standards include major modifications as to how AGNSS (Assisted GNSS) assistance data is transferred from the cellular network to the cellular terminal. One of the major changes in the coming 3GPP GSM/UMTS Release 7 is the introduction of multi-mode navigation model. This navigation model supports all the native formats, including Galileo ephemeris.

Multi-mode navigation model also enables an interesting opportunity to provide navigation model assistance in non-native formats. For instance, the multi-mode model allows, if desired, representing GLONASS orbits and clocks using GPS/Galileo Keplerian ephemeris and GPS clock model. Such a solution simplifies the network and terminal implementation as well as partly harmonizes performance between different GNSSs.

The current work examines, if GLONASS orbits and clocks can be represented with the bit counts and scale factors available in GPS/Galileo ephemeris. The analysis is based on the final precise orbits for GPS and GLONASS as well as clocks for GPS from International GNSS service over

19 weeks. The final GLONASS clocks were obtained from the Russian Mission Control Centre, MCC. Included in the analysis are GPS weeks 1437-1455. Keplerian parameters and clock models were fitted to these data in 4-h blocks.

The results show that GLONASS orbits and clocks can be expressed in 4-h blocks using GPS/Galileo-style ephemeris without difficulties. The study also reveals interesting differences in the characteristics of GPS and GLONASS orbits.

INTRODUCTION

The annual sales of AGNSS-enabled (Assisted GNSS) handsets is estimated to rise to 400 million units by 2011 [1]. Currently the size of the market is approximately 100 million units annually. High growth requires developing constantly more efficient and capable methods to improve user experience in terms of availability, accuracy and short time-to-first-fix. The assistance data available from the network are a significant factor affecting the user experience. The advantages and benefits of assistance are discussed in [23].

Figure 1 shows the high-level view of AGNSS-architecture. The core of the architecture is the AGNSS-server, or more precisely, server centers that are geographically distributed. These centers serve the AGNSS-subscribers in each geographical area. Assuming that the AGNSS-terminal is to receive assistance over user plane (IP-network) the terminal takes an HSPA (High-Speed Packet Access) connection to the pre-set server and requests for the assistance data. The assistance data is then delivered to the terminal as specified in the associated standards.

The AGNSS-server may obtain its data from various sources. These may include physical GNSS-receivers distributed geographically (left hand side in figure 1). These receivers can provide integrity information as well as broadcast ephemerides to the AGNSS server for distribution. On the other hand, the orbit and clock models (as well as other data) can originate from an external service providing, for instance, precise ephemerides and orbit/clock predictions (right hand side in figure 1). Such services include International GNSS Service or IGS [17]. In case predictions are available, AGNSS-enabled terminals can be provided with extended ephemerides, in which case the terminal does not need to connect to the assistance server in the beginning of each positioning session. This improves user experience due to the time saved in not having to set up an HSPA con-

nection and download the assistance [19]. Currently it is only possible to provide assistance for GPS L1 in GSM (Global System for Mobile Telecommunications) and UMTS (Universal Mobile Telecommunications System) networks. In GSM the assistance is specified in the Radio Resource LCS (Location Services) Protocol (RRLP, [2]) and in UMTS in the Radio Resource Control (RRC, [3]). Moreover, there are also user plane solutions, such as Open Mobile Alliance (OMA) Secure User Plane Location (SUPL, [4]), that effectively carry the same information elements over packet networks that are specified for circuit-switched networks in RRLP and RRC.

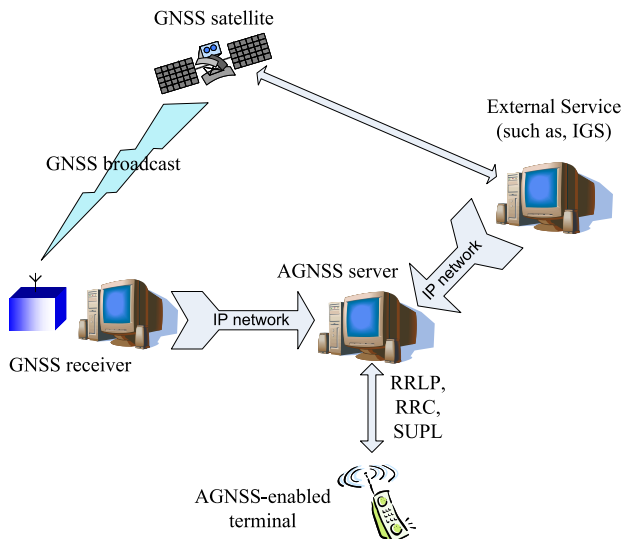


Figure 1: AGNSS architecture. RRLP, RRC and SUPL refer to different carriers used to deliver assistance.

However, due the upcoming changes in the GNSS infrastructure [23], such as modernization of GPS and GLONASS as well as the introduction of Galileo amongst others, the 3GPP standardization body [5] accepted a proposal which opened the way for the addition of new GPS bands as well as other GNSSs to the assistance standard [6] in autumn 2006. This decision concerned RRLP only, but the same solution was later approved into RRC [7] as well as SUPL 2.0 [8].

The AGNSS introduces common and per GNSS elements into the standards. The superstructure is detailed in [20]. The common elements are GNSS-independent and include, for instance, ionosphere model and reference location. In the future, for instance, troposphere models or Earth-Orientation Parameters (EOP) can be added without obstacles.

The per GNSS elements, on the other hand, are by definition GNSS-dependent (as well as signal-dependent) and include differential corrections, real-time integrity, GNSS-common time relation, data bit assistance, reference measurements as well as orbit and clock models (ephemerides). The variety of orbit models found in the GNSS families, however, cause problems in assistance standards, for instance, due to the high number of required information elements and unsynchronized update intervals. The dif-

ferent orbit and clock models are shown in figure 2 and thoroughly discussed in [23] and [22]. This study shows that GLONASS orbits and clocks can be expressed in terms of GPS/Galileo models, which possibility significantly reduces the system complexity.

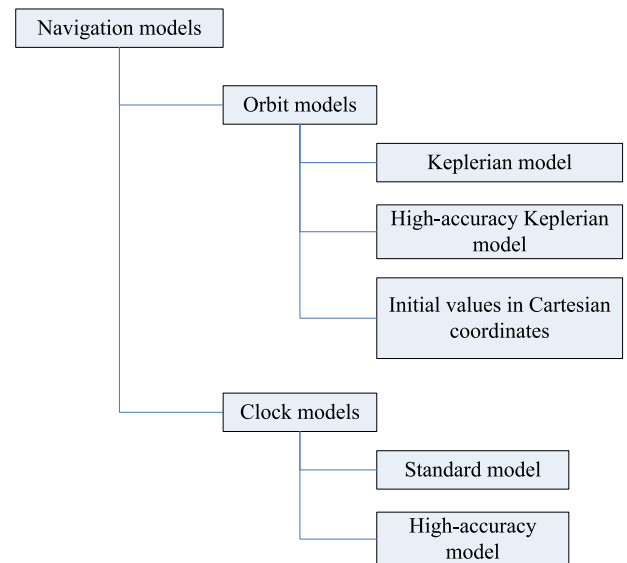


Figure 2: The division of orbit and clock models. Reproduced from [23].

ORBIT AND CLOCK MODELS

The variety of orbit and clock models found in the current and future GNSSs is shown in figure 2. In high level there are three types of orbit models and two types of clock models. The three orbit models are the standard Keplerian representation utilized by GPS [9] as well as Galileo [10], the high-accuracy Keplerian parameterization utilized by modernized GPS [11] as well as Japanese Quazi-Zenith Satellite System (QZSS, [12]) and the Cartesian coordinate representation utilized in Space-Based Augmentation Systems (WAAS, EGNOS) as well as in Russian GLONASS [13]. A full account of the orbit models is given in [23].

The two clock models are the standard clock model utilized by Galileo, which model can also be utilized for GPS using conditional scale factors [23], and the high-accuracy clock model utilized by modernized GPS.

Although the upcoming Release 7 of RRLP supports only GPS L1 and Galileo, the new assistance branch is future-proof in such a way that it can easily be updated to support the rest of the required models [21]. Having performed the update, assistance could be delivered for at least seven GNSSs [22].

The information elements used to carry orbit and clock models also include a non-broadcast flag that identifies, if the orbit/clock models are not obtained from the broadcast, but from an external service [21]. An example of such a situation is that the orbit predictions for GLONASS SVs (SV, Space Vehicle) were used to represent orbits for GLONASS SVs in Keplerian parameterization.

NATIVE GLONASS EPHEMERIS

The native GLONASS orbit model differs significantly from the GPS orbit model. The GLONASS broadcast includes the satellite position $\underline{x}(toe)$ and velocity $\dot{\underline{x}}(toe)$ at time-of-ephemeris or toe (t_b in [13]) as well as the acceleration of the satellite due to luni-solar forces (\underline{a}^{LS}). The life-time of the ephemerides is 30, 45 or 60 minutes. [13] The broadcast parameters $\underline{x}(toe)$ and $\dot{\underline{x}}(toe)$ are used as initial conditions in the integration of equations of motion given in equation 1.

$$(1) \begin{bmatrix} \partial_t \underline{x} \\ \partial_t \underline{v} \end{bmatrix} = \begin{bmatrix} -\frac{\mu}{r^3} \underline{x}_1 - \frac{3}{2} J_0^2 \frac{\mu a_e^2}{r^5} \underline{x}_1 \left(1 - \frac{5\underline{x}_3^2}{r^2}\right) + \omega_e^2 \underline{x}_1 + 2\omega_e \underline{v}_2 + \underline{a}_1^{LS} \\ -\frac{\mu}{r^3} \underline{x}_2 - \frac{3}{2} J_0^2 \frac{\mu a_e^2}{r^5} \underline{x}_2 \left(1 - \frac{5\underline{x}_3^2}{r^2}\right) + \omega_e^2 \underline{x}_2 - 2\omega_e \underline{v}_1 + \underline{a}_2^{LS} \\ -\frac{\mu}{r^3} \underline{x}_3 - \frac{3}{2} J_0^2 \frac{\mu a_e^2}{r^5} \underline{x}_3 \left(3 - \frac{5\underline{x}_3^2}{r^2}\right) + \underline{a}_3^{LS} \end{bmatrix}$$

where t is time, $r = \|\underline{x}\|_2$, \underline{v} velocity, $\mu = G \cdot M_e$, G gravitational constant, M_e the Earth mass, a_e the semi-major axis of Earth, J_0^2 the 2nd zonal harmonic of the geopotential and ω_e the Earth rotation rate. The Luni-Solar acceleration components \underline{a}^{LS} are constant over the ephemeris life-time. The satellite clock model is given by

$$(2) \quad \Delta t^s(t) = \tau^s(toe) - \gamma^s(toe) \cdot (t - toe),$$

where $\tau^s(toe)$ and $\gamma^s(toe)$ are the zeroth and first order clock correction terms, respectively, as specified in [13]. From the AGNSS point-of-view the GLONASS ephemeris has certain issues. First of all, the high update-rate requires the terminal to request assistance for GLONASS SVs significantly more often than for GPS SVs. This results in high amount data traffic, increased power consumption and cost to AGNSS subscribers. Moreover, the unsynchronized updates for different GNSSs add complexity in the network as well as in the terminal. Secondly, there is currently a major performance gap between GPS and GLONASS orbit models. While the GPS orbits are accurate to few meters RMS [18], the current GLONASS orbits have 5-meter RMS accuracy in radial direction. However, the accuracy in the radial component is expected to improve to 1.5 meters RMS with GLONASS-M [13].

