

## ОБЗОР ПАРАМЕТРИЧЕСКИХ МЕТОДОВ ПОЗИЦИОНИРОВАНИЯ НА ОСНОВЕ КОНЦЕПЦИИ ОТПЕЧАТКА ПАЛЬЦА

*Под термином «позиционирование на основе отпечатка пальца» (англ. *fingerprinting* – метод отпечатка пальца, метод *фингерпринтинга*) понимается большое разнообразие методов определения местоположения приемника с использованием базы данных мощностей радиосигнала, ранее измеренных и привязанных к координатам. Непараметрические методы, напр. метод *k* взвешенных ближайших соседей (*WKNN*), невозможно использовать в крупномасштабных службах для мобильных устройств вследствие больших объемов данных и требований к передаче информации. В предлагаемой работе представлен обзор параметрических методов *фингерпринтинга*, которые используют представления данных на основе моделей. Рассмотрены три различные группы параметрических методов: методы, использующие области покрытия, методы, использующие потери при распространении сигнала, и методы, использующие смесь нормальных распределений. В рамках каждой группы рассматриваются различные подходы, их достоинства и недостатки. Обсуждается качество позиционирования в помещении с использованием некоторых из приведенных подходов в различных сценариях по данным беспроводных локальных сетей (*WLAN*). Полученные результаты сравниваются с результатами для непараметрического метода *WKNN*.*

**Ключевые слова:** *фингерпринтинг, параметрические методы позиционирования, сравнение*

---

**Мюллер Филипп.** Докторант факультета автоматизации и проектирования Технологического университета Тампере, Финляндия.

**Райтохарью Матти.** Доктор наук, научный сотрудник факультета автоматизации и проектирования Технологического университета Тампере.

**Али-Лойту Симо.** Преподаватель факультета математики Технологического университета Тампере.

**Вирола Лаура.** Научный сотрудник факультета автоматизации и проектирования Технологического университета Тампере.

**Пише Роберт.** Доктор наук, профессор Технологического университета Тампере.

Научный редактор перевода д.т.н. О. А. Степанов

## Введение

В последние десятилетия методам позиционирования уделяется значительное внимание, т.к. они стали основой для различных приложений в военной и коммерческой областях, в частности в области, связанной с общественными услугами [1, 2]. К таким приложениям относятся интеллектуальные транспортные системы, включая навигацию транспортных средств, отслеживание производственных ресурсов, управление парком воздушных судов, управление ресурсами, контроль состояния окружающей среды, аварийные службы (E911 в Северной Америке, E112 в Европе), медицинские службы (наблюдение за пациентами и оборудованием в больницах) и спасательные операции (напр., определение положения пожарных в горящем здании). В смартфонах используются такие пользовательские приложения, как идентификация местоположения, локальный поиск, предложение местных точек интереса, геолокация фото и видео, выставление счета в зависимости от местоположения и целевая реклама [2–8]. Многие из этих приложений работают на маленьких мобильных устройствах, поэтому алгоритмы позиционирования при их реализации имеют строгие ограничения по мощности, памяти, полосе частот и вычислительным ресурсам.

При работе на улице методы позиционирования основаны на использовании сигналов глобальных навигационных спутниковых систем (ГНСС). В настоящее время почти все новые смартфоны оснащены ГНСС приемниками, поэтому возможно точное постоянное позиционирование мобильных пользовательских устройств (ПУ). Однако ГНСС приемники потребляют значительное количество энергии. Кроме того, в помещении, на полосах лесонасаждений и в определенных городских условиях – например, в городских каньонах, возникают трудности при проникновении сигналов ГНСС, что обычно приводит к недостоверности или отсутствию данных о местоположении. Т.о., позиционирование в таких условиях должно опираться на другие измерения, напр., от инерциального измерительного модуля, или радиосигналы сотовых сетей, Bluetooth, сетей WLAN или сверхширокополосных сетей.

В зависимости от используемого алгоритма, позиционирование с помощью сотовых телефонов может быть простым, экономичным и не требовать обновления программного обеспечения ПУ или сетевого оборудования [3, 9, 10]. Хотя изначально сотовые сети не создавались для позиционирования<sup>1</sup>, они в большинстве случаев обеспечивают достаточную точность (ок. 100 м в помещении и в городской среде, ок. 200 м за городом) для таких приложений, как локальный поиск или прогноз погоды [12]. Поскольку сети WLAN повсеместно работают в городах и области покрытия точек доступа WLAN значительно меньше области покрытия ячеек сотовых сетей, методики позиционирования на основе WLAN имеют предпочтение перед сотовыми технологиями в качестве альтернативы для ГНСС [13], когда требуется большая точность позиционирования. Подобно сотовым сетям, сети WLAN изначально не создавались для позиционирования. Позиционирование на основе WLAN использует уже существующую инфраструктуру, и точки доступа WLAN, а приемники в ПУ также широко доступны.

В настоящей работе мы будем говорить о сетях WLAN, хотя многие обсуждаемые вопросы применимы и к сотовым сетям.

---

<sup>1</sup> Исключение составляет стандарт «Долгосрочное развитие сетей связи» (LTE), в девятой версии которого описан опорный сигнал позиционирования, что делает возможным достичь точности позиционирования порядка 10 м.

Большинство алгоритмов позиционирования на основе WLAN используют корреляцию между мощностью принимаемого сигнала (RSS) и местоположением ПУ (см. напр. [14, стр. 47]). Поскольку моделировать распространение сигнала, особенно внутри помещений, достаточно сложно, для методов позиционирования на основе фингерпринтинга (МПФП) широко используются непараметрические методы [15]. В этих методах положение пользователя оценивается путем сравнения перечня текущих измерений индикатора мощности принятого сигнала (RSSI) или RSS точек доступа с базой данных «отпечатков» точек доступа («радиокарта»), представляющих собой мощности сигналов, соответствующие значений известным значениям координат.

Параметрические методы включают разные подходы. В отличие от непараметрических методов, только некоторые суммирующие данные (отчеты о расположении, приеме или наблюдении) и параметры (напр., параметры модели распространения сигнала) сохраняются в радиокарте, значительно сокращая объем данных. Положение пользователя можно оценить с помощью параметрических методов, например, путем вычисления оценок расстояния между пользователем и точкой доступа по полученным измерениям RSS/RSSI в модели распространения сигнала.

Основная цель предлагаемой статьи состоит в том, чтобы представить обзор новейших разработок в области параметрических методов позиционирования на основе фингерпринтинга и проанализировать их достоинства и недостатки. Обзоры непараметрических МПФП приводятся, в частности, в [15, 16]. Результаты натурных испытаний опубликованы в [17]. Однако в настоящей статье дается более детальный анализ результатов и сравниваемых параметрических методов.

Статья структурирована следующим образом. В разделе I обсуждаются сходства и различия параметрических и непараметрических МПФП, а также методов сбора отпечатков и т.д. В параметрических методах используется правило Байеса и байесовская фильтрация, поэтому в разделе II кратко излагается идея определения места с помощью правила Байеса и байесовской фильтрации. В разделе III рассматриваются параметрические методы МПФП, использующие довольно простые модели для описания области покрытия точки доступа. Эти методы позволяют значительно сократить размер радиокарты, в то же время обеспечивая достаточную точность позиционирования для большого количества приложений. Раздел IV посвящен моделям потерь сигнала при распространении, которые калибруются по данным фингерпринтинга. В разделе V представлен обзор частично параметрических МПФП и смесь распределений для моделирования оценок положения и т.д. Эти методы используются в нелинейных и/или негауссовских системах, где традиционные методы типа фильтра Калмана и обобщенного фильтра Калмана работают плохо. Несколько параметрических МПФП, как включающих, так и не включающих фильтрацию, сравниваются в разделе VI в рамках оценочных испытаний с использованием реальных измерений WLAN в помещении. В заключении подводятся итоги и приводятся выводы.

#### **Список использованных сокращений**

WKNN	(weighted k-nearest neighbor) – метод k взвешенных ближайших соседей,
WLAN	(wireless local area network) – беспроводная локальная сеть,
ГНСС	(глобальная навигационная спутниковая система) – GNSS (global navigation satellite system),

ПУ	– пользовательское устройство
RSS	(received signal strength) – мощность принимаемого сигнала,
RSSI	(received signal strength indicator) – индикатор мощности принятого сигнала,
МПФП	– метод позиционирования на основе фингерпринтинга,
NN	(nearest neighbor method) – метод ближайшего соседа,
MAC	(media access point) – точка доступа,
ФПРВ	(функция плотности распределения вероятности) – pdf (probability density function)
ПНР	(потери на распространение) – PL (path loss),
IRLS	(iterative reweighted least squares) – метод взвешенных наименьших квадратов,
MMSE	(minimum mean square error) – минимальная средняя квадратическая погрешность,
МГС	(модель гауссовских сумм) – GM model (Gaussian mixture model)
СТФ	сигма-точечный фильтр – SPGMF (sigma point Gaussian mixture filter),
GMB-REM	(GM Bayes' with regularized expectation maximization) – байесовский подход на основе МГС с регуляризованной максимизацией ожидания,
EGMF	(efficient GM filter) – эффективный МГС фильтр,
GGM	(generalized version of GM) – обобщенная версия МГС,
GGMF	(filtered version of the GGM) – отфильтрованная версия GGM,
GMEM	(signal strength estimation model) – модель оценки мощности сигнала,
CA	(coverage area) – область покрытия

## I. Параметрические и непараметрические методы

В этом разделе определим, что понимается под параметрическими и непараметрическими МПФП, в чем они схожи, а в чем – различны. Далее обсудим некоторые вопросы и возможные решения, относящиеся к МПФП.