NATIVE GPS/GALILEO EPHEMERIS

The GPS/Galileo ephemeris and the associated algorithm for the calculation of the SV position is well-known. The required parameters and their ranges are given in table I. The position of the SV may be calculated by

$$(3) \quad \underline{x}(t, toe, \underline{p}) = \underline{x}(t, toe, \sqrt{a}, e, \Omega_0, \omega, i_0, M_0, \Delta n, \partial_t \Omega, \partial_t i, C_{us}, C_{uc}, C_{rs}, C_{rc}, C_{is}, C_{ic})$$

where a is the semi-major axis of the orbit, e the orbit eccentricity, Ω_0 right ascension, ω argument of perigee, i_0 the inclination, M_0 mean anomaly, Δn mean motion correction, $\partial_t \Omega$ rate of change of right ascension, $\partial_t i$ rate

of change of inclination and $C_{us}, C_{uc}, C_{rs}, C_{rc}, C_{is}, C_{ic}$ the harmonic correction terms. The algorithm $\underline{x}(t, toe, \underline{p})$ is given in [9]. \underline{p} is the vector of the ephemeris parameters. Finally, the SV clock model is given in equation 4.

$$(4) \quad \Delta t^s(t) = a_{f0} + a_{f1} \cdot (t - toe) + a_{f2} \cdot (t - toe)^2,$$

where a_{f0}, a_{f1} and a_{f2} are the 0th, 1st and 2nd order coefficients, respectively. The associated bit counts and ranges are given in table II.

Coefficient	Range
$\Omega_0, \omega, M_0, i_0$	[-0.5, 0.5] sc
\sqrt{a}	$[0, 2^{32} - 1] \cdot 2^{-19} \text{ m}^{\frac{1}{2}}$
e	$[0, 2^{32} - 1] \cdot 2^{-33}$
Δn	$[-2^{15}, 2^{15} - 1] \cdot 2^{-43} \text{ sc/s}$
$\partial_t i$	$[-2^{13}, 2^{13} - 1] \cdot 2^{-43} \text{ sc/s}$
$\partial_t \Omega$	$[-2^{23}, 2^{23} - 1] \cdot 2^{-43} \text{ sc/s}$
C_{is}	$[-2^{15}, 2^{15} - 1] \cdot 2^{-29} \text{ rad}$
C_{ic}	$[-2^{15}, 2^{15} - 1] \cdot 2^{-29} \text{ rad}$
C_{rs}	$[-2^{15}, 2^{15} - 1] \cdot 2^{-5} \text{ m}$
C_{rc}	$[-2^{15}, 2^{15} - 1] \cdot 2^{-5} \text{ m}$
C_{us}	$[-2^{15}, 2^{15} - 1] \cdot 2^{-29} \text{ rad}$
C_{uc}	$[-2^{15}, 2^{15} - 1] \cdot 2^{-29} \text{ rad}$

Table I: Orbit parameters and ranges. sc denotes semi-circle.

Coefficient	Range
a_{f0} (GPS)	$[-2^{21}, 2^{21} - 1] \cdot 2^{-31} \text{ s}$
a_{f1} (GPS)	$[-2^{15}, 2^{15} - 1] \cdot 2^{-43} \text{ s/s}$
a_{f2} (GPS)	$[-2^7, 2^7 - 1] \cdot 2^{-55} \text{ s/s}^2$
a_{f0} (Galileo)	$[-2^{27}, 2^{27} - 1] \cdot 2^{-33} \text{ s}$
a_{f1} (Galileo)	$[-2^{17}, 2^{17} - 1] \cdot 2^{-45} \text{ s/s}$
a_{f2} (Galileo)	$[-2^{11}, 2^{11} - 1] \cdot 2^{-65} \text{ s/s}^2$

Table II: Clock parameters and ranges for GPS and Galileo.

PRECISE ORBITS AND CLOCKS FROM IGS AND MCC

The International GNSS Service [17] provides orbit and clock products for GPS and GLONASS. The products are listed in table III. All the orbit products have a 15-min interval while the GPS clock data intervals are 15, 5 and 5 minutes in ultra-rapid, rapid and final products, respectively. For GLONASS the clock data interval is 15 min. Note, however, that the GLONASS clock data are identical to GLONASS broadcast. [17]

Product	Accuracy
GPS Ultra-Rapid (predicted)	10 cm / 5 ns
GPS Ultra-Rapid (observed)	< 5 cm / 0.2 ns
GPS Rapid (17-h latency)	< 5 cm / 0.1 ns
GPS Final (13-d latency)	< 5 cm / <0.1 ns
GLONASS Final (2 weeks)	15 cm

Table III: The IGS products as well as the expected orbit and clock accuracies.

The Russian Mission Control Center, MCC [14], provides both rapid and final orbits as well as clocks for GLONASS and GPS. The rapid data seems to become available in the afternoon next day. The final data from MCC is available with a typical latency of 5-7 days.

In the current study the GPS/Galileo ephemerides were fitted to the final GLONASS orbit (from IGS) and clock (from MCC) products.

METHOD

The optimization of Keplerian parameters was performed in two phases. In the first phase the four angular parameters (Ω_0 , ω , M_0 and i_0), e and \sqrt{a} were initialized using a grid search over the applicable search space. The correction terms (Δn , $\partial_t \Omega$, $\partial_t i$) and the harmonic correction coefficients were set to zero in the first phase of the optimization.

The grid size for the angular parameters was 5 degrees in the range [0,355] degrees. The grid size for eccentricity was 0.005 and the range [0,0.03]. Finally, the grid size for a was 500 meters in the range [$a_{GLONASS}-65$ km, $a_{GLONASS}+65$ km]. The 65-km interval was taken from GPS-ICD-800 [11]. The interval was also believed to be suitable for GLONASS semi-major axis variation.

The difference between the final IGS orbit and the orbit calculated based on the grid point was evaluated. The second phase of the optimization was initialized with the parameter set having the smallest deviation from the IGS orbit data in the sense of infinity norm, i.e.

$$(5) \quad \min_{\hat{p}} \max_i \|\underline{x}(t_i) - \hat{\underline{x}}(t_i, toe, \hat{p})\|_{\infty},$$

where $i = 1 \dots 17$, since a 4-h block consists of 17 data points with 15-min intervals. Moreover, \hat{p} denotes the reduced ephemeris parameter set (with correction terms clamped to zero) and $\hat{\underline{x}}(t_i, toe, \hat{p})$ the orbit calculated based on the parameter set \hat{p} . $\underline{x}(t_i)$ is the final orbit data from IGS.

Having performed the initialization, all the parameters were optimized using the conjugate gradient method with the termination conditions

$$(6) \quad \begin{aligned} \|\underline{x}(t_i) - \hat{\underline{x}}(t_i, toe, \hat{p})\|_{\infty} &< \epsilon_1, \forall i \\ \|\hat{\underline{x}}(t_i, toe, \hat{p}_{k+1}) - \hat{\underline{x}}(t_i, toe, \hat{p}_k)\|_{\infty} &< \epsilon_2, \forall i \end{aligned}$$

with ϵ_1 set to 0.5 meters and ϵ_2 set to 1 cm. \hat{p} is the optimized parameter set and k the iteration counter. If the 2nd condition fulfilled without the first condition being fulfilled, the data set was discarded. Such divergence may be caused by inappropriate initial conditions. A typical residual plot is given in figure 3.

The approach was verified by calculating orbit data at 15-min intervals from the GPS broadcast ephemeris and fitting ephemeris parameters to that orbit data. The obtained parameter set was then compared to the original set.

In addition, 2nd degree polynomials were fitted to clock data in 4-h blocks in the least-squares sense. Also, 4-h clock predictions and their usability was examined.

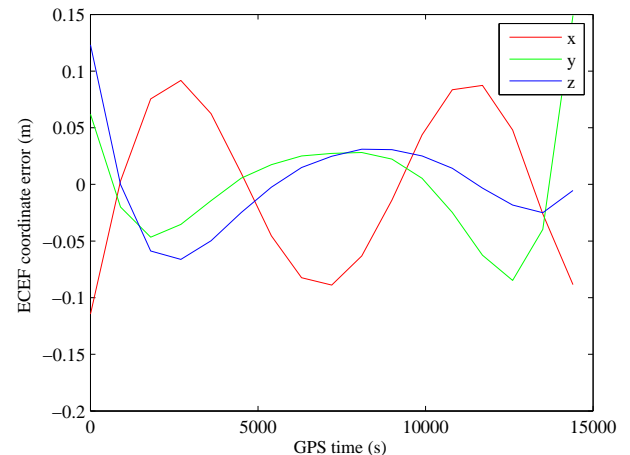


Figure 3: Typical fit residuals for a GLONASS orbit over a 4-h fit.

RESULTS

Orbit parameters

The following analysis is based on 2039 sets of 4-h fits of GLONASS orbits into the GPS/Galileo-style orbit parameterization. The temporal span of the orbit data is 19 weeks ranging from the GPS week 1437 to the week 1455. On each day for each SV two 4-h fits were made: the first between 00:00-04:00 (GPS time) and the second between 12:00-16:00.

The maximum residuals in each spatial directions were approximately 0.5 meters. The mean as well as median residuals were in the order of decimeter in each spatial direction. A typical residual plot is shown in figure 3.

Since the angular parameters (Ω_0 , ω , M_0 and i_0) always lie in the range $[-\pi, \pi)$, they are not of any concern in the following analysis, since they necessarily fit into the GPS/Galileo ephemeris. Moreover, the square-root of the semi-major axis (\sqrt{a}) and eccentricity (e) are well-defined physical parameters that cannot overflow from the parameter ranges available in GPS/Galileo ephemeris. However, they reveal other interesting characteristics about the orbits and are discussed in the following chapters. All the other parameters (Δn , $\partial_t \Omega$, $\partial_t i$ as well as harmonic correction terms) may overflow from the parameter ranges fixed by the bit counts and scale factors in the GPS [9] and Galileo ICDs [10].

Moreover, in order to obtain data for comparison between GPS and GLONASS orbits, the GPS broadcast ephemerides starting 1st July and ending 1st December 2007 were analyzed. The range of GPS parameters in normalized scale is shown in figure 4 in blue. The range [-1,1] denotes the full range representable by the bit count and scale factor available for the parameter. Naturally, eccentricity and square-root of semi-major axis have a range of [0,1]. Moreover, the range [0,1] for eccentricity does not denote the true parameter range [0,0.5), but the effective parameter

range [0,0.03) as given in the GPS ICD [9].