Задача параметрических и непараметрических МПФП состоит в определении вектора состояния  $\mathbf{x}_k \in \mathbb{R}^{n_x}$  размерности  $n_x$  по векторным измерениям  $\mathbf{y}_k \in \mathbb{R}^{n_y}$  размерности  $n_y$ . МПФП включают этапы оффлайн и онлайн. На этапе оффлайн состояние  $\mathbf{x}_k$  означает вектор параметров, а на этапе онлайн – положение ПУ в момент времени  $t_k$  и, возможно, дополнительные данные, например, скорость устройства.

Измерения в режиме оффлайн – это так называемые отпечатки, зафиксированные в известных точках. Иногда их называют отпечатками соответствующих положений. При локализации в помещении на основе WLAN отпечатки обычно фиксируются в узловых точках – одна узловая точка на  $m^2$  [18]. Данные отпечатка для одной такой точки включают идентификаторы точек доступа, от которых приходят сигналы на ПУ в этой узловой точке, а также соответствующие значения мощности сигнала. По собранным отпечаткам формируется радиокарта. Далее обсуждаются вопросы построения радиокарты с использованием различных параметрических и непараметрических методов.

На этапе онлайн измерения в виде отпечатков собираются в ПУ, положение которого неизвестно. Данные отпечатков зависят от местоположения ПУ. Эта зависимость между положением ПУ  $\mathbf{x}_k$  и измерениями  $\mathbf{y}_k$  используется при оценивании местоположения ПУ, при этом используются также данные радиокарты. Далее рассматриваются и анализируются некоторые методы определения местоположения.

### A. Сравнение параметрических и непараметрических МПФП

В непараметрических методах МПФП используются радиокарты, содержащие отпечатки  $\mathbf{y}_k$ . В методе WLAN радиокарта содержит координаты места, идентификаторы точек доступа, от которых получены сигналы в этой месте и

соответствующие мощности сигналов. На этапе позиционирования измерения ПУ (идентификаторы точек доступа и соответствующие значения RSS) сравниваются с данными радиокарты для получения оценки положения. Самый простой подход – это метод ближайшего соседа (NN). Он выдает в качестве оценки местоположения положение отпечатка из радиокарты, чье измерение наиболее близко к измерению ПУ. Такой отпечаток называется ближайшим соседом и находится путем оптимизации заданной функции стоимости [19].

Более усовершенствованная, широко применяемая версия метода ближайшего соседа – метод  $k$  взвешенных ближайших соседей (WKNN). Здесь оценка положения является взвешенным средним  $k$  положений, чьи отпечатки наиболее близки к измерению ПУ [15]. В соответствии с работой [7] (Liu и др.), в методе WKNN совмещается средняя сложность и средняя стоимость с высокой робастностью и точностью. Поэтому мы используем его в разделе VI для сравнения с параметрическими методами местоопределения. Обзоры непараметрических МПФП приведены в [15, 20] и в ссылках из этих публикаций.

Параметрические МПФП используют радиокарты, содержащие запоминаемые параметры  $x_k$ , аккумулирующие отпечатки  $y_k$ . Радиокарта содержит набор параметров для каждой точки доступа в наблюдаемых отпечатках (в отличие от принципа «один отпечаток на каждую узловую точку»). Для позиционирования можно использовать разные методы, зависящие от параметров, используемых в радиокарте. Далее мы рассмотрим эти методы.

Непараметрические МПФП имеют следующее преимущество: в них не требуется моделирование распространения сигнала. Эти алгоритмы проявили достаточную точность и надежность при работе в помещении (см. напр. [15, 21]), где часто встречаются ситуации отсутствия прямой линии видимости. Многие параметрические МПФП требуют моделирования распространения сигнала, что вызывает проблемы при работе в помещении. Далее обсудим этот вопрос более подробно.

Главное различие между параметрическими и непараметрическими методами состоит в том, что в первых объем данных радиокарты зависит от количества точек доступа и количества параметров, сохраненных для каждой точки, а во вторых – от количества отпечатков (т.е. узловых точек). Это означает, что в непараметрических методах при сборе большого количества отпечатков для повышения качества радиокарты ее размер увеличивается, в то время как в параметрических методах он остается неизменным, если не обнаруживаются дополнительные точки доступа.

Чтобы лучше представить себе, о каких объемах данных идет речь, следует отметить, что для каждой узловой точки собирается ок. 100 реализаций [18], чтобы получить достоверный отпечаток для радиокарты. Поскольку непараметрические МПФП работают непосредственно с данными отпечатков, объем этих данных может быть решающим для случаев использования МПФП в качестве крупномасштабной службы для мобильных устройств. В работе [12] Wirola и др. указывается, что радиокарта площади размером всего 0,1% поверхности Земли (ок. 130 000 км<sup>2</sup>, или размер Греции или американского штата Луизиана) со средней плотностью 1 отпечаток/м<sup>2</sup> и минимум один 6-байтовый MAC адрес/точку займет по меньшей мере 780 Гб. При работе в помещении часто есть возможность получить сигналы от более чем 5 точек доступа. Также, кроме MAC адресов точек доступа хранятся значения мощности сигналов.

Один из методов уменьшения объема радиокарты в непараметрическом методе включает использование технологии сжатия данных [22, 23]. Более фунда-

ментальный путь решения данной проблемы – использовать параметрические МПФП (методы на основе модели). Поскольку сигнал от точки доступа обычно можно получить во многих узловых точках, радиокарта в параметрических методах обычно значительно меньше [17].

Перед тем как переходить к обсуждению различных параметрических методов, рассмотрим, как происходит сбор отпечатков, а также некоторые вопросы и решения, связанные с ним.

#### *Б. Фингерпринтинг и связанные вопросы*

Для позиционирования методом фингерпринтинга производится сбор отпечатков на этапе оффлайн съемки местности, вардрайвинга (англ. war-driving) или краудсорсинга (crowd sourcing). Кроме текущего положения пользователя, каждый отпечаток содержит характеристики радиосигнала. В сетях WLAN эти характеристики обычно содержат как минимум идентификаторы точек доступа и соответствующие значения RSS или RSSI. Под съемкой понимается, что сбор отпечатков производится в разных точках для создания базы данных отпечатков. Чаще всего ввод точек местоположения производится вручную. В этом состоит серьезное отличие от метода вардрайвинг, где местоположение обычно является оценкой по данным ГНСС. Как следует из названия, этот метод состоит в сборе измерений с движущегося транспортного средства. Недостатком двух названных методов является то, что сбор данных представляет собой трудоемкую и дорогостоящую процедуру. Т.о., предпочтительно использовать метод краудфандинга, где сбор отпечатков осуществляют несколько человек (напр. пользователи МПФП). На практике может использоваться комбинация этих методов.

При работе на улице сбор отпечатков можно производить по методу краудсорсинга или вардрайвинга, что делает обновление радиокарт менее трудоемким. Однако вследствие отсутствия сигналов ГНСС в помещении такую технологию использовать нельзя, и возникает необходимость применения более сложных альтернативных методов. Следует отметить, что возможно и использование отпечатков пальцев без ГНСС (т.н. неопределенные отпечатки), но тогда точность положения снижается [24].

Важным вопросом в МПФП является «поддержка» радиокарты. Поскольку топография сети постоянно меняется (точки доступа могут добавляться, удаляться, перемещаться или модифицироваться), как и радиотехническая обстановка, необходимы постоянные обновления отпечатков радиокарты, чтобы избежать снижения качества работы [15, 25]. Даже небольшие изменения в среде могут оказывать огромный эффект на измеряемые значения RSS. Если, например, отпечатки были собраны в пустом здании, а позиционирование производится в нем, когда оно заполнено людьми, значения RSS в определенном положении будут значительно отличаться вследствие затенения фигурами людей [18]. Влияние тела человека на RSS было проанализировано, напр. в работе [26] (Kaemarungsi и Krishnamurthy).

Еще одна проблема, требующая рассмотрения – это разнородность устройств. Она состоит в том, что полученные мощности сигналов, измеренные разными устройствами в одном месте и в одно время, могут значительно различаться вследствие разнородности этих устройств [13, 21, 27–29]. Это явление включает, напр., тот факт, что многие устройства отображают только безразмерные значения RSSI, а не RSS, которые всегда измеряются в дБм. Производители наборов микросхем используют разные шкалы RSSI, отличающиеся границами и степенью детализации, что мешает производить сравнение измерений

RSS от разных устройств [27, 29–31]. Чтобы обойти проблему разнородности RSS, некоторые авторы используют ранжирование значений RSS в отпечатке [21, 32], соотношения [31] или различия значений RSS [33] вместо измеренных их значений. Далее мы будем разделять RSS и RSSI, в зависимости от того, какие значения были использованы цитируемыми авторами или требуются для специфической задачи.