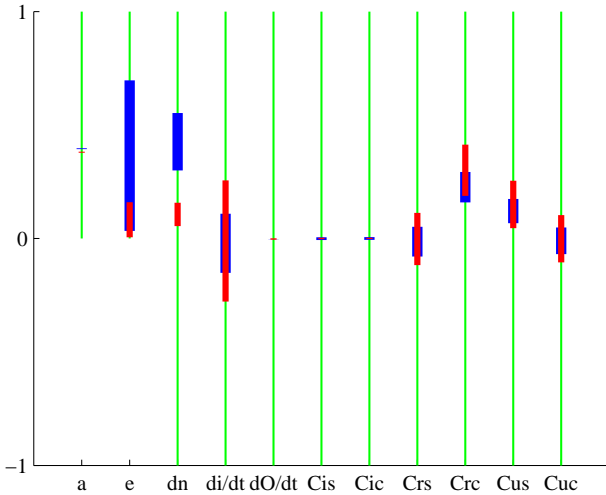


Figure 4: RED: Span of GLONASS ephemerides (fitted). BLUE: Span of GPS ephemerides (broadcast). The range [-1,1] denotes the full range as representable by GPS/Galileo ephemeris.

Semi-major axis: In the data analyzed the semi-major axes for GLONASS orbits vary 700 meters between 25507965 and 25508665 meters. The mean value is 25508141 m, which is 1051.569 km below the nominal GPS orbit. On the other hand, the GLONASS ICD [13] states that the orbit altitude is 19100 km. Given the Earth semi-major axis of 6378.136 km, the nominal GLONASS orbit should be 1081.574 km below the nominal GPS orbit. The reason for this 30-km discrepancy is unknown.

In the GPS/Galileo ephemeris the square-root of the semi-major axis is presented by 32 bits and a scale factor of $2^{-19} m^{\frac{1}{2}}$. Hence, the range presentable is approximately [0,67108] km. Therefore, the GLONASS semi-major axis can be represented without difficulties:

$$(7) \quad \lfloor \sqrt{25508141} \cdot 2^{19} \rfloor = 10011101110101000111011111000110_2,$$

where the least-significant bit represents resolution of 1 cm.

Eccentricity: Representing the GLONASS orbit eccentricity in GPS/Galileo-type ephemeris does not cause concerns either. While the GPS ICD [9] quotes that the effective range of eccentricity for GPS SVs is [0,0.03], the bit count (32) and the scale factor (2^{-33}) allow for representing the range [0,0.5). However, in the GLONASS orbits analyzed the eccentricity varies in the range [0.00017,0.0048]. These values are considerably smaller than eccentricities observed for GPS, which lie in the range [0.001,0.02] in the ephemerides analyzed. The difference is clearly visible in figure 4. Therefore, it can be concluded that the GLONASS orbits are significantly more circular than GPS orbits.

Mean motion correction: The purpose of the mean motion correction is to obtain the correct angular velocity for the SV. However, from the Newtonian mechanics it is known that for a perfectly circular orbit the mean motion is given by

$$(8) \quad n = \sqrt{\frac{\mu}{a^3}} + \Delta n,$$

with $\Delta n = 0$. Due to non-zero eccentricity, the GPS/Galileo ephemeris includes the mean motion correction, or Δn . Now, figure 4 clearly shows, how the size and the range of mean motion corrections is considerably smaller for GLONASS than GPS. This is attributable to the higher circularity of GLONASS orbits: the smaller eccentricity requires smaller mean motion correction.

Rate-of-change parameters: The rate-of-change parameters refer to $\partial_t i$ and $\partial_t \Omega$. Figure 4 shows that they both fit well within the range available in GPS/Galileo ephemeris. The behavior of $\partial_t \Omega$ is very similar in GPS and GLONASS. However, $\partial_t i$ has somewhat greater range in GLONASS. This may be attributable to the greater inclination of GLONASS orbits (64.8°) as compared to the GPS inclination (approximately 55°) as well as to the approximately 1000-km difference in orbit heights.

Harmonic corrections: As with all the other parameters, the harmonic correction coefficient do not raise any concerns. Figure 4 shows that all the six correction terms behave in a similar fashion in GPS and GLONASS. However, the GLONASS orbits seem to require somewhat greater harmonic corrections terms than GPS orbits. The behaviour may be due to the same reasons as given above for the differences in the rate-of-change parameters.

Clock parameters

Table II summarizes the clock model coefficients in the GPS/Galileo ephemeris. The GPS/Galileo clock model as defined in equation 4 is the 2nd order polynomial while the GLONASS clock model is of the 1st order. However, the life-time of the GLONASS clock model is only 30-60 minutes, whereas the GPS/Galileo clock model suffices for 4-6 hours.

Figure 5 shows the GLONASS and GPS clock coefficients in the similar fashion as the orbit parameters were presented in figure 4. The analysis is based on 2768 4-h fits of 2nd order polynomials to the final GLONASS clock data from MCC. The median of the maximum residuals is 1.1 ns (0.33 m). The analysis for GPS clocks, on the other hand, is based on approximately 60000 sets of broadcast clock models having a temporal span from 1.7. to 1.12.2007. It is interesting to note that during this interval the 2nd order coefficients (a_{f2}) were zero across all the sets and all the GPS SVs.

Figure 5 shows that the GLONASS clocks can be represented by GPS clock model. Interestingly, while the 0th and 1st order terms behave in similarly in GPS and GLONASS clocks, the 2nd order coefficient shows significant difference. It was noted that the 2nd order parameter was zero in the GPS ephemerides examined. However, 4-h fits to GLONASS clocks require a 2nd order polynomial. This shows that the GLONASS clocks are somewhat more unstable than GPS clocks.

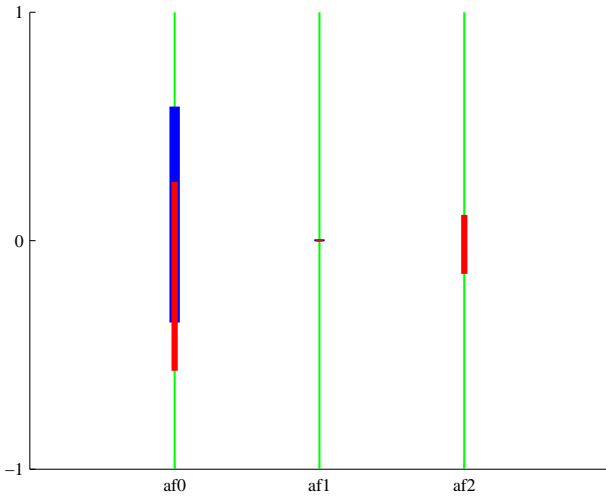


Figure 5: RED: Span of GLONASS clock parameters (fitted). BLUE: Span of GPS clock parameters (broadcast). Range [-1,1] denotes the maximum range in the GPS clock model.

CLOCK PREDICTIONS

While orbit predictions can be made at high accuracy (see table III), the clock predictions are more demanding. For instance, CODE [15] produces 24-h and 48-h orbit but not clock predictions for GLONASS. However, the ability to provide assistance for GLONASS for 4 hours is dependent upon the capability to obtain clock predictions in addition to orbit predictions.

The predictability of GLONASS clocks using a very simple approach was tested based on the 4-h fits made earlier. The method used was to utilize the 4-h fits (2nd order polynomial) to predict the clock for the next four hours. The prediction error was then compared to the MCC final data. In sum, 1015 4-h predictions were tested.

Figure 6 shows the cumulative distribution of clock prediction errors for GLONASS and GLONASS-M satellites separately. Figure shows that clocks on-board GLONASS-M satellites are more predictable (stable) than the older ones on-board conventional GLONASS SVs. For GLONASS-M the prediction error is smaller than 10 meters in 97.6% of the cases.

Figure 7 shows an exemplary comparison between the broadcast clock model, final clock data from MCC as well as the 4-h clock fit and the 4-h clock prediction (based on the fit) using real data. The broadcast and MCC final data are plotted as such in blue line and red dots, respectively. Figure shows the piecewise-behavior of the 30-min broadcast model. The 4-h fit to the MCC data between 08-12 hours UTC is shown in solid red line and the prediction for the following 4 hours (12-16 hours UTC) in solid green. Figure shows that while the broadcast clock model is 5-10 meters off, the fit accuracy is <0.5 meters and the maximum prediction error is in the order of 3 meters in this case.

While the prediction-capability of the 4-h fit may not be considered to be sufficient for AGNSS-purposes, more accurate clock predictions can be made using longer fits and taking other factors into account. For instance, reference

[16] states that whereas GPS clocks can be predicted 12 hours ahead at 3-ns (1 m) accuracy, the 6-h prediction accuracy for GLONASS clocks is 10 ns (3 m). The method utilized is a long fit to a polynomial as well as addition of cyclic terms, for instance, to reflect satellite temperature changes in the orbit. Hence, as GLONASS clocks can be predicted at sufficient accuracy several hours ahead and the resulting clock model can be presented using GPS/Galileo clock model, giving AGNSS assistance for GLONASS in GPS/Galileo ephemerides is feasible.

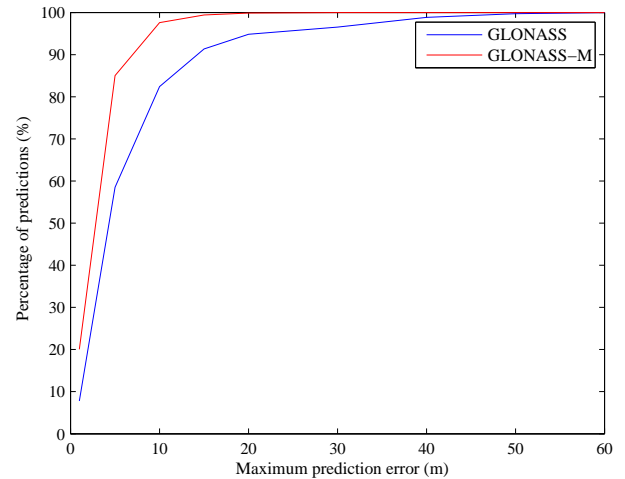


Figure 6: Cumulative 4-h prediction errors for GLONASS and GLONASS-M clocks. GLONASS and GLONASS-M data consist of 347 and 668 samples, respectively.

ABOUT TIME BASES

GPS, GLONASS and Galileo use different time bases to which, for instance, ephemerides are referred to.

$$(9) \quad \begin{aligned} |t_{GPS} - (UTC_{USNO} + \Delta_{Ls})| &\leq 1ms \\ |t_{GLONASS} - (UTC_{SU} + 10800s)| &\leq 1ms \end{aligned}$$

where UTC(USNO) is maintained by US Naval Observatory [9] and UTC(SU) is the Russian National Reference Time [13]. GPS time is continuous, whereas GLONASS time is discontinuous, whenever leap seconds are added to UTC. GLONASS and GPS time scales can be related either by solving the time difference in the GNSS-receiver or using parameter τ_{GPS} ($|\tau_{GPS}| < 30$ ns) included in GLONASS-M broadcast. τ_{GPS} is defined by $t_{GPS} - t_{GLONASS} = \Delta T + \tau_{GPS}$ with ΔT containing the full-second difference between UTC(USNO) and UTC(SU) (10800 seconds) as well as leap seconds between UTC(USNO) and GPS system time. Moreover, GPS and Galileo time scales can be related by the information included in Galileo broadcast [10].

The broadcast ephemerides are naturally referred to the GNSS system times. However, in the assistance standards the complexity may further be reduced, when using non-broadcast ephemeris assistance, by referring ephemerides to common time. The common time may be some arbitrary

time scale or a chosen GNSS system time. The assistance standards carry information on the GNSS-Common Time-relation. Having such a relation in the assistance also removes the potential need to solve the system time differences explicitly.