Чтобы разрешить проблему объема данных RSSI, можно применить калибровку приборов; краткий обзор этих методов дан в [28]. В более ранних методиках используется ручная калибровка: пользователи производят сбор измерений в известных точках для калибровки своего устройства. Обычно там применяются линейные функции калибровки, подобранные методом наименьших квадратов [27, 30]. В работе Park и др. [13] авторы возражают, утверждая, что калибровка мощности сигнала недостаточна и различия в дисперсии мощности сигнала также важны, и описывают метод базовой оценки (kernel estimation) для калибровки приборов. Методики автоматической калибровки, которые позволяют одновременно производить калибровку и позиционирование и не требуют проведения измерений в известных опорных точках, предложены, напр., в [28] и [29]. В работе [29] (Koski и др.) предлагается построение гистограммы RSSI по значениям, собранным за долговременный период. Допуская, что распределение значений RSSI для всех приборов одинаковы, включая пересчет и масштаб, эти специфические для приборов параметры могут быть подобраны по гистограмме. В работе Laoudias и др. [28] более тщательно исследуется метод гистограмм. Демонстрируется, что допущение о том, что распределение реальных значений RSS одинаково для всех пользователей, позволяет оценивать параметры для линейного отображения RSS приборов. В отличие от работы Park и др. [13], авторы считают, что эффект разнородности устройств в значительной степени снижается за счет линейного отображения, и достаточно будет произвести простую установку нуля. Описывается метод постоянного обновления калибровки при позиционировании, и определяется, что позиционирование таким методом такое же точное, как ручная калибровка, за исключением начала этапа обучения.

## **II. Позиционирование и Байесовская фильтрация**

Все параметрические методы, представленные в статье, используют правило Байеса для определения параметров и статического позиционирования. В задачах позиционирования, связанных с фильтрацией, применяется байесовская фильтрация. Далее мы кратко излагаем байесовские методы для статических и фильтрационных методов позиционирования.

В статическом случае состояние системы  $\mathbf{x}_k$  оценивается следующим образом. Измерения вектора  $\mathbf{y}_k$  используются для определения апостериорной функции плотности распределения вероятности (ФПРВ) состояния по правилу Байеса:

$$p(\mathbf{x}_k|\mathbf{y}_k) \propto p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k), \quad (1)$$

где  $p(\mathbf{x}_k)$  – априорная ФПРВ.

Задача позиционирования формулируется как задача байесовской фильтрации для модели в пространстве состояний для дискретного времени. В рамках данного раздела мы рассматриваем общую модель в пространстве состояний

для дискретного времени, а именно, нелинейную систему с негауссовским аддитивным шумом:

$$\mathbf{x}_k = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1} \quad (2a)$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k \quad (2b)$$

где  $\mathbf{w}_{k-1}$  и  $\mathbf{v}_k$  считаются белыми шумами, не зависимыми друг от друга и от начального состояния  $\mathbf{x}_0$ ;  $\mathbf{f}_{k-1}(\cdot)$ ,  $\mathbf{h}_k(\cdot)$  – нелинейные функции, которые полагаются известными. Далее, ФПРВ  $\mathbf{w}_k$  и  $\mathbf{v}_k$  обозначены  $p(\mathbf{w}_k)$  и  $p(\mathbf{v}_k)$ , соответственно. Задача фильтрации заключается в нахождении условной (апостериорной) ФПРВ  $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ , где  $\mathbf{y}_{1:k} \triangleq (\mathbf{y}_1, \dots, \mathbf{y}_k)$ . Апостериорная ФПРВ определяется рекуррентно в соответствии с соотношениями [34].

*Прогноз (априорная функция):*

$$p(\mathbf{x}_k|\mathbf{y}_{k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1})d\mathbf{x}_{k-1}, \quad (3)$$

*Коррекция (апостериорная функция):*

$$p(\mathbf{x}_k|\mathbf{y}_{1:k}) = \frac{p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k-1})}{\int p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k-1})d\mathbf{x}_{k-1}} \quad (4)$$

где переходная ФПРВ равна  $p(\mathbf{x}_k|\mathbf{x}_{k-1}) = p_{\mathbf{w}_k}(\mathbf{x}_k - \mathbf{f}_{k-1}(\mathbf{x}_{k-1}))$ , а функция подобия

$$p(\mathbf{y}_k|\mathbf{x}_k) = p_{\mathbf{v}_k}(\mathbf{y}_k - \mathbf{h}_k(\mathbf{x}_k)) \quad (5)$$

Исходное условие для рекурсии – ФПРВ начального состояния  $p(\mathbf{x}_0|\mathbf{y}_{1:0}) = p(\mathbf{x}_0)$ .

В общем случае и в случаях, проанализированных далее, условную ФПРВ невозможно определить аналитически. Поэтому существует множество аппроксимирующих методов для вычисления апостериорного среднего (см. напр. [35]). Кроме апостериорного среднего, эти методы обычно обеспечивают получение апостериорной матрицы ковариаций. В случае, если истинное состояние известно, можно проверить, согласована ли апостериорная матрица, напр. с помощью проверки гауссовской согласованности (Gaussian consistency test) [36, стр. 235]. Идея этих проверок состоит в оценке точности матрицы ковариаций оценки состояния. Например, в такой проверке при уровне риска 5%, апостериорное значение считается согласованным, если апостериорная ковариационная матрица  $\mathbf{P}_k$  и апостериорное среднее  $\hat{\mathbf{x}}_k \in \mathbb{R}^2$  соответствуют неравенству  $(\hat{\mathbf{x}}_k - \mathbf{x}_k)^T \mathbf{P}_k^{-1} (\hat{\mathbf{x}}_k - \mathbf{x}_k) \leq \chi_2^2 = 5.99$ , где  $\mathbf{x}_k$  – истинное состояние.

### III. Модели области покрытия

#### A. Оценка области покрытия

Простой с вычислительной точки зрения метод параметрического фингерпринтинга предложен в [20, 29]. Чтобы сократить размер радиокарты, авторы представляют область покрытия (CA, coverage area) (или область приема) любого узла сети (в сети WLAN узел сети называется точкой доступа, в сотовых радиосетях – базовой станцией) в виде эллиптического распределения вероятности, для которого возможно произвести аналитическую оценку параметров распределения и местоположения. Распределение вероятности представляет только ту область, в которой может быть получен сигнал от точки доступа, он не содержит данных об RSS, кроме предполагаемого порога мощности приема. Поскольку реальные области покрытия зачастую имеют неправильную форму, моделирование их посредством простых форм дает возможность сохранять ком-



пактный размер базы данных этих областей, обеспечивая при этом приемлемую точность. Это также обеспечивает возможность быстрой передачи данных ПУ [20, 29] и быстрый расчет положения этого устройства, т.к. любой эллипс можно представить пятью параметрами: три параметра эллипса с центром в точке начала координат (симметричная положительно определенная матрица  $2 \times 2$ ), и еще два параметра для указания местоположения центра эллипса [37]. В работах Koski и др. [29] и Piché [37] указывается, что для моделирования областей покрытия узлов сети можно использовать также и другие формы, напр. окружности (три параметра) или многогранники (минимум шесть параметров). Более того, стандарт 3GPP TS 23.032 поддерживает использование таких геометрических форм, как эллипсы, многогранники и эллипсоиды [4, стр. 98].

Область покрытия моделируется в [20, 29] апостериорным распределением для параметров эллипса  $\theta$  при расположениях отпечатков  $\mathbf{z} = \{z_1, z_2, \dots, z_n\}$ , где воспринимается узел. Распределение описывается правилом Байеса

$$p(\theta | \mathbf{z}) \propto p(\mathbf{z} | \theta) p(\theta) \quad (6)$$

где функция правдоподобия и априорная ФПРВ являются гауссовскими. Другими словами, область покрытия моделируется путем подбора среднего и ковариации многомерного гауссовского распределения к данным. На Рис. 1 показано расположение отпечатков  $\mathbf{z}$ , а также среднее и ковариация подобранного многомерного гауссовского распределения, параметры которого содержатся в  $\theta$  (для среднего и для симметричной матрицы ковариации). Эллипс соответствует примерно 40% двумерной вероятности попадания в него двумерного гауссовского вектора.

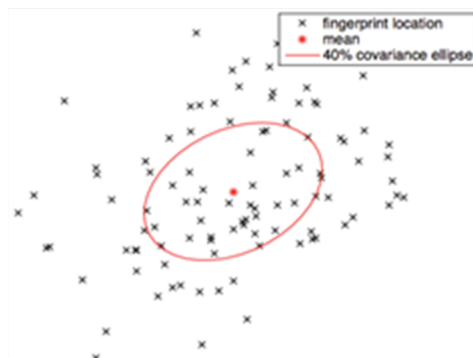


Рис. 1. Расположения отпечатков  $\mathbf{z}$ , где получен сигнал точки доступа, а также среднее и эллипс ковариаций подобранного двумерного гауссовского распределения. Расположение отпечатков, среднее, эллипс ковариации 40%

Использование байесовской постановки задачи регрессии имеет два преимущества. Во-первых, байесовская априорная ФПРВ  $p(\theta)$  позволяет использовать данные о типичных областях покрытия, что имеет большое значение, если имеется небольшое количество отпечатков [20, 29]. Эти данные можно получить путем экспериментальных исследований. Например, для сети WLAN типичный диапазон приема в помещении составляет 20–50 м [6, стр. 9]. В соответствии с работами Trevisani и Vitaletti [3], а также Molisch [14, стр. 17], размер области покрытия в сотовой сети сильно зависит от типа ячейки; он может

разниться от 1 м для наноячеек до 30 км для макроячеек. Далее, авторы Trevisani и Vitaletti [3] указывают, что на размер влияет множество факторов – помехи, местная ожидаемая нагрузка и чувствительность антенны ПУ. Такую вариативность также можно смоделировать с помощью априорной байесовской ФПРВ путем, применяя распределение с большей дисперсией.

Во-вторых, использование правила Байеса с заданием модели в пространстве состояний (2) для нахождения параметров эллипса позволяет производить рекурсивное оценивание и обновление оценок [29]. Обновление апостериорной ФПРВ  $p(\theta | \mathbf{z})$  по мере появления новых отпечатков осуществляется с помощью

правила Байеса [38, стр. 14] или расчета ее как временного ряда [38, стр. 29]. В работе Koski [38] указывается, что во втором методе учитывается, что параметры могут меняться со временем. Эти изменения встречаются часто и могут быть вызваны, например, постройкой новых зданий, изменением планов этажей или модификацией топологий радиосетей [4, 6, 14, 37, 38].

Одно из возможных замечаний к вышеописанному методу состоит в том, что в работах [20, 29] допускается, что отчеты о расположении моделируются с помощью гауссовского (нормального) распределения. Общеизвестно, что модель нормальной регрессии недостаточно робастна, в том смысле, что резко отклоняющиеся отчеты могут привести к переоценке области покрытия (см. напр. [39]). Причины таких отклоняющихся отчетов включают условия приема, сбои программного обеспечения или аппаратуры ГНСС (при использовании ее для определения координат отчетов) или прием радиосигналов [37]. В работе Riché [37] автор возражает, что большие «выбросы» можно легко обнаружить эвристическими методами, средние же выбросы обнаружить сложно, особенно если в перечне отпечатков, используемых для определения области покрытия, встречается большое количество выбросов. Поэтому автор рекомендует моделировать отчеты о расположении как отчеты с распределением Стьюдента, так чтобы выбросы автоматически устранялись за счет более тяжелых хвостов распределения. В [37] показано, что регрессию Стьюдента можно рассчитать при помощи алгоритма выборки Гиббса [40] или алгоритма максимизации ожидания (Expectation Maximization) [41]. Для обоих алгоритмов в [37] указано, что отчеты о расположении со стьюдентовым распределением также могут включать данные о типовых областях покрытия как информативные априорные ФПРВ, как и в случае с отчетами с гауссовским распределением. Введение таких априорных ФПРВ требует небольших изменений в алгоритме максимизации ожидания или модели Гиббса.

Рассмотренные выше методы полностью игнорируют значения RSS/RSSI, соответствующие идентификаторам воспринимаемых узлов. Поэтому они менее чувствительны к изменениям в радиосвязи, чем МПФП, в которых эти значения используются. Однако такое повышение робастности приводит к меньшей точности по сравнению с непараметрическими МПФП (напр. WKNN), которые, помимо расположения отпечатков и идентификаторов узлов, наблюдаемых в каждом отпечатке, также сохраняют соответствующие значения RSS или RSSI. Значения RSS, а следовательно и RSSI, зависят от расстояния между узлом сети (излучатель, передатчик) и пользовательским устройством (приемник) и обычно моделируются как функция этого расстояния по моделям потерь на распространение, которые будут оговорены в разделе IV.

Метод оценки области покрытия, в котором используются данные RSS, предлагается в [20, 39]. Для каждого узла сети в базе данных сохраняется не одна, а несколько областей покрытия, которые моделируются по данным отпечатков, сгруппированных в соответствии с RSS. В работе [39] авторы исследуют использование одной, двух и трех областей покрытия для одного узла сети при допущении о гауссовском или стьюдентовом распределении отчетов о расположении. Отпечатки группируются по данным RSS, и разные области покрытия создаются только по отчетам соответствующей группы. Рассмотрены три различные правила группировки: уровень RSS,  $n$  самых сильных узлов сети каждого отпечатка и  $x\%$  сильнейших узлов сети каждого отпечатка.

*Б. Позиционирование с использованием областей покрытия*

Оценку положения ПУ с помощью областей покрытия [20, 29, 39] можно получить, применив правило Байеса. Оценку положения и меру ее неопределенности можно извлечь из гауссовской апостериорной ФПРВ  $p(\mathbf{x}|\mathbf{c})$  положения  $\mathbf{x}$  пользовательского устройства, где  $\mathbf{c} = (c_1, c_2, \dots, c_N)$  – перечень узлов сети, наблюдаемых ПУ. Для гауссовской априорной ФПРВ этого положения выбираются подходящее среднее и ковариация, которые представляют априорные знания о положении ПУ. Если таких данных нет, оправдано задание ковариации [20] с большими значениями для диагональных элементов. Для вычисления функции правдоподобия  $p(\mathbf{c}|\mathbf{x})$  [20, 29, 39] полагается, что априорные вероятности наблюдения  $c_n$  равны для всех  $n = 1, \dots, N$ , и что наблюдения условно независимы от данного  $\mathbf{x}$ .

Это допущение рассматривается как недостаток модели [20]. Например, для соседних узлов  $c_i$  и  $c_j$ , предположение о независимости часто нарушается. Если  $c_i$  наблюдается в  $\mathbf{x}$ , это очевидно влияет на вероятность наблюдения  $c_j$ , т.к. области покрытия обычно перекрываются, то есть оба узла могут наблюдаться в одной и той же области. В работе Koski и др. [20] авторы предполагают, что использование данных по узлам сети, ненаблюдаемых в  $\mathbf{x}$ , может значительно повысить точность позиционирования. Однако, в силу невозможности получения аналитических решений для таких данных, такой подход не может конкурировать с предлагаемым алгоритмом в плане вычислительной сложности.

В работе Koski и др. [29] авторы указывают, что точки доступа в сети WLAN (т.е. узлы сети) часто могут быть наблюдаемы по всему зданию, что делает бессмысленным использование их областей покрытия для позиционирования. Поэтому удаление некоторых узлов на этапах обучения и позиционирования может повысить точность позиционирования [29], [42, стр. 71]. Для выбора узлов, подлежащих удалению, рассмотрено несколько методов, включая прямую «селекцию» и обратное исключение, взвешивание узлов с помощью общей перекрестной проверки на достоверность, выбор узлов по приросту данных, мере расхождения или показателю исключения. Более подробный обзор дан в работе [43].

В работах [29] и [44] авторы предлагают использовать значение порога мощности сигнала для сокращения количества наблюдений. Узлы, наблюдаемые с RSSI ниже этого порога, удаляются из отпечатков, т.е. исключаются из определения области покрытия и позиционирования. Чтобы обеспечить сопоставимость RSSI разных значений, авторы в [29] применяют метод гистограмм RSSI, упомянутый в разделе I. Для позиционирования по измерениям WLAN, средние и среднеквадратические ошибки позиционирования были значительно снижены путем применения такой методики. Вместо метода гистограмм можно использовать другие калибровочные методы, которые были описаны ранее. Однако при использовании метода исключения по мощности сигнала необходимо убедиться, что удаление узлов не нарушает преимущественности оценок положения, и что не произойдет излишнего исключения, когда (почти) все узлы оказываются удалены из списка наблюдений. Такая проблема, например, может возникнуть, при измерениях WLAN вне помещения, где значения мощности сигнала и их динамика значительно слабее, чем в помещении, т.к. сигналы обычно проходят через толстые стены.

Это объясняет, почему использование нескольких областей покрытия для одного узла при WLAN-позиционировании вне помещений дает лишь неболь-

шие улучшения в сравнении с использованием одной области для одного узла [39], в то время как в помещении оно дает значительно более высокую точность [20, 39]. Альтернативно, можно избежать излишнего исключения, используя правило  $n$  сильнейших или  $x\%$  сильнейших измерений, которые не требуют калибровки RSS.

#### IV. Модели потерь сигнала на распространение

Модели потерь на распространение (ПНР) относятся к моделям потерь мощности сигнала  $L_p$  или мощности принимаемого сигнала  $P_{RSS}$  по радиоканалу, осредненного по сильному и слабому затуханию (large-scale and small-scale fading) [14, стр. 127]. В простейших моделях потери на распространение зависят только от мощности передачи и расстояния  $d$ , проходимого радиоволной; в более сложных моделях учитываются и другие факторы. Обзор механизмов распространения и моделей потерь на распространение см. в работах [6, 14, 45] и содержащихся в них ссылках.

Соотношение значения RSS и расстояния, пройденного радиоволной, можно использовать при позиционировании. По измерениям RSS и моделям ПНР оцениваются расстояния между набором опорных узлов и целевым узлом, что затем дает возможность оценить расположение целевого узла. Однако оценка положения чувствительна к шуму сигнала и неопределенности параметров модели потерь на распространение, т.к. градиент «расстояние-мощность» относительно мал [1]. Следовательно, эти оценки обычно бывают менее точными, чем оценки на основе радиосигналов, которые получены с использованием угла наклона траектории сигнала или измерений временных задержек. Однако авторы Patwari и др. [46] показывают, что алгоритмы позиционирования с достаточно высокой плотностью узлов сети, опирающиеся на модели ПНР (и т.о. на RSS) могут обеспечить такое же качество работы, как алгоритмы на основе временных задержек.