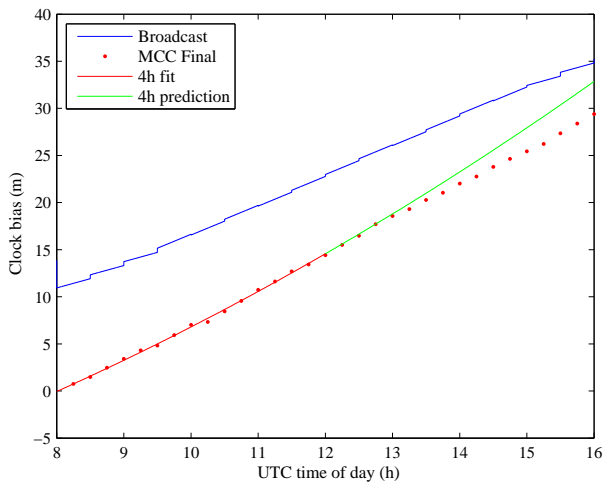


Figure 7: Broadcast, final and fitted/predicted clock model for GLONASS SV in the slot 23 (GLONASS number 714) for 17th Nov 2007. A bias of -20338 meters has been removed for clarity from all the data.

CONCLUSIONS

The introduction of new GNSSs has resulted in the modification of the 3GPP and OMA AGNSS standards. Amongst other, new orbit and clock models have been added into the standards in the form of multimode navigation model that is capable of carrying orbit and clock assistance to AGNSS-enabled mobile terminals for all the current and future GNSSs.

Although the benefits of introducing assistance for new GNSSs are evident from the user experience point-of-view, simultaneously complexity increases due to the need to support a plethora of information elements. In addition, for instance, unsynchronized ephemeris update-intervals in different GNSSs result in complex book-keeping in the network as well as in the terminal. Hence, it is advantageous to reduce the number of information elements and harmonize the orbit and clock parameterizations.

The current study has shown that assuming the availability of high-quality orbit and clock predictions for GLONASS satellites, the orbit and clock models may be given in the GPS/Galileo-style ephemeris. The ranges of the ephemeris parameters in GPS/Galileo ephemeris fixed by the bit counts and scale factors are sufficient for GLONASS orbits and clocks as well.

The study has also revealed that while some of the Keplerian parameters lie approximately in the same range in GPS and GLONASS, there are some major differences also. For instance, the harmonic correction coefficients lie in the same range in GPS and GLONASS. On the other hand, eccentricity and mean motion correction, which are coupled,

show significant differences between the GNSSs. This is due to GLONASS satellites having significantly more circular orbit than GPS satellites.

In conclusion, the performance of GPS and GLONASS can be harmonized with respect to orbit and clock models in a simple manner by parameterizing GLONASS orbits and clocks in GPS/Galileo-format.

References

- [1] Analyst report: Global Handsets, GPS/A-GPS Phone Sales. Strategy Analysts, November 2006.
- [2] 3GPP TS 44.031, Radio Resource LCS (Location Services) Protocol (RRLP), <http://www.3gpp.org>.
- [3] 3GPP TS 25.331, Radio Resource Control (RRC) protocol specification, <http://www.3gpp.org>.
- [4] OMA-TS-ULP-V1-0-20050719-C, User Plane Location Protocol, <http://www.openmobilealliance.org>.
- [5] The 3rd Generation Partnership Project, <http://www.3gpp.org>.
- [6] Meeting report: Report of TSG GERAN meeting#32, Sophia-Antipolis, France, 13th-17th November 2006, <http://www.3gpp.org>.
- [7] Meeting report of the 36th 3GPP TSG RAN meeting, Busan, Korea, 29th May - 4th June, 2007, <http://www.3gpp.org>.
- [8] Open Mobile Alliance Location Working Group meeting minutes OMA-LOC-2007-0290-MINUTES_20Aug2007Seoul, Seoul, Korea, 20th-24th August 2007, <http://www.openmobilealliance.org>.
- [9] IS-GPS-200 Revision C, IRN-200C-003: NAVSTAR GLOBAL POSITIONING SYSTEM Navstar GPS Space Segment/Navigation User Interface, 11th October 1999.
- [10] Galileo OS SIS ICD draft 0, 23rd May 2006.
- [11] IS-GPS-800 draft Navstar GPS Space Segment/User Segment L1C Interfaces, 19th April 2006.
- [12] IS-QZSS ver 1.0, Japan Aerospace Exploration Agency, 30th November 2007.
- [13] GLONASS interface document version 5.0. Moscow, 2002.
- [14] The Russian Mission Control Centre MCC, <http://www.mcc.rsa.ru/>.
- [15] Center for Orbit Determination in Europe CODE, http://www.aiub.unibe.ch/content/research/gnss/code___research/index_eng.html.
- [16] V. Broederbauer, A. Kostadinov, and R. Weber. Modelling of GPS and GLONASS satellite clocks. *Geophysical Research Abstracts*, 7, 2005.

- [17] J.M. Dow and R.E. Neilan. The International GPS Service (IGS): Celebrating the 10th Anniversary and Looking to the Next Decade. *Advanced in Space Research*, 36(3):320–326, 2005.
- [18] A.Q. Le. Achieving Decimetre Accuracy with Single Frequency Standalone GPS Positioning. In *Proceedings of the ION GNSS 17th International Technical Meeting of the Satellite Division, 21st-24th September, Long Beach, CA, USA*, pages 1881–1892, 2004.
- [19] D. Lundgren and F. Diggelen. Long-Term Orbit Technology for Cell Phones, PDAs. *GPS World*, pages 32–36, 2005. October issue.
- [20] J. Syrjärinne and L. Wirola. Setting a New Standard - Assisting GNSS Receivers That Use Wireless Networks. *InsideGNSS*, pages 26–31, 2006.
- [21] L. Wirola and J. Syrjärinne. Introduction to multi-mode navigation model. Technical Report GP-060179, Nokia Inc., 2006. Presented in 3GPP TSG GERAN2#29bis, 22nd-24th May, Sophia-Antipolis, France.
- [22] L. Wirola and J. Syrjärinne. Bringing All GNSS into Line. *GPS World*, 18(9):40–47, 2007.
- [23] L. Wirola and J. Syrjärinne. Bringing the GNSSs on the Same Line in the GNSS Assistance Standards. In *Proceedings of the 63rd ION Annual Meeting 2007, 23rd-25th April 2007, Boston, MA, USA*, pages 242–252, 2007.

Publication 7

L. Wirola and I. Halivaara and J. Syrjärinne. Requirements for the next generation standardized location technology protocol for location-based services. *Journal of Global Positioning Systems*, Vol. 7, No. 2, pp. 91–103, (2008)

Copyright ©2008 Journal Of Global Positioning Systems. Reprinted, with permission, from Journal Of Global Positioning Systems.

Requirements for the next generation standardized location technology protocol for location-based services

Lauri Wirola, Ismo Halivaara and Jari Syrjärinne
Nokia Inc., Finland

Abstract

The booming location-based services business requires more accuracy and availability from positioning technologies. While several proprietary location and positioning protocols have been developing in the market, scalable and cost-effective solutions can only be realized using standardized solutions.

Currently the positioning protocol standardization is concentrated in the 3GPP and 3GPP2 that define Control Plane (CP) positioning technologies for Radio Access Networks' native use. The limitations of the control plane in terms of architecture and bearer protocols are necessarily reflected in the CP positioning protocols and limit the feature sets offered. In addition to 3GPP/2 positioning technologies are also defined in WiMAX Forum and in IEEE for WLAN networks.

Location protocols in IP-networks, such as OMA SUPL (Open Mobile Alliance Secure User Plane Location protocol), encapsulate the CP positioning protocols. Thus the limitations of the CP protocols have also been copied to the User Plane, although the bearer there would be much more capable.

Due to the shortcomings in the CP positioning protocols, standardization activity for a new bearer-independent positioning protocol is proposed in order to fulfil the needs of the future location-based services. This paper discusses the current solutions, trends in the location technologies and outlines requirements for the future location technology protocol in terms of protocol features and data content.

The development of a generic positioning technology protocol is seen as an important development towards a convergence in the location protocols and the capability to provide location-based services irrespective of the bearer network. This has a major impact on the service development as well as user experience.

Key words: User Plane, Positioning protocol, Assisted GNSS, Fingerprint, Hybrid positioning

1. Introduction

Developing positioning and location standards has substantial market demand. Already now AGPS-enabled (Assisted GPS) mobile terminals constitute a significant share of the global navigation device market. The 2008 annual GPS-enabled smart phone sales are estimated above 30 million units and the analysts estimate that in 2011 the annual sales surpass 90 million units (Canalys, 2008). Moreover, modern smart phones are location-aware at least through the cellular network base station information. Finally, laptops can be made location-aware using WLAN-based positioning methods.

Positioning protocol standardization is concentrated in 3GPP (The Third Generation Partnership Project) and 3GPP2, which define positioning protocols for the Control Planes of GERAN (GSM EDGE Radio Access Network), UTRAN (UMTS Terrestrial RAN), E-UTRAN (Enhanced UTRAN) and CDMA (Code Division Multiple Access) networks. GERAN, where EDGE stands for Enhanced Data rate for Global Evolution, is better known as GSM (Global System for Mobile communications). UTRAN, where UMTS stands for Universal Mobile Telecommunications System, is commonly referred to as WCDMA (Wide-band CDMA). Finally, E-UTRAN is also known as LTE (Long-Term Evolution).

The Release 8 of GERAN standard will include the possibility to provide terminals with assistance data for all the existing and some future GNSSs (Global Navigation Satellite System). The assistance includes, among other things, the navigation model (orbit and clock parameters), reference location and reference time. In an assisted situation, the receiver does not need to download the navigation model from the satellites, but receives it over the cellular network to considerably reduce the time-

to-first-fix. Moreover, location and time data improve sensitivity significantly. The positioning is thus enabled in adverse signal conditions such as urban canyons. The improvement in user experience is significant compared to the performance of the autonomous GPS or simple cell-ID based positioning.

In addition to the RAN-independent AGNSS data each 3GPP location protocol also contains RAN-specific items. For instance, RRLP (3GPP-TS-44.301) (Radio Resource LCS Protocol, LCS LoCation Services) for GERAN networks and RRC (3GPP-TS-25.031) (Radio Resource Control protocol) for UTRAN networks include time difference and round trip time measurements, respectively, allowing for native RAN-based network positioning. Moreover, the reference time is given in a RAN-specific way by binding the cellular frame timing to the GNSS time.

Solutions for IP-networks include OMA (Open Mobile Alliance) SUPL (Secure User Plane Location protocol) Release 1 (OMA-TS-SUPL-1-0, 2007) and (draft) Release 2 (OMA-TS-SUPL-2-0, 2009) that encapsulate Control Plane positioning protocols defined by 3GPP/3GPP2 as sub-protocols to ULP (User plane Location Protocol). In addition to the capabilities of 3GPP and 3GPP2 positioning protocols the (draft) SUPL Release 2 adds items for, for instance, WLAN- (Wireless Local Area Network, IEEE 802.11) and WiMAX-based (Worldwide Interoperability for Microwave Access) positioning.