##### *A. Оценка параметров моделей потерь на распространение*

В более ранних исследованиях предполагалось, что параметры моделей ПНР известны априори, что является излишним упрощением в некоторых практических приложениях, и, следовательно, слабо обосновано [47]. Т.о., параметры модели должны оцениваться по данным отпечатков, включающим идентификаторы узлов сети и соответствующие значения RSS.

Существуют разные методы оценки параметра(ов) модели. В некоторых сначала оценивается положение узла (или оно полагается известным), а по этим данным оценивается параметр(ы) (напр. [48, 49]); в других, положение узла и параметр(ы) модели ПНР оцениваются одновременно (напр. [47, 50]).

В работе [48] предлагается метод статистического моделирования, в котором положение ПУ оценивается по статистической модели мощности сигнала, а положение узла считается известным. Параметры модели оцениваются по алгоритму максимизации ожидания [41], который используется, чтобы найти значения, максимизирующие функцию правдоподобия. Авторы указывают, что основное отличие их алгоритма в сравнении с геометрическим подходом состоит в том, что он выводит данные о свойствах сигнала из расположения узлов. Однако зачастую расположение узлов неизвестно и также подлежит оцениванию. В работе Li [47] автор обнаружил, что оценка положения узла при фиксированных параметрах модели ПНР может приводить к большим ошибкам, если значения параметров выбраны неточно, и рекомендует одновременное оценивание.

Далее указывается, что совместная оценка исключает необходимость расширенных измерений в каналах.

В работе [47] оцениваются положение узла и экспонент ПНР  $n$  (градиент «расстояние-мощность») классической модели ПНР распространения радиосигнала в узком диапазоне<sup>2</sup>

$$L_p(d) = L_p(1) + 10n \log_{10}(d) + w, \quad (7)$$

где  $L_p(d)$  – потери мощности сигнала в дБ на расстоянии  $d$  м от узла. Центрированная гауссовская случайная величина  $w$  с дисперсией  $\sigma_w^2$  используется для моделирования медленного затухания. Приблизительно логарифмически нормальное распределение медленного затухания, которое подразумевает гауссовское распределение  $w$  в (7), наблюдалось эмпирическим образом, напр. в [51–54]. Для решения задачи оценивания в работе Li [47] используется метод Левенберга-Маркара, модифицированный алгоритм Гаусса-Ньютона в виде системы нелинейных уравнений.

В работе Nurminen и др. [50] авторы идут на один шаг дальше и дополнительно оценивают кажущуюся мощность передачи  $A = P_{RSS}(1)$  для разновидности логарифмической модели расстояния:

$$P_{RSS}(d) = A - 10n \log_{10}(d) + w, \quad (8)$$

используя итеративный метод взвешенных наименьших квадратов (IRLS), который также является алгоритмом Гаусса-Ньютона. В соответствии с работой Dil и Havinga [55], (8) можно использовать для описания зависимости расстояния  $d$  от  $P_{RSS}$  при работе в помещении. В работе [50] применяется байесовский алгоритм, который хорошо подходит для обновления оценки положения узла и параметров модели ПНР по мере появления новых данных отпечатков. В алгоритме используются неинформативные априорные гауссовские ФПРВ. Имея достаточное количество отпечатков, в соответствии с работой Nurminen и др. [50], априорные средние значения для параметров модели ПНР можно выбрать произвольным образом, т.к. при большом количестве отпечатков апостериорное распределение обычно бывает одномодальным. Эта идея поддерживается выводом в исследовании Li [47] о том, что влияние неточного априорного значения экспонента ПНР  $n$  на результаты оценки ничтожно мало. Однако автор указывает, что, особенно для случаев с ограниченным объемом данных, правильно выбранное информативное априорное значение будет преимуществом. В различных исследованиях выявлены значения экспонента ПНР для разных условий и сетей (напр. [45, 56]); для кажущейся мощности передачи имеется меньше материалов (напр. [56]).

При учете априорно известного положения узла следует проявить особую осторожность, чтобы алгоритм IRLS не поместил узел в область слабых значений RSS [50]. Тем не менее, даже при таких мерах нельзя гарантировать, что алгоритмы обеспечат нахождение правильного положения узла. Например, карта RSS может содержать несколько пиковых значений, реальное положение узла может находиться вне пределов карты RSS, или количество измерений может быть недостаточным для определения пика [50]. Оба метода: байесовский метод IRLS в [50] и метод Левенберга-Маркара в [47] дают пользователю «ин-

<sup>2</sup> В работе Roos и др. [48] используется более сложная модель ПНР, включающая такой параметр, как направление передачи. Поскольку работа посвящена сотовым сетям, в которых узлы (базовые станции) в основном имеют направленные антенны, это должно обеспечить более точные оценки ПНР. Для изотропных точек доступа WLAN, которые используются у Li [47], ПНР должны быть (теоретически) одинаковы во всех направлениях.

струмент» для различения надежных и ненадежных оценок положения и модели ПНР в виде ковариационной матрицы. В последнем методе, ковариацию можно вычислить после нахождения оптимальных оценок, а в первом – она автоматически доступна в виде апостериорной ковариационной матрицы. Также, метод, рассмотренный в [50], учитывает корреляцию ошибок измерений путем добавления малой постоянной диагональной матрицы к матрице ковариаций положения узла (в работе Li [47] все наблюдения полагаются статистически независимыми). Взаимной корреляцией между положением узла и параметрами модели ПНР тем не менее пренебрегают, главным образом с целью ограничения количества их параметров.

Одна из возможных претензий к методам, описанным в [47, 50] состоит в том, что авторы считают среднеквадратическое отклонение  $\sigma_w$  компонента затухания постоянным, хотя в этих же работах говорится, что он сильно зависит от среды распространения, что подтверждается, напр., в работе Ghassemzadeh и др. для сверхширокополосных сетей [54]. Типовые значения разнятся от 1 до 6 дБ [56, 57] для сетей WLAN. Также наблюдались и большие значения, особенно в высоких зданиях (см. напр. [57, стр. 139]). В сотовых сетях наблюдались значения от 5 до 16 дБ [52, 56]. В работе Nurminen и др. [50] используется фиксированное значение  $\sigma_w = 6$  дБ, в то время как в работе Li [47] изучается влияние изменяющихся значений, зафиксированных при оценивании ошибок в оценках положения узлов. Испытания показывают, что значение  $\sigma_w$  оказывает значительное влияние на систематическую погрешность и эффективность алгоритмов оценивания положения в зависимости от используемого метода оценивания.

В работе [18] Nan и др. авторы игнорируют среднеквадратическое отклонение компонента затухания и вычисляют точечные оценки для параметров модели ПНР. Однако этот подход также заслуживает внимания, т.к. позволяет построить радиокарту по значительно меньшему объему данных отпечатков. Далее модель ПНР, которая является продолжением модели Зейделя [58], учитывает углы между траекторией сигнала и преградами. Это означает, что моделируемые ПНР зависят от угла, под которым сигнал сталкивается с преградой. Например, если сигнал сталкивается со стеной под углом 90 градусов, расстояние, которое он проходит через стену, значительно короче, чем при угле 60 градусов. Модель ПНР, описанная в [18], учитывает это.

Для построения радиокарты авторы предлагают сбор данных отпечатков только для небольшой части области, покрываемой радиокарткой. Затем отпечатки используются для построения линейного уравнения, решаемого методом наименьших квадратов для получения параметров модели ПНР. Далее модель ПНР используется для выработки отпечатков для тех мест, по которым не производилось измерений. Этот второй шаг, однако, необязателен, если для позиционирования используются параметрические методы, как станет видно в следующем подразделе. В работе [18] авторы показывают, что достаточно собрать отпечатки от 20% узловых точек, для которых имеются отпечатки на радиокартке. Как показали проведенные испытания, радиокартка, использующая эти точки и такой метод генерирования отпечатков для других 80% узловых точек, имеет такую же кумулятивную функцию распределения ошибки прогноза, как и карта, где были собраны отпечатки всех узловых точек.

В работе Shrestha и др. [19] используются методы обратной свертки, чтобы сократить размер карты в 10 раз по сравнению с картой для непараметрических методов. Кроме модели ПНР (8), они также исследуют многотактную (multi-slope) модель ПНР, которая принимает различные значения  $n$  в зависимости от

расстояния между передатчиком и приемником. Далее, авторы расширяют обе модели с помощью дополнительного параметра потерь при прохождении этажей и рассматривают трехмерные положения.

Параметры моделей ПНР оцениваются в [19] следующим образом. Для каждого отпечатка вычисляются оценки параметров по методу наименьших квадратов, взвешенных наименьших квадратов или минимальной средней квадратической погрешности (MMSE), при этом под положением отпечатка понимается положение точки доступа. Затем вычисляются ожидаемые значения RSS в этой точке по данным о текущих оценках параметров. Расположение отпечатков, для которого среднеквадратическая ошибка измеренных и ожидаемых значений RSS минимальна, выбирается в качестве оценки положения точки доступа. Альтернативно, выбирается среднее из  $k$  отпечатков, которые дают минимальную среднюю квадратическую погрешность. В этом случае параметры модели ПНР следует пересчитать по новой оценке положения точки доступа.

#### *Б. Позиционирование по моделям потерь на распространение*

Когда произведена оценка параметров модели ПНР и расположения всех узлов, по модели ПНР и измеренным значениям RSS можно получить оценки дальности. Затем для вычисления оценки положения ПУ можно применить метод трилатерации или другую методику нелинейного оценивания.