Currently the AGNSS-based positioning methods are essentially the only standardized positioning solutions available for global LBS (Location-Based Services). The native RAN-based methods are not widely deployed and their accuracy is varying. Moreover, the WLAN-based positioning capability in the (draft) OMA SUPL Release 2 is quite limited.

The future location services require more accuracy and availability from the standardized positioning solutions, because the use cases and user appetite for LBS will become more demanding. The requirement for increased availability can be understood by noting that people tend to stay indoors majority of the time, which is also the environment, where AGNSS performance is the worst. Accuracy requirement is, naturally, a question of application – location-sharing in a social network may require only cell-based positioning. On the other hand guiding a person to the correct entrance instead of just address requires higher accuracy than the current standardized solutions can offer.

However, when it comes to introducing new features into the standards, the problems lie in the currently utilized Control Plane positioning standards being bound to the

architecture and protocol limitations in the respective RANs. Moreover, because those protocols evolve from RAN needs (mainly emergency call positioning requirements), Control Plane positioning protocols will not add support, for example, for novel signal-of-opportunity -based positioning technologies. Therefore, a new standardized location protocol is required to introduce and implement new novel features and positioning technologies.

This article reviews the existing positioning protocols, discusses the future location needs and shows the limitations of the current solutions. Finally, based on the future needs and use cases, requirements are outlined for the new standardized location technology protocol that is flexible, scalable and comprehensive. In the long term the sought goal is the convergence towards a single generic User Plane location technology protocol.

It should be emphasized that in this article the term location technology protocol refers strictly to protocols associated with obtaining the position estimate of the user using different location technologies. The services including sharing location with third parties, security and privacy are out-of-scope of this article. The same also applies to the term location technology which refers to technologies related to obtaining the plain position information.

2. Radio Positioning Protocols in Different Radio Access Networks

3GPP TS 44.031

Radio Resource LCS Protocol (RRLP) 3GPP TS 44.031 (3GPP-TS-44.031) for the Control Plane of GERAN networks is defined in the 3GPP GERAN (General Radio Access Network) Working Group 2. RRLP is a stand-alone positioning protocol used in the communication between the Mobile Station (MS) and the SMLC (Serving Mobile Location Centre). RRLP carries information on the positioning methods, such as MS-assisted and MS-based modes, as well as assistance data.

RRLP also enables reporting Enhanced Observed Time Difference (EOTD) measurements as well as delivering information about the cell tower locations and real-time differences (RTD) between the base stations to the MS for MS-based EOTD that can be used as an alternative to or in combination with Assisted GNSS (AGNSS). EOTD is based on trilateration of the MS with respect to the base stations. However, EOTD requires relatively expensive infrastructure investments in the network (LMUs, Location Measurement Unit for measuring the RTDs) and, hence, its deployment has been very limited.

The release 98 of RRLP defined the support for Assisted GPS and EOTD. The Release 7 of the RRLP brought in

the support for A-Galileo and for multi-frequency measurements (including carrier-phase measurements), but also a generic structure for easy addition of other satellite systems. Finally, the Release 8 adds the support for GLONASS (Global Navigation Satellite System), QZSS (Quasi-Zenith Satellite System), Modernized GPS as well as various SBAS (Space-Based Augmentation System), such as WAAS (Wide-Area Augmentation Service) and EGNOS (European Geostationary Navigation Overlay Service).

Fig. 1 shows a simplified functional LCS architecture (3GPP-TS-43.059) in the GERAN network. The functional components in addition to MS, SMLC and LMU are BSC (Base Station Controller), BTS (Base Transceiver Station), CBC (Cell Broadcast Centre) and BSS (Base Station System). Note that the GERAN network consists of several BSS entities. The location requests originating from location clients are directed to the SMLC that handles the requests.

In the example of Fig. 1 the SMLC and CBC are integrated in BSC, although they can also be standalone components. From the positioning point of view the role of SMLC is to be the termination point of RRLP and CBC is responsible for broadcasting the assistance data to all the MSs within the cell (this is an alternative channel to distributing data over RRLP).

Moreover, an LMU can be, for instance, an entity with a GPS-receiver measuring the cellular time – GPS time relation for assistance data purposes. As mentioned, LMU is also needed for EOTD for measuring time relations between base stations. The LMU can also be a separate entity from BTS.

Although the GERAN LCS architecture is not the primary focus of this paper, the example given works to show that the Control Plane positioning protocols are strictly bound to the underlying architecture, which linkage is necessarily reflected also in the RRLP.

The RRLP includes two primary choices for the location of the position determination. In an MS-based mode the terminal autonomously determines its position taking advantage of the assistance that the terminal receives from the network. In contrast, in an MS-assisted mode the terminal typically receives only a measurement request and minimal assistance for fast signal acquisition, such as code phase search window in the AGNSS case, from the network and reports measurements to the SMLC, which determines the position. The measurements may either include GNSS measurements or EOTD measurements or alternatively both types for a hybrid solution. One to three sets of measurements can be requested and delivered to the SMLC.

In addition to MS-based and MS-assisted modes the range of methods in RRLP also includes MS-assisted preferred, MS-assisted allowed, MS-based preferred and MS-based allowed.

Apart from RRLP, it should be noted that due to the limited adoption of EOTD, emergency services primarily utilize UTDOA (3GPP-TS-43.059) (Uplink Time Difference Of Arrival) in GERAN networks. Moreover, the actual assistance data requests are not delivered in RRLP, but in the BSS Application Part LCS Extension protocol (3GPP-TS-49.031).

3GPP TS 25.331

Radio Resource Control (RRC) 3GPP TS 25.331 (3GPP-TS-25.331) is the radio resource control protocol for the User Equipment (UE) - UTRAN interface. RRC defines, amongst other items, similar functionality for positioning of an UE in an UTRAN network as RRLP does for positioning of an MS in a GERAN network. It should, however, be noted that whereas RRLP is a standalone positioning protocol with termination points at MS and SMLC, RRC carries in addition to positioning payload also a plethora of other data. Hence, RRC is not only a standalone positioning protocol. RRC is terminated at UE and RNC (Radio Network Controller) of the UTRAN. RRLP and RRC therefore differ in scope and implicated architecture, although both can carry the same type of positioning and location information.

In addition to AGNSS-based positioning, RRC also provides a RAN-based trilateration method called IPDL-OTDOA (Idle Period DownLink - Observed Time Difference Of Arrival). Similarly to EOTD, IPDL-OTDOA requires infrastructure investments and, hence, the deployment has been limited.

3GPP2 C.S0022-A

C.S0022-A (or IS-801-A) (3GPP2-C.S0022-A) defines a position determination protocol for IS-95/IS-2000 and HRPD (High Rate Packet Data) systems and is maintained by 3GPP2. The capabilities of C.S0022-A are similar to its 3GPP counterparts. The support for additional satellite systems will be included in the coming release C.S0022-B.

The IS-95/IS-2000 networks also support TOA-based (Time Of Arrival) positioning method called Advanced Forward Link Trilateration (AFLT). AFLT is based on the time synchronized base stations that allow the network or the terminal to calculate the position estimate of the terminal based on the TOA measurements.

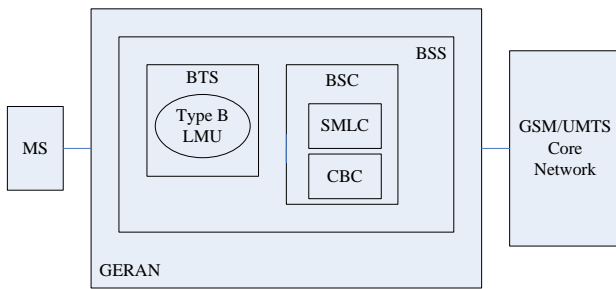


Fig. 1 GERAN LCS architecture.

OMA SUPL 1.0 and 2.0

The previously considered RRLP, RRC and TIA-801 are positioning protocols for Control Plane of the cellular networks - they are an integral part of the cellular network. However, in addition to Control Plane solutions, there are also User Plane solutions, which provide assistance and positioning data over IP-networks.

Examples of User Plane solutions are the SUPL (Secure User Plane Location protocol) Release 1 (OMA-TS-SUPL-1-0, 2007) and (draft) Release 2 (OMA-TS-SUPL-2-0, 2009) standardized in OMA (Open Mobile Alliance). SUPL architecture provides a wide-range of services, such as authentication, security and charging, through other enablers (defined by OMA, 3GPP or other standardization fora) as well as various location services including triggered periodic and area events. Therefore, the OMA LCS architecture with SUPL can be considered to be a complete end-to-end solution as required of OMA enablers.

In positioning technologies OMA SUPL relies on Control Plane protocols, such as RRLP and RRC, which the SUPL encapsulates as sub-protocols (see Fig. 2). Over the recent years the importance of SUPL has increased due to the growth in the LBS business. Increasingly the primary use for the Control Plane methods, such as A-GPS as well as AFLT in the CDMA networks and U-TDOA in GERAN networks, is in emergency services, whereas LBSs are based on User Plane positioning solutions.

Fig. 3 introduces a simplified OMA Location Architecture - for full architecture see (OMA-AD-SUPL-2-0, 2008). The architecture shows the major entities including SUPL Location Platform (SLP), Short Messaging Service Centre (SMSC), WAP PPG (Wireless Application Protocol Push Proxy Gateway) and SET (SUPL-Enabled Terminal), which is the terminal to be positioned. The functional entities of SLP, SUPL Location Centre (SLC) and SUPL Positioning Centre (SPC), handle amongst other items subscription, authentication, security, charging, privacy, positioning and assistance data delivery.

In the SUPL framework the positioning session can either be SET-initiated or Network-Initiated. In the SET-initiated case the SMSC and WAP PPG do not have a role, but the SET directly connects to the SLP (in proxy mode - the behaviour in the non-proxy mode in CDMA networks is somewhat different) and, for example, retrieves the required assistance data from the SLP. In the Network-Initiated case the SET must be notified so that it knows to set up data connection to the SLP. The channels to deliver the notification are, for example, over an SMS (text message) or over WAP. In an exemplary case of the network-initiated session a SUPL Agent external to the SET (an application, for instance) requests SLP to position the SET. Having received the request the SLP sets up a Network-Initiated session with the SET using an SMS and positions the SET.

The advantages of OMA SUPL lie in the possibility to rely on other OMA enablers and also on other standardized architectures including SMS. The Network-Initiated sessions can, for example, be utilized in various services as well as in lawful interception and positioning of emergency calls.

LTE and WiMAX considerations

The emerging RANs, 3GPP E-UTRAN and WiMAX (based on IEEE 802.16), also need positioning solutions. The 3GPP LTE has a work item open for an LTE-native Control Plane solution that will incorporate AGNSS as well as time difference-based methods. WiMAX Forum has agreed to use SUPL as one option for positioning. Moreover, in WiMAX Forum there is also a work item towards a WiMAX-native positioning solution called WLP (Wireless Location Protocol).

It should be noted that the (draft) SUPL Release 2 supports both LTE and WiMAX and, hence, the use of SUPL might be an adequate solution in these networks. The draft release also supports LTE-native positioning protocol, even though it has not been defined yet. However, its inclusion can be justified by respective timelines of future SUPL releases and LTE Release 9.

RRLP	RRC	LTE RRC	TIA-801
ULP			
TLS			
TCP/IP			

Fig. 2 OMA SUPL Release 2 protocol stack.

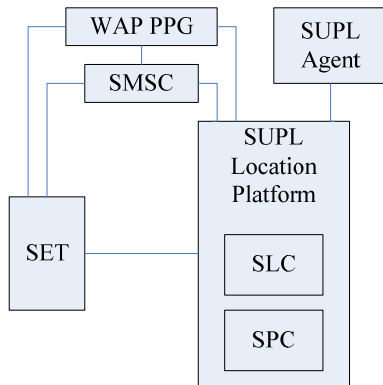


Fig. 3 OMA Location Architecture

3. Shortcomings in the Existing Positioning Protocols

The Control Plane positioning protocols have been developed and evolve based on RANs' needs (mainly emergency call positioning) and capabilities. The positioning protocols defined in 3GPP and 3GPP2 contain RAN-specific items that are not needed outside the scope of the respective RAN. These include, for instance, the native RAN methods including EOTD and AFLT. Hence the use of RAN-specific positioning protocols complicates the User Plane deployments. This is especially true of RRC, which is the protocol for the radio resource control in general.

Moreover, the Control Plane protocols suffer from the limitations of the protocols lower in the hierarchy in the RAN protocol stack and prevent realization of novel features. For instance, in the Control Plane it has been impossible to realize solutions for high-accuracy sub-meter positioning methods, such as Real-Time Kinematic (Leick, 2004) because of bandwidth, architecture and protocol limitations. To be more specific, for instance RRLP is designed for one-time point-to-point assistance and measurement delivery and is, therefore, unsuitable for positioning methods requiring constant stream of reference measurements (Wirola et al., 2007b and 2008b).

Because the Control Plane protocols are also being utilized in the IP-networks via the use of OMA SUPL, the in-built protocol limitations have also been copied to the User Plane solutions. This leads to sub-optimal solution, because in the User Plane the bearer-networks and -protocols would in fact be capable of providing more services and significantly larger bandwidth for positioning purposes.

Considering the current and future needs one of the most serious flaws in the 3GPP-based protocols is the lack of support for the signal-of-opportunity -based, such as WLAN, positioning. The Control Plane protocols do

include the support for the RAN-native network measurements, but they lack the capability to transfer measurements made from other RANs or radio networks. For instance, because IEEE networks are out-of-scope of 3GPP, it is highly unlikely that 3GPP would define positioning methods that are based on IEEE technologies including WLAN. The same also applies vice versa.

In the (draft) SUPL Release 2 this deficiency has been overcome to some extent by incorporating items for signal-of-opportunity positioning in the ULP-layer (User plane Location Protocol) of SUPL shown in Fig. 2. These capabilities include the possibility to report radio network measurements from various networks, including GSM, WCDMA, LTE and WLAN in the ULP-layer messages.

The ULP-layer defines messaging, for example, for initiating and terminating the SUPL session as well as capabilities handshakes and service subscriptions. However, because the AGNSS and the other RAN-based positioning methods (EOTD, IPDL-OTDOA) are encapsulated in the sub-protocols to the ULP, the layer-approach typically adopted in the protocol design is dismantled due to the positioning technology additions made to the ULP layer. Therefore, the structure and capabilities of SUPL have suffered significantly from inheriting the limited Control Plane positioning protocols. Hence, also OMA SUPL would benefit from developing a flexible and comprehensive positioning technology protocol solely for the User Plane.

Another challenge in the user plane is the GNSS fine time assistance. The more accurate the time assistance is the more precisely the AGNSS receiver can predict the Doppler and code phase in order to improve sensitivity and, hence, the time to first fix. In the Control Plane the GNSS time assistance is tied to the cellular frame timing. However, as mentioned, this requires the deployment of LMUs in the network. The same capability is also available in the User Plane. However, this approach has a drawback that it makes SUPL inherently operator-tied service, because access to the core network is needed in order to obtain the GNSS-cellular time relations.

In the User Plane alternative solutions to be considered include Network Time Protocol (NTP) and timing services available in the Internet. However, the latency of the IP-network, especially over the air, results in unpredictable errors in the time assistance. Another future option is to obtain the cellular time -tied time assistance from another terminal in the same cell over a peer-to-peer network or via a server caching the cellular timing data collected from terminals. However, in any case these latter methods would not provide a solution for, say, WLAN-only devices. Therefore, time assistance over the IP-network remains a challenge.

Finally, one of the drawbacks in performing positioning in the User Plane using OMA SUPL is that WLAN-only devices cannot utilize all the SUPL services unless the terminal and the WLAN network are I-WLAN -enabled (3GPP Interworking WLAN). For example, authentication in OMA SUPL requires having a SIM card (Subscriber Identification Module) in the terminal and a subscription to the 3GPP network. I-WLAN provides a mechanism to support 3GPP specified mechanisms, including authentication, over the WLAN bearer.

4. Trends in location technologies

Positioning services can be characterized by four attributes: availability, accuracy, integrity and authenticity of the source. Availability refers to the fraction of time, when positioning is possible. For example, GNSS-based positioning has excellent availability in rural outdoor conditions. However, in urban and indoor environment the availability degrades rapidly.

Accuracy, on the other hand, refers to how precise location information a given positioning technology may yield. Typically GNSS is considered an accurate technology, whereas cell-based methods are referred to as inaccurate technologies with a potential position error of several kilometres.

Integrity refers to the reliability of the positioning service. For instance, in GNSS-based positioning integrity may be compromised by a faulty satellite. Because of this satellites send their health (or integrity) data to the user equipments. The integrity information is also provided in the AGNSS assistance.

Finally, authenticity refers to the authenticity of the signal source. Typical examples for signal authentication include the methods to prevent the spoofing of GNSS signals. Spoofing can be understood to mean misguiding of users by means of forged signals (Günter, 2007). While military users have always been concerned with the potential spoofing and jamming of the signals, these aspects are also of growing importance in the civilian sector now that, for instance, location-based security solutions are being introduced. Moreover, in addition to deliberate forging attempts, unintentional interference from in-device or from other devices are potential sources of errors.

While integrity and authenticity are major concerns in the emergency services, they are not currently considered as major drivers in developing positioning technologies for location-based services. This is due to the inherent problems with availability and accuracy in consumer solutions, such as positioning services in mobile terminals. These issues must be solved first. However, as

technologies develop in these areas, solutions in the areas of integrity and authenticity will be required as well. For instance, applications requiring or providing location-based charging necessitate integrity and authenticity guarantees. One option to tackle both spoofing and interference is to have at least two independent positioning technologies enabled in the device.

Therefore, the two near-future driving factors in the location technologies are accuracy and availability. Accuracy requirements can be addressed by enabling more advanced GNSS-based positioning methods and AGNSS assistance to the consumers. From technology point-of-view it would be possible to provide the end users with high-accuracy GNSS positioning methods, such as Real-Time Kinematics (RTK) (Wirola et al., 2006) and Precise Point Positioning (PPP) (Leick, 2004). However, these methods both require new protocol messaging as well as new types of assistance data services, such as high-accuracy navigation models as well as regional atmosphere models in the case of PPP. These methods cannot be realized in the Control Plane protocols and extending SUPL to sew up the sub-protocol (RRLP or RRC) shortcomings has been shown unfeasible (Wirola, 2008b).

Although a full-scale multi-frequency RTK may be an overkill for a handset integrated GNSS, it is feasible to realize at least a light-version of RTK using an external GNSS-receiver connected via Bluetooth to the device (Wirola et al. 2006 and 2008a). The increasing availability of satellite systems and civilian signals in consumer-grade GNSS devices will eventually enable the technologies now in professional use also to the wider audience. By the light-RTK the authors refer to abandoning rigorous integrity requirements of professional RTK solutions to some extent and also on being satisfied with a float solution.

Moreover, the availability of GNSS reference networks and, hence, the availability of virtual reference measurement services introduce interesting opportunities for future high-accuracy positioning technologies for consumers. However, in order to realize this potential the standardized positioning solutions must be able to carry appropriate data content, which they are not capable of doing at the moment.

The same also applies to PPP. A rigorous professional-quality PPP may not be feasible for consumer devices, but significant performance improvements can already be achieved by enabling high-accuracy navigation models and, say, regional troposphere and ionosphere models. Again, such a PPP solution might be called light-PPP.

The discussed high-accuracy AGNSS methods, however, have low availability due to the requirement to have good

or excellent satellite signal conditions. The availability aspect, on the other hand, can be addressed by the radio network -based methods based on fingerprinting, fingerprint databases and associated positioning technologies.

A fingerprint database is defined as a grid, in which each grid point is associated with a set of measurements from a set of radio networks (Honkavirta, 2008). The measurement types include time delay measurements, time difference measurements between the base stations, channel or signal quality measurements (power histograms, number and spread of RAKE fingers, pulse shapes) and measurements from multiple antennas (diversity receiver). The databases may have wide or even global coverage.

An important aspect in signal-of-opportunity -based positioning is that it must be based on existing infrastructure. Limited areas, such as hospitals, can be populated with special positioning tags or similar, but a global scale positioning solution must take advantage of already existing wide-spread infrastructure. This can be seen as one of the drivers for WLAN-based positioning. The WLAN infrastructure is widely available and various devices are already equipped with a WLAN chip. Hence its utilization in positioning is a natural step. The only remaining aspect is the availability and transfer of WLAN access point maps, which transfer is currently in the scope of no positioning standards.

The IEEE 802.11 has activity towards standardizing an interface that allows the access point to report its position to the terminal or vice versa (IEEE, 2008). However, it takes time to replace the existing WLAN infrastructure with new equipment supporting new standards. And even then, not all the access points may have their coordinates set for further distribution. Hence, the current WLAN-based positioning solutions rely on databases with records of access points versus their coordinates in the simplest form of databases.

As discussed, the fingerprint positioning solutions almost completely lack support in the location standards. Although the (draft) SUPL Release 2 supports reporting GSM, WCDMA, LTE, CDMA, HRPD, UMB (Ultra Mobile Broadband), WLAN and WiMAX network information, SUPL is not designed for the fingerprint collection. Moreover, OMA SUPL is based on an assumption of network-based SET-assisted RAN-based positioning and, hence, it is impossible to transfer a fingerprint database to the terminal for positioning purposes using the current SUPL versions. Again benefits for SUPL can be seen in defining a new positioning technology package for the User Plane LBS needs.

Moreover, the support for sensor-generated measurements and information originating from e.g. accelerometers, magnetometers and barometers is not covered by the current standards. For instance, although heading information is supported in various standards, motion state (walking, running, etc.), which can be extracted from the accelerometer data, is not. Moreover, taking advantage of the full potential of barometers requires availability of either pressure reference data or troposphere models. Sensors are also expected to play a major role in addressing the indoor positioning challenge (Alanen et al., 2005).

Due to the limitations in the currently utilized standardized Control Plane/User plane solutions several User Plane -oriented proprietary systems have been developing in the market. Examples include WLAN-based positioning solutions as well as proprietary GNSS assistance data services. These are differentiators in the market and all the techniques can never be standardized due to the intellectual property right and business secret issues. Although the standards cannot provide unified interface towards these services, the standards could still, however, provide generic containers for proprietary payloads so that both, standardized and proprietary assistance, could be carrier within the same standardized framework. The advantage of such an approach is that each new assistance service would not then have to define a new protocol.