В работе Nurminen и др. [50] испытываются три различных метода, использующие модель ПНР (8) с реальными данными WLAN в офисных условиях: использование интегрирования по Монте-Карло для расчета правдоподобия в каждой точке пространственной сетки, алгоритм Метрополиса-Гастингса и IRLS. Для сравнения авторы применяют метод оценивания области покрытия, представленный в [20, 29], отфильтрованный с помощью стандартного фильтра Калмана, и WKNN с  $k = 3$  и неотфильтрованными измерениями. Этаж в процессе испытаний считается известным. Далее производится анализ метода сеток, алгоритма Метрополиса-Гастингса и IRLS с помощью точечных оценок и гауссовских распределений значений параметров модели ПНР.

При использовании большого объема данных отпечатков для оценки положения узлов и оценок параметров, метод WKNN обеспечивает наилучшую точность, за ним идут метод сеток, алгоритм Метрополиса-Гастингса и IRLS, которые дают примерно одинаковую точность. Метод оценивания области покрытия работает хуже всего. Однако при ограниченном объеме данных отпечатков для каких-то узлов, точность метода WKNN снижается до уровня метода сеток, алгоритма Метрополиса-Гастингса и IRLS. Это соответствует ранее сделанным выводам. Например, в работе Dil и Havinga [59] выясняется, что при ограниченном объеме данных отпечатков непараметрические МПФП, напр. WKNN, работают хуже, чем алгоритмы на основе дальности (range-based algorithms).

Испытания, проведенные в [50], показывают, что допущение о гауссовском распределении (а не точечных оценках) параметров в общем случае дает положительный результат; преимущество использования распределения становится более ясно в испытаниях с ограниченным набором отпечатков. Эти результаты не вызывают удивления, т.к. модель ПНР содержит ошибки аппроксимации [60]. Если для оценки модели ПНР доступно меньшее количество отпечатков, эти ошибки, как правило, выше. Следовательно, в таких ситуациях с теоретической точки зрения было бы выгодно предположить большую неопределенность оценок параметров.

Другое возможное объяснение того, что допущение о распределении экспонента ПНР дает лучшие результаты, чем допущение о точечных оценках, состо-

ит в том, что экспонент ПНР  $n$  может считаться постоянным только в течение ограниченного времени в среде [47]. Однако, если значение меняется, оно должно быть близко к предыдущему значению, пока окружающие условия не меняются. Это можно в некоторой степени отразить путем допущения некоторой неопределенности в оценках экспонента ПНР. Дополнительно, сохранение неопределенностей также позволяет рекурсивно обновлять параметры и использовать модели эволюции во времени по мере появления новых отпечатков.

Что касается вычислительной нагрузки, метод сеток и алгоритм Метрополиса-Гастингса не имеют преимуществ перед WKNN, в то время как IRLS значительно быстрее, время его работы близко к методу оценивания области покрытия.

Как отмечалось в предыдущем подразделе, в работе Хан и др. [18] используется модель ПНР и некоторые данные отпечатков для построения полной радиокарты определенной области. Для позиционирования затем используется непараметрический метод WKNN. Испытания показывают, что точность позиционирования таким методом немного ниже, но построение карты происходит гораздо быстрее. Оно происходило бы еще быстрее, если бы авторы могли сохранить параметры модели ПНР для каждого узла радиокарты, а на этапе позиционирования вывести оценки дальности из модели ПНР и значений RSS, а затем, применив какую-либо методику нелинейного оценивания, получить оценку положения ПУ.

В работе Shrestha и др. [19] авторы идут таким же путем, как Хан и др. [18], строя полную радиокарту по параметрам ПНР на этапе позиционирования. Их идея состоит в том, чтобы создать искусственную сетку и для каждой узловой точки рассчитать разности мощности сигнала между измеренными и ожидаемыми значениями RSS ПУ в этой точке. Затем можно получить оценку положения, применив один из вариантов метода ближайшего соседа. В статье применяется метод  $k$  ближайших соседей с  $k = 4$ . Для трех из четырех зданий применение такого метода с полной радиокартой (т.е. оригинальные данные отпечатков, содержащиеся в радиокarte) дало лучшие результаты, чем применение его же с радиокартой, построенной по карте, содержащей только параметры ПНР, но разница не очень значительная. По результатам испытаний, наилучшим способом вычисления параметров модели ПНР на этапе оффлайн является метод минимальной среднеквадратической ошибки. Более того, модель ПНР (8) обеспечивает лучший компромисс между точностью позиционирования и вычислительной сложностью.

Опять же, было бы разумно пропустить этап реконструкции радиокарты и просто использовать параметры ПНР и измеренные значения RSS для расчета оценок дальности, а затем применить какую-либо нелинейную методику оценки. Однако при применении многотактной модели ПНР сложно найти правильную оценку экспонента ПНР.

Можно еще более повысить точность позиционирования на основе модели ПНР путем замены изотропных моделей, рассматриваемых до настоящего момента, на анизотропные. Хорошо известно, что характеристики канала в разных направлениях от узла сети различны даже для всенаправленных антенн вследствие меняющихся условий (напр., точка доступа WLAN установлена в углу комнаты). Более того, на практике широко используются направленные антенны, т.к. они снижают помехи [14], обеспечивают большую пропускную способность данных и увеличение диапазона [61]. Т.о., применение различных параметров модели ПНР для разных направлений может дать преимущество в точ-



ности. Однако опять это преимущество стоит увеличения базы данных. Сохранение параметров модели ПНР для двух направлений уже удваивает размер радиокарты. Кроме того, оно усложняет этап позиционирования, т.к. необходимо решать, какой параметр использовать, или вычислять (взвешенное) среднее параметров.

### V. Модель гауссовских сумм<sup>3</sup> и аналогичные методы

Известным недостатком метода оценивания области покрытия, рассмотренного в разделе III, является то, что основная часть вероятностей расположена вблизи центра эллипса, описывающего область покрытия узла сети. Однако при слабых сигналах более вероятно, что ПУ находится ближе к краю области покрытия. Т.о., использование областей покрытия в таких случаях дает плохие оценки при позиционировании [20]. В предыдущем разделе мы рассмотрели методы, использующие данные RSS в дополнение к идентификаторам узлов сети путем применения моделей ПНР. Альтернативный вариант – применение моделей смеси нормальных распределений – модель гауссовских сумм (МГС).

Под МГС понимается выпуклая комбинация функций гауссовских плотностей распределения вероятности  $N(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ :

$$p(\mathbf{x}) = \sum_{n=1}^N \omega_n N(\mathbf{x}; \boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n) \quad (9)$$

где веса  $\omega_n$  неотрицательны и в сумме дают 1. Смысл МГС и фильтров на ее основе состоит в том, что любую функцию плотности можно произвольно близко аппроксимировать (кроме разрывов) выпуклой комбинацией гауссовских плотностей [62–64]. По мере роста числа гауссовских компонентов в смеси и приближения к нулю нормы всех ковариационных матриц аппроксимирующая функция плотности равномерно сходится к желаемой функции плотности ([62], Lo (1969) и Alspach (1970) [65]). В отличие от других методик аппроксимации, напр. разложение Грама-Шалье и Эджворта, МГС сама по себе является достоверной функцией плотности [62, 65].

В работе Sorenson и Alspach [62] указывается, что качество аппроксимации зависит не только от количества компонентов смеси, но и от их расположения. Предлагается выбирать их так, чтобы минимизировать  $L^k$ -норму ошибки аппроксимации или чтобы аппроксимация смеси точно соответствовала каким-то моментам реальной плотности. Дополнительно отмечается, что назначение одинаковой ковариации всем компонентам значительно упрощает вычислительную нагрузку. Однако во многих случаях последнее приводит к неудовлетворительной аппроксимации.

#### A. Представление данных отпечатков с помощью гауссовских сумм

В работе Kaji и Kawaguchi [44] предлагается более сложный метод сокращения размера радиокарты, чем подход на основе одной гауссовской области покрытия [20, 29], который использует мощности сигналов. Авторы предлагают распределение RSS узла сети в виде модели МГС. Хотя при этом подходе обычно увеличивается количество данных, содержащихся в радиокarte, по сравнению с методом оценки области покрытия (раздел III), оно все равно сильно уступает объему данных в традиционных базах данных отпечатков.

<sup>3</sup> В отечественной литературе эти модели получили наименование полигауссовских см. [86] (прим. ред.).

В предлагаемом алгоритме собранные данные отпечатков сначала преобразуются в точечное распределение, в котором значения плотности в точке зависят от мощности сигнала, принятого в отпечатке (чем выше значение RSS или RSSI, тем выше значение плотности в точке). Затем параметры модели МГС, а именно средние значения  $\{\mu_n\}_{n=1}^N$ , матрицы ковариации  $\{\Sigma_n\}_{n=1}^N$  и веса компонентов  $\{\omega_n\}_{n=1}^N$ , оптимизируются по методу максимизации ожидания [41]. Авторы Kaji и Kawaguchi отмечают, что их подход позволяет обновлять модели МГС по мере появления новых отпечатков. В работе не дается уравнения или правила определения количества компонентов  $N$ . В наших испытаниях в разделе VI мы используем  $N = \max(\lfloor K/40 \rfloor, 8)$ , где  $K$  – количество отпечатков, где воспринимаются сигналы определенного узла.