The introduction of proprietary containers would also work to prevent the fragmentation of the positioning protocols. Currently each new RAN is forced to define a new positioning protocol for its native use. In addition, each new assistance service is compelled to define a proprietary protocol for carrying the data. The negative effects of such fragmentation include increased costs due to the need to support multiple protocols.

Yet another driver for the location technologies is the location-awareness and power consumption, which are closely related. Being location-aware requires performing positioning periodically or based on some other criteria such as change of an area. However, such frequent positioning events lead to increased power consumption and also to data costs. Hence, the location technologies being developed generally try to minimize the data connections – an example of such are predicted ephemeris services. Moreover, such power consumption requirements also lead to positioning being performed by the technology that just and just fulfils the required quality-of-service. For example, if only crude position estimate is required, GNSS shall not be used. Instead, the terminal is always aware of its serving cell and, hence, assuming an availability of an appropriate fingerprint database, the terminal can be positioned without significant additional energy consumption.

Also, the applications utilizing location data, such as Nokia Maps, operate on the User Plane and are becoming more interactive. Therefore, it is natural that the development of the location technology protocols is concentrated in User Plane, not in the Control Plane.

In conclusion, the discussion above shows that there is a need for standardization activity in location technology protocols in the User Plane. Authors have been proposing a work item for the LTP (Location Technology Protocol) in the OMA Location working group, but so far such a work item has not been approved.

5. Use Cases

In the current location solutions (for example see Figs. 1 and 3) there is a strict architecture with the location server (for example SMLC in GERAN and SLP in SUPL) providing the terminals (MS or SET, respectively) with assistance and positioning instructions. The network element has been given the control of the positioning session - it is the network element that decides, or recommends, which positioning method shall be used. These both issues must change since they imply heavy architecture and network-controlled positioning session, respectively.

Firstly, the LTP must not limit the roles of different entities – instead, the LTP can find its use between various different types of entities. Any entity (for instance, handset, laptop, server and service/data provider) can work in any role. For example, traditionally there has been a server providing terminals assistance data – however, it is equally feasible for the server to request assistance data from terminals for distribution to other terminals.

Similarly, the LTP messaging must also flow between any types of entities. For instance, in device-to-device relative positioning measurement messages are exchanged between two devices, for example two terminals, not between a terminal and server. This implies that the LTP must not be tied to any specific architecture, because any entity can request and deliver almost any data.

Such a concept is shown in Fig. 4, in which different entities are represented as nodes that are termination points of the LTP. Any node should, in principle, be allowed to work as a data producer (i.e. allowed to publish, for example, the satellite ephemerides the node has received) in the location network. Also, any node should be able to function as a data provider (i.e. work, for instance, as a cache server for assistance data) in the network. Such a scheme opens up a possibility to set up community-based assistance networks.

Note that the complexity with, for instance, security, charging, privacy, setting up the point-to-point or even point-to-multipoint as well as multipoint-to-point connections is hidden in the bearer protocol. Such aspects are not in the scope of the location technology protocol, but are taken care by the bearer protocol encapsulating the LTP. The bearer protocol is indicated in Fig. 4 by the notation B(LTP), which refers to the LTP being encapsulated by a bearer (B) protocol.

The retrieval of assistance data from an external source to the location server for distribution has thus far also been out-of-scope of location technology protocols. However, the data from the Wide-Area Reference Networks (WARN) is essentially similar to the data provided by the location server to the terminals. Therefore, the third use case to consider for the LTP is in this interface. The requirement can again be achieved by considering the WARN feed provider as a node (see Fig. 4) with certain capabilities. Cost savings can be induced by standardizing also the channel between the data provider and the assistance server.

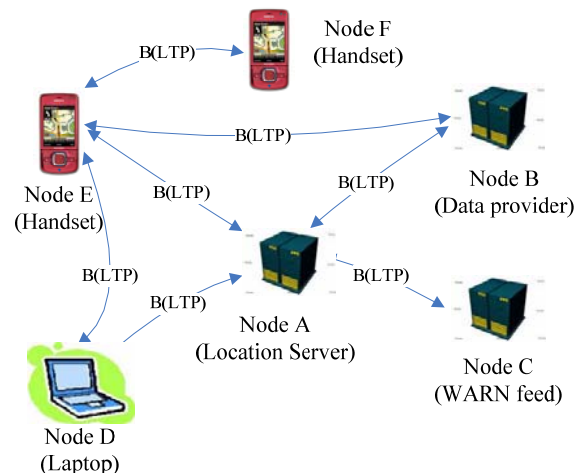


Fig. 4 The LTP is designed to not to limit the flow of information between the nodes in the location network. However, the bearer protocol may limit the actual connections. The notation B(LTP) is introduced to highlight that LTP needs to be encapsulated by a lower level bearer protocol.

The fourth use case to consider is the event-based assistance data. Currently, the protocols are designed so that the terminal either requests assistance data or the server pushes assistance data to the terminal in the beginning of the positioning session. However, there are emerging assistance data types including atmosphere models and long-term GNSS navigation models, which require that a serving node must be capable of pushing updated assistance data to the terminal as the data changes. Also, the nodes must be able to subscribe this data.

Finally, the fifth use case for the data content in the LTP is broadcasting. Majority of, for instance, GNSS assistance data is global by nature. Therefore, server loads and bandwidth requirements can be eased, if GNSS assistance were broadcasted, for example, over OMA BCAST (OMA-TS-BCAST, 2008). Also, some data may be regional. For example, Europe-wide ionosphere maps could be broadcasted using OMA BCAST, because the enabler also provides means to control the distribution geographically.

Also RTK measurements from GNSS reference networks are suitable for broadcasting. In such a case measurements and locations of the reference stations are distributed so that the terminal can process all the measurements (from the network and the terminal) in a single filter. The approach has been shown to produce superior results (Dao, 2005). Another option is to distribute reference measurements and spatial correction terms to the terminals. Finally, also updates to the geographically-segmented fingerprint databases could be delivered over the OMA BCAST channel.

6. Data content requirements

In the Assisted GNSS -side, the LTP must offer the same types of AGNSS assistance as today - for thorough discussion about AGNSS assistance refer to (Syrjärinne et al., 2006). This includes being able to provide data common to all the GNSSs (such as, ionosphere model) and GNSS-specific data (such as navigation models) in a generic format. Also, an important aspect is having a multi-mode navigation model enabling providing GNSSs navigation models also in non-native formats (Wirola, 2007a). The present protocols also support differential GNSS, data bit assistance, earth-orientation parameters and real-time integrity. All these must be supported by any subsequent protocols. In general, the current positioning standards for AGNSS support effectively all the content available in the GNSS broadcasts.

In addition to the broadcast data types, the data must also support reference location and time. Reference location may be given based on radio network -data. The reference time must be defined in such a way that it can be given with respect to any given radio system. Currently the RAN-specific positioning standards support only giving GNSS time with respect to the specific RAN time. However, the frame timings that are typically given in RAN-specific units for reference time purposes can be reduced into common units including SI-units.

Also, the future protocol must have suitable content to support novel high-accuracy GNSS positioning methods. This means having certain measurement types, namely code phases and carrier phases at suitable resolution, in the standard (Wirola et al. 2007b) in order to be able to

support RTK. For PPP the new required data content includes high accuracy navigation models, differential code biases and regional atmosphere (ionosphere, troposphere) models. Additionally also, for instance, antenna information may be considered.

Finally, the AGNSS side must also consider the emerging predicted navigation model services and their derivatives. The 3GPP specifications already include one implementation of predicted navigation models that can provide the terminal navigation model data for several days ahead. However, there are also other implementations and also data transfer needs for proprietary services including autonomous predicted ephemeris generation in the terminal.

In the radio network -based positioning the assistance to be carried by the LTP consists of fingerprint database. These items are not consistently included in any location standard. Although the (draft) SUPL Release 2 specification defines measurement parameters for a number of networks, the parameters are not equal between the systems (contents of data elements, resolutions, ranges). The hybrid use of different networks is, therefore, very difficult due to profound differences in the measured parameters and the measurement report contents in the User Plane specifications.

The generic fingerprint to be included in the LTP must therefore equalize the systems by providing, for example, such generic timing (or time difference) and observed signal strength measurement report that it is applicable to all the systems. Only then are the real hybrid methods feasible. This addresses especially the availability challenge. The systems considered may include GSM, WCDMA, WLAN, WiMAX, Near-Field Communications, Bluetooth and DVB-H (Wirola, 2008c).

The positioning using the fingerprint database is based on statistically comparing the measured fingerprint to the database records (Honkavirta, 2008). Another type of data, based on fingerprint database however, suitable for positioning are radiomaps (Wirola, 2009) that contain access point and/or base station coverage area models in terms of shapes defined in 3GPP GAD (3GPP-TS-23.032) including ellipses, ellipsoids or polygons.

Finally, the support for sensor-generated measurements and information originating from e.g. accelerometers, magnetometers and barometers must be covered. For example, supporting barometer fingerprints could allow the server to keep its pressure assistance grid up-to-date for assistance data purposes instead of relying on weather forecasts or similar.

Finally, the possible broadcast of the data elements over the OMA BCAST introduces no additional requirements for data coding. The data content carried within the OMA BCAST can be anything, for instance a file. However, the broadcasting possibility should be borne in mind, when defining the data content so that, for instance, geographical applicability aspects are adequately taken into account.

7. Protocol stack requirements

Fig. 5 shows the schematic protocol stack used with the Location Technology Protocol, which is the highest protocol layer. It handles all the positioning-related messaging and data transfer.

In addition to the LTP, a lower level protocol is required to handle transporting the LTP payload from one node to another simultaneously handling, for example, user authentication, security, privacy and charging issues, if required. This protocol encapsulating the LTP is called the Routing Protocol in Fig. 5. This can be thought to be the bearer protocol indicated in Fig. 5.

The protocol requirements to the LTP itself include that it must be capable of error handling and recovery – a typical situation with AGNSS assistance is that the entity providing assistance data cannot provide all the data the terminal requested. The protocol must therefore not expect to get all the requested data, but be capable of a re-requesting other assistance data, if applicable. For this purpose the termination points must also be able to exchange their capabilities, namely to report what their positioning method and assistance data capabilities are.

Furthermore, the LTP messaging shall be symmetric so that the LTP does not imply that it is always the terminal that requests assistance data – equally well a server may request assistance data from some entity. Symmetric messaging, therefore, enables abandoning the current scheme of strict division between the MS and Location Server.

While version control is a natural requirement of any protocol, the LTP is also envisioned to be stateless in order to maintain the scalability of the infrastructure that includes, for example, multiple servers. However, depending upon the services provided the Routing Protocol may need to have states if the deployment supports, for instance, sessions for continuous periodic exchange of measurements (streaming). This is required, for example, in RTK that requires a possibility to request and deliver a stream of measurements from one node to the other. The streaming is then realized in the Routing level and the exchange of measurements in the LTP layer.

The lack of states also means abandoning the conventional methods such network-based MS-assisted mode, in which the network orders the MS to take measurements and return them to the network for position determination. Giving up such schemes is natural, because the terminal capabilities have increased and terminals are nowadays fully capable of performing, for instance, all the calculations required for position determination. Hence, the role of the network (server) side should be more supporting than imperious.

The Routing Protocol may either be a very simple or arbitrarily complex one depending upon the services it must provide. However, the possibility to have a simple Routing Protocol, which in its simplest form only need to open a data pipe between two nodes, serves research and development work as well as academics. In certain deployments this also yields cost advantage. The realization of the Routing Protocol in each deployment ultimately depends upon the environment.

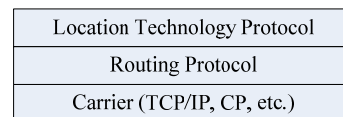


Fig. 5 Schematic protocol stack

Note that the definition of this bearer protocol is tied to the architecture. In an exemplary case in Fig. 5 the Node A may work as a master node (location server) to which all the other nodes register with their capabilities. A node might for instance register with a capability that it can provide the other nodes broadcast A-GLONASS assistance limited to the satellites visible to the node. Therefore, in addition to registration, the bearer protocol must, in this case, provide the means to route, say, assistance data requests originating from one node to another node capable of providing the assistance data.

Now, the architecture defined in this example sets requirements to the bearer protocol, but the underlying LTP is unchanged. The adaptation is, hence, in the Routing level and is transparent to the LTP. This is essential from the protocol transferability point-of-view. It is also the Routing protocol that limits the roles of the entities based authentication, security privacy and charging requirements.

Although the Routing Protocol is out-of-scope of the LTP and this article, it should be recognized that different features provided by the LTP require different levels of service from the Routing Protocol. It is therefore advisable to categorize the LTP features into service packs. The service pack definition consists of the subset of the LTP features and of the requirements their implementation sets for the Routing Protocol.

8. Exemplary implementation

In the 3GPP systems the Routing Protocol already exists to some extent, because for example authentication is inherent to the architecture.

In the IP-networks the Routing Protocol could be the ULP, because it can rely on, for instance, security (based on TLS, Transport Layer Security), authentication (based on 3GPP GBA (3GPP-TS-33.220), Generic Bootstrap Architecture) and charging mechanisms already defined in 3GPP, OMA and other fora. However, it should be noted that the (draft) SUPL Release 2 cannot support the LTP, but the future releases of SUPL could consider the LTP, if standardized, as a location technology enhancement exerting certain requirements on the ULP-layer as well as on the OMA LCS architecture.

Fig. 6 shows the realization of the LTP in the OMA SUPL architecture. The LTP has been introduced alongside the current 3GPP/2 positioning protocols as a sub-protocol to the ULP. Within the LTP there are modules for different positioning technologies including GNSS and Radio Network –based positioning. The LTP itself contains the capabilities handshake and positioning requests for different positioning technologies or their hybrids.

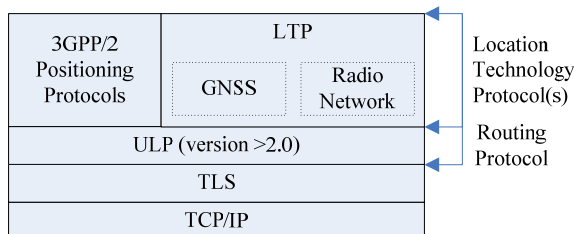


Fig. 6 Exemplary stack implementation

The GNSS module in the LTP is the protocol for Assisted GNSS data. It includes the content, request and delivery mechanisms found in the 3GPP AGNSS specifications. Also included are, for instance, capabilities to request and deliver regional atmosphere models, measurements required for high-accuracy methods and the multitude of non-native navigation models. Moreover, the historical division to MS-based and MS-assisted methods is not required, because similar functions can be realized through simply requesting and providing assistance data and measurements.

The Radio Network –module is, on the other hand, a protocol for transferring generic fingerprints and radiomap data. For example, through this module it is possible to transfer a WLAN access point coverage area map.

The individual modules can be coded in the wanted formats – they can follow the same coding or have different codings. Exemplary codings include XML and ASN.1. XML has the advantage of being flexible and robust as well as easy to debug. The drawbacks include high bandwidth consumption. The situation can be improved by binary XML such as Efficient XML (W3C, 2008). Even though the binary XML typically achieves good compression ratio, ASN.1 with PER (Packet Encoding Rules) encoding is still superior bit-consumption-wise especially, when the unaligned version is used. On the other hand, aligned PER is more efficient to decode, but consumes more bandwidth. However, extending ASN.1 in future releases results in the code being challenging to follow.

The actual choice of the encoding depends on the anticipated environment as well as future needs. For example, it could be argued that the GNSS-based methods and required assistance data elements are well-established and known and, hence, no major future changes are expected to take place. Hence, ASN.1 is the choice for the GNSS package. On the other hand, the Radio Network –package with new and novel fingerprint databases requires flexibility and expandability. Hence, XML might be the choice for that package.

In order to support, for instance, the delivery of basic GNSS assistance data (navigation models etc.) the OMA LCS architecture and the ULP layer need not be modified to a large extent. In principle the only modification required is adding the indication of the support for the LTP to the ULP-layer. The LTP can then be carried in the same container as the 3GPP/2 positioning protocols.

Bigger changes are, however, required, for instance, for streaming of GNSS measurements between two (or more) users. This requires changes to the OMA LCS architecture and additional messaging so that one user can request such a data pipe to be opened, a network-initiated method to request such measurements from the other user as well as an architecture enabling such routing of measurements for a pre-defined time.

9. Conclusions

Several shortcomings in the currently utilized positioning protocols have been identified in the view of the future location and positioning needs. The current approach of each standardization forum working with location technology protocols for their own domain leads to continued fragmentation of location technology standards and to domain-specific implementations. These domain-specific standards differ in scope and capabilities depending upon the bearer network capabilities as well as the development cycles. Therefore, harmonized positioning performance cannot be guaranteed across all

the networks and access network handovers. Moreover, due to the long development cycles of standards, various proprietary location technology protocols have been developing in the market leading to further fragmentation.

Instead, the domain-specific items must be addressed in a lower level adaptation protocol, which is transparent to the location technology protocol. For the location-based services it is important that the location experience is independent of the access network. Such a location technology protocol free of domain-specific hooks can address the needs of every domain (IP, RAN) and lead to convergence in location standards by being re-usable in every domain.

The location technology protocol itself must address in their entirety positioning procedures, messages, measurements and assistance for GNSS-, sensor- as well as radio network -based positioning methods. The protocol must also be as flexible and comprehensive as possible so that additions can be made in fast schedule, when needed. Also, placeholders for proprietary extensions reduce the need for the proprietary protocols in the market.

The authors see that there is a market demand for a comprehensive standardized location technology protocol for User Plane needs. In the long term the standardized solutions are the most cost effective approaches and lead to the widest adoption of the technologies.

Acknowledgements

The authors wish to thank Tommi Laine (Nokia Inc.) and Kimmo Alanen (Nokia Inc.) for providing insightful comments.

References

- 3GPP-TS-23.032 *Universal Geographical Area Description (GAD)*
- 3GPP-TS-25.331 *Radio Resource Control (RRC) protocol specification*
- 3GPP-TS-33.220 *Generic Authentication Architecture (GAA); Generic Bootstrapping Architecture*
- 3GPP-TS-44.031 *Radio Resource LCS (Location services) Protocol (RRLP)*
- 3GPP-TS-43.059 *Functional Stage 2 Description of Location Services (LCS) in GERAN*

- 3GPP-TS-49.031 *Location Services (LCS); Base Station System Application Part LCS Extension (BSSAP-LE)*

- 3GPP2-C.S0022-A *Position determination service for CDMA2000 Spread Spectrum Systems*

- Alanen, K. and Käppi, J. (2005) *Enhanced assisted barometric altimeter AGPS hybrid using the Internet*. In Proceedings of ION GNSS 2005, 13th-16th September, Long Beach, USA.

- Canalys (2008). *Mobile Navigation Analysis*. Issue 2008.04 / 9th December 2008

- Dao, T.H.D (2005) *Performance evaluation of Multiple Reference Station GPS RTK for Medium Scale Network*. Master of Science thesis, University of Calgary.

- Günter, H. and Kneissl, F. and Ávila-Rodríguez, J. and Wallner, S. (2007) *Authenticating GNSS, Proofs against Spoofs*. InsideGNSS, July-August 2007.

- Honkavirta, V. (2008) *Location fingerprinting methods in wireless local area networks*. Master of Science thesis, Tampere University of Technology.

- IEEE (2008) *802.11k-2008*

- Leick, A. (2004) *GPS Satellite Surveying, 3rd ed.* John Wiley & Sons.

- OMA-TS-BCAST (2008) *OMA-TS-BCAST_Distribution-V1_0-20081209-C File and Stream Distribution for Mobile Broadcast Services*

- OMA-TS-SUPL-1-0 (2007) *OMA-TS-ULP-V1-0-20070615-A, User Plane Location Protocol Release 1.0*

- OMA-TS-SUPL-2-0 (2009) *OMA-TS-ULP-V2-0-20090226-D, User Plane Location Protocol Draft Release 2.0*

- OMA-AD-SUPL-2-0 (2008) *OMA-AD-SUPL-V2-0-20082706-C, User Plane Location Protocol Release 2.0 Architecture Document Candidate*

- Syrjärinne J. and Wirola L. (2006) *Setting a New Standard - Assisting GNSS Receivers That Use Wireless Networks*. InsideGNSS, pages 26–31.

- W3C (2008) *Efficient XML Interchange (EXI) Format 1.0*. W3C Working Draft 19th September.

- Wirola, L. and Alanen, K. and Käppi, J. and Syrjärinne, J. (2006) ***Bringing RTK to Cellular Terminals Using a Low-Cost Single-Frequency AGPS Receiver and Inertial Sensors***. In Proceedings of IEEE/ION PLANS 2006, 25th-27th April, San Diego, CA, USA, pages 645–652.
- Wirola, L. and Syrjärinne, J. (2007a) ***Bringing the GNSSs on the Same Line in the GNSS Assistance Standards***. In Proceedings of the 63rd ION Annual Meeting, 23rd-25th April, Boston, MA, USA, pages 242–252.
- Wirola, L., Verhagen, S., Halivaara, I. and Tiberius, C. (2007b) ***On the feasibility of adding carrier phase –assistance to the cellular GNSS assistance standards***. Journal of Global Positioning Systems, vol. 6, no. 1, pages 1-12.
- Wirola, L. and Kontola, I. and Syrjärinne, J. (2008a) ***The effect of the antenna phase response on the ambiguity resolution***. In proceeding of IEEE ION PLANS 2008, 6th-8th May, Monterey, CA, USA.
- Wirola, L. (2008b) ***RRLP Shortcomings***. Technical report LOC-2008-0385, presented in OMA LOC WG meeting, Chicago, USA, 18th-22nd August.
- Wirola, L. (2008c) ***OMA-LOC-2008-0303: Generic Fingerprinting in SUPL2.1***, presented in OMA LOC WG interim meeting in Wollongong, Australia, 19th-21st May.
- Wirola, L. (2009) ***OMA-TP-2008-0470R02 Socialization of WID 0181 – Generic Location Protocol 1.0***, provided to OMA Technical Plenary 8th January.