В работе Koshizen [66] предложен другой подход, разработанный для локализации роботов. Это байесовский подход на основе МГС с максимизацией регуляризованного ожидания (GM Bayes' with Regularized Expectation Maximization – GMB-REM). На этапе оффлайн отпечатки, содержащие измерения датчиков, собираются в разных местах сетки. Затем для каждой точки  $\mathbf{x}$  сетки рассчитывается условная функция плотности измерений датчиков  $\mathbf{y}$  по  $\mathbf{x}$  как регуляризованная модель МГС. Параметры компонентов выбираются с использованием метода максимизации ожидания так, чтобы они максимизировали логарифмическую функцию правдоподобия, если имеются данные обучения (в режиме оффлайн). На этапе позиционирования функция правдоподобия для каждой точки  $\mathbf{x}$  умножается на значение априорной ФПРВ и затем значения перенормируются для получения апостериорного значения. В работах [66–68] авторы предлагают другие методики для выбора датчиков и новой системы комплексов датчиков для метода GMB-REM. Однако следует отметить, что данному алгоритму свойственны типичные недостатки баз данных отпечатков, описанные в разделе I, и его широкое применение нерационально.

#### *Б. Позиционирование с использованием моделей гауссовских сумм*

В задачах позиционирования система (2) зачастую является негауссовской и/или в значительной степени нелинейной (см. напр. [69] о критерии нелинейности). В этом случае байесовская рекурсивная формула обычно неразрешима в аналитическом виде [62]. Применение обобщенного фильтра Калмана для решения таких обыкновенно мультимодальных систем имеет тот недостаток, что в этом случае отслеживается пиковое значение ФПРВ, что означает, что вырабатывается скорее оценка максимального правдоподобия, а не оценка минимальной дисперсии [65]. Кроме вычислительно сложных методов, напр. фильтр частиц (см. напр. [70], 87), для обеспечения высококачественной работы алгоритмов нелинейных и/или негауссовских систем, можно также применять фильтры на основе МГС.

В [65] сообщается, что еще в 1965 г. Aoki предложил аппроксимировать апостериорную ФПРВ как МГС. В работе Ali-Löytty [71] описывается аппроксимация, которая называется «эффективный МГС фильтр» (EGMF). В работе Ali-Löytty показано, что предлагаемый метод, в отличие от большинства фильтров МГС, дает оптимальные результаты в части среднего и ковариации в линейном случае. Также отмечается, что метод EGMF обеспечивает большую точность, чем традиционные фильтры калмановского типа (обобщенный, ансамбленный и проч) и сигма-точечные фильтры (СТФ, англ. SPGMF) [72], т.к. требуют наличия меньшего количества компонентов, чем SPGMF. Если априорная ФПРВ соответствует гауссовскому распределению, то количество компонентов

в методе EGMF соответствует размерности переменной состояния  $x$ . Один из недостатков SPGMF состоит в том, что для него требуется аналитическое дифференцирование. Чтобы его избежать, в работе Raitoharju и Ali-Löytty [73] предлагается метод адаптивного разделения. В этом методе, сначала находится направление максимальной нелинейности в гауссовской априорной ФПРВ. Если нелинейность измерения представляется значительной по критерию, предложенному Ali-Löytty и Sirola [72], то гауссовский компонент, описывающий измерение, разделяется на смесь двух гауссовских компонентов. Разделение продолжается до тех пор, пока ни один компонент не прекратит проявлять значительную нелинейность. Результаты в [73] говорят о том, что для предлагаемого метода требуется меньше компонентов, чем для сигма-точечного фильтра, в то же время обеспечивается лучшая аппроксимация опорной ФПРВ.

Другой важный момент заключается в том, что допущение о гауссовском распределении априорной ФПРВ не всегда возможно. Очевидно, что для апостериорного значения в момент  $k$ , описываемого СТФ, априорная ФПРВ в момент  $k + 1$  также должна описываться СТФ. В общем, если начальное состояние негауссовское, и/или при сильной нелинейности для решения задачи следует использовать банк гауссовских фильтров, а именно СТФ фильтр с использованием МГС [72]. В работах Anderson и Moore [74, стр. 212] и Ali-Löytty [75] показано, что СТФ аппроксимация априорной ФПРВ сходится к реальной априорной ФПРВ по мере уменьшения его ковариации, СТФ аппроксимация апостериорной плотности также сходится к истинной апостериорной плотности. В работе Lo [64] представлено применение СТФ для фильтрации системы с линейной динамикой и произвольно распределенной априорной ФПРВ, что прекрасно иллюстрирует концепцию.

Следует стараться ограничить количество компонентов СТФ, об этом уже говорилось ранее при введении этого понятия [62, 65]. В работе Sorenson и Alspach [62] предлагается объединять компоненты с приблизительно одинаковым средним и ковариацией или опускать компоненты достаточно малого веса (т.н. «исключение»). В работах [69, 76, 77] представлены альтернативные методы. Более подробный и широкий обзор методов для сокращения количества компонентов см. в работах [78, 79] и ссылках, содержащихся в них.

Кроме снижения количества гауссовских компонентов, было бы желательно иметь малое количество компонентов уже на этапе аппроксимации. В работе [80] Müller и др. предлагается обобщенная версия МГС (GGM), которая снимает требования к неотрицательности весов компонентов – СТФ, допускающая отрицательные веса. Для изотропной модели дальности, рассмотренной в [80], функция правдоподобия для измерений имеет кольцевую форму, как показано на левом графике на Рис. 2. При традиционной МГС требуется большое количество компонентов вследствие бесконечного количества пиков подобия, в то

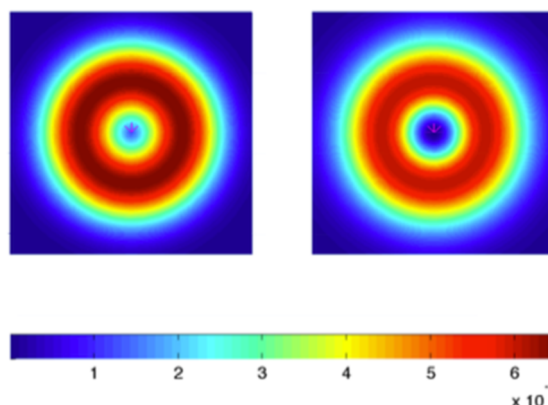


Рис. 2. Нормализованное точное подобие (слева) для измерения изотропного узла (обозначен звездочкой в круге) и его аппроксимация GGM (справа).

Том 24 № 1 (92), 2016 21

время как GGM обеспечивает удовлетворительную аппроксимацию при наличии всего двух компонентов (см. правый график на Рис. 2), один с положительным, а другой с отрицательным весом. Следует быть особо внимательным при назначении весов компонентов, чтобы обеспечить, что результирующее подобие остается неотрицательным, а значит и достоверным. В фильтре на основе GGM авторы сворачивают апостериорную МГС на каждом временном этапе, используя совпадение моментов, т.к. ранее указанные методы сокращения обычно подходят только для МГС с неотрицательными весами компонентов.

В работе Müller и др. метод GGM применяется в [80] для позиционирования в сотовых сетях, и результаты показывают, что GGM превосходит однокто-вый обобщенный фильтр Калмана и гауссовский метод оценки области покрытия [20, 29] в плане точности и последовательности. Отфильтрованная версия GGM (GGMF) также работает лучше, чем обобщенный фильтр Калмана и метод фильтрации на основе оценки области покрытия. В работе [81] эти выводы подтверждаются для сверхширокополосных сетей.

В работе Ali-Löyty и Sirola [69] проводится моделирование МГС при интегрировании сотовых измерений и измерений ГНСС. Согласно результатам моделирования, только мультимодальные функции правдоподобия должны аппроксимироваться МГС.

#### *В. Альтернативы модели гауссовских сумм*

Все методы, рассмотренные в разделе до настоящего момента, могли сильно зависеть от ранее упомянутой чувствительности гауссовской регрессии к выбросам. В работе Bishop и Svensen [82] указывается, что эта чувствительность могла привести к переоценке количества требуемых гауссовских компонентов. Авторы предлагают байесовский метод моделирования смеси на основе распределения студента, который более устойчив к негауссовости данных (в работе McLachlan и Peel [83] делается такое же предположение). Основным недостатком применения распределения студента состоит в том, что в отличие от применения компонентов с гауссовским распределением, не существует аналитического решения для максимизации подобия [82, 83]. Однако, как показано в [37, 82], любое распределение Стюдента может быть представлено как бесконечная смесь масштабированных гауссовских распределений. Т.о., для нахождения максимального правдоподобия можно использовать метод максимизации ожидания, в то время как вычислительная нагрузка предлагаемого алгоритма [82] будет незначительно выше, чем при использовании максимизации ожидания для нахождения параметров модели МГС. Однако до настоящего момента смеси с использованием распределения Стюдента не использовались в МПФП.

## **VI. Сравнительные испытания**

В этом разделе сравнивается качество работы нескольких параметрических МПФП методов и методов позиционирования, описанных в предыдущих разделах. Мы провели оценку этих методов путем анализа точности позиционирования по WLAN в ходе шести тестовых маршрутов по двум зданиям Технологического института Тампере. Первое здание имеет площадь ок. 10 000 м<sup>2</sup>, второе – ок. 6 600 м<sup>2</sup>; оба здания трехэтажные. Общее количество обнаруженных точек доступа в обоих зданиях – 506. Для двух маршрутов измерения были собраны несколько месяцев спустя после сбора измерений для первых четырех и одно-временного сбора данных для построения радиокарт. Некоторые из маршрутов включали переходы с этажа на этаж, которые полагаются известными. Радиокарты строились отдельно для каждого этажа. В таблице 1 для каждого этажа

двух зданий приведено количество обнаруженных точек доступа, отпечатков и контрольных точек четырех маршрутов, собранных одновременно с данными для радиокарт. Контрольные точки тестовых маршрутов – это точки, координаты которых определялись в процессе оценивания.

Т а б л и ц а 1

Размеры наборов данных. Некоторые точки доступа могли быть восприняты на нескольких этажах и/или в обоих зданиях

Здание	Этаж	Точки доступа	Отпечатки	Контрольные точки
1	1	200	889	19
1	2	289	243	47
1	3	212	160	22
2	1	154	1530	168
2	2	186	1582	33
2	3	148	333	19

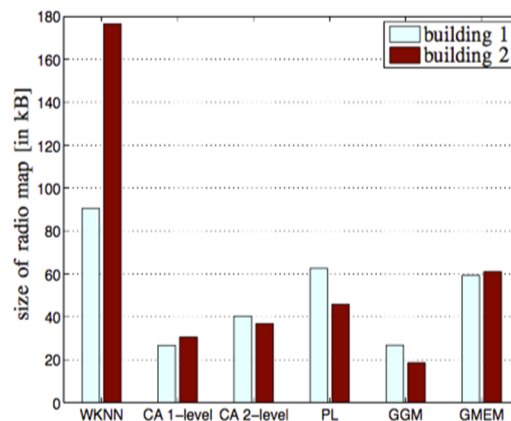


Рис. 3. Требования к объему данных радиокарты для испытываемых методов в двух тестовых зданиях. Объем данных радиокарты (в кб), здание 1,2, Метод оценивания областей покрытия (одноуровневый и двухуровневый), ПНР, СТФ

Для сравнения мы провели позиционирование на основе оценивания области покрытия одноуровневым [29] и двухуровневым методом области покрытия с ограничением -70 дБм [39], моделью ПНР [50], аппроксимацией модели ПНР методом СТФ [80] и моделью оценки мощности сигнала из работы [44] (GMEM). В дополнение к этим параметрическим методам мы использовали метод WKNN с  $k = 5$  в качестве опорного. На Рис. 3 показано, какой объем данных требуется каждому методу для построения радиокарты. В методе WKNN отпечатки никаким образом не суммируются, поэтому требуются радиокарты максимального размера для обоих зданий. В наших испытаниях применение параметрических методов сокращают размер радиокарты до 30% – 90%. Однако, поскольку в методе WKNN размер радиокарты зависит от количества отпечатков, а в других методах – от количества точек доступа, эти цифры нельзя обобщить.

На Рис. 4 показаны контуры, содержащие 50% вероятностей всех исследуемых методов (одноуровневая область покрытия соответствует двухуровневой модели с измерением -75 дБм), за исключением WKNN. Для вычисления вероятностей по подобиям мы использовали прямоугольную единообразную априорную ФПРВ, которая покрывает все здание. Отпечатки, соответственно, разделяются на группы, вероятно содержащие и вероятно не содержащие отпечатки с 50% вероятностей. Стандартное отклонение для методов на основе RSS

задано на уровне 6 дБ. Визуально формы контуров совершенно разные, кроме форм для ПНР и GGMF, но если рассмотреть количество вероятных отпечатков внутри контуров, можно увидеть, что они похожи, кроме метода GMEM со слабой мощностью сигнала.

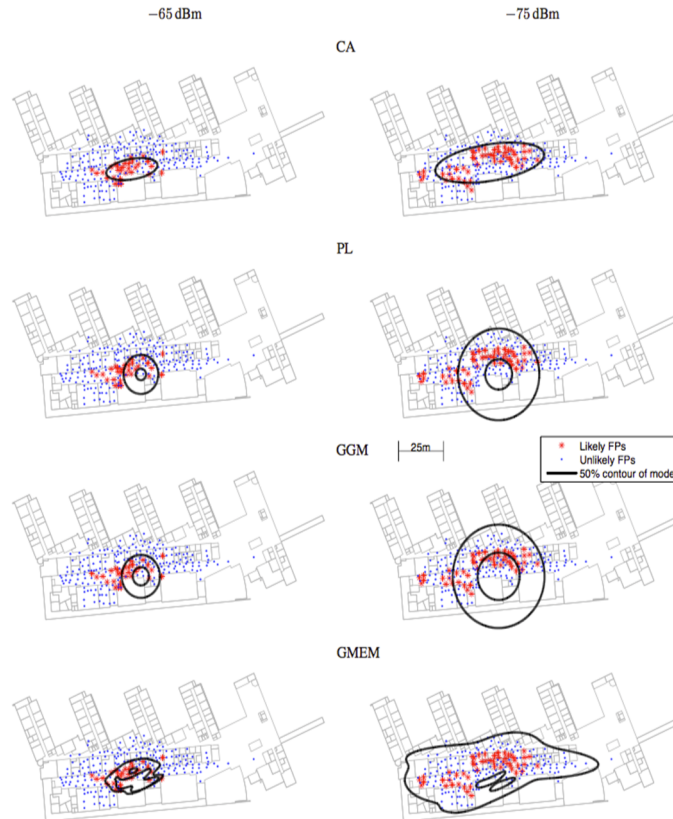


Рис. 4. Подобия разных моделей и значений RSS. Каждый вероятный отпечаток имеет большее количество вероятностей, чем каждый не вероятный.

Вероятно содержащие отпечатки  
Вероятно не содержащие отпечатки  
50% контур модели

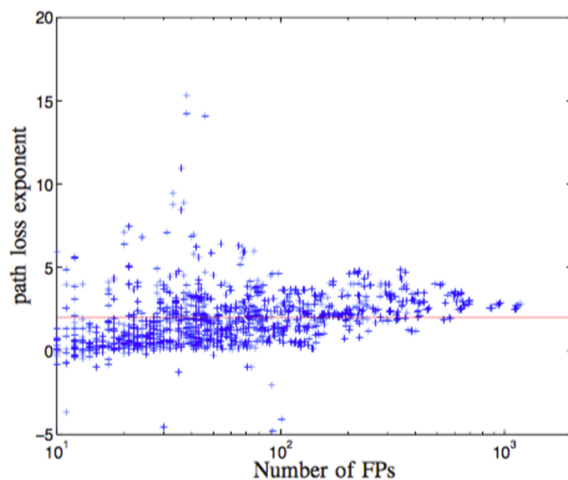


Рис. 5. Экспонент ПНР как функция отпечатков, используемых при обучении.  
Экспонент ПНР, кол-во отпечатков

На Рис. 5 для всех точек доступа показаны экспоненты ПНР, оцененные по методу Nurminen [50] как функция количества отпечатков, где наблюдалась конкретная точка доступа. Из рисунка видно, что при малом количестве отпечатков экспоненты ПНР часто принимают значение меньше 2 (68% всех точек доступа, полученные в менее чем 100 отпечатках, имеют  $n < 2$ , но только 27% точек доступа, полученных в более чем 100 отпечатках, имеют  $n < 2$ ). Экспонент ПНР, равный 2, означает, что сигнал

распространяется в открытом пространстве; если значение меньше 2, в наших испытаниях это может означать, что коридоры, где были собраны отпечатки, работали как волноводы [14, стр. 66].

Реальные траектории всех шести маршрутов измерялись путем нанесения отметок на карте на сенсорном экране при прохождении и интерполяции между этими отметками, и оценивались как для статического случая, так и для временного ряда (при наличии фильтрации). Для фильтрации мы рассмотрели вектор состояния  $x_k$ , содержащий положение и скорость ПУ. Обе модели с областями покрытия и метод GGMPF были доработаны с использованием обычного фильтра Калмана. Дополнительно, в методе GGMPF мы сократили количество компонентов до одного после 5 измерений и после каждого временного такта. В методе GMEM используется сетка для оценки статического положения и частичный фильтр с 300 частицами для оценки временного ряда; в модели ПНР используется алгоритм Гаусса-Ньютона для статического положения и фильтр ПНР [73] для временного ряда. Во временном ряде, эффект неопределенностей параметров менялся в зависимости от местоположения и поэтому был рассчитан по априорной средней оценке. В методе WKNN было задано стандартное отклонение 10 м для калмановской фильтрации.

Для фильтрации выбирается линейное уравнение перехода состояния (2а) с аддитивным центрированным шумом, т.е.  $x_k = F_{k-1}x_{k-1} + w_{k-1}$  при

$$F_{k-1} = \begin{bmatrix} I & \Delta t I \\ 0 & I \end{bmatrix}, \quad (10)$$

где  $\Delta t$  – интервал измерений в с, и  $w_{k-1} \sim N(0, Q)$  при

$$Q = 0.1m^2 \begin{bmatrix} \frac{(\Delta t)^3}{3} I & \frac{(\Delta t)^2}{2} I \\ \frac{(\Delta t)^2}{2} I & \Delta t I \end{bmatrix}. \quad (11)$$

Уравнение измерения (2b) зависит от используемого метода позиционирования. При статическом методе WKNN  $y_k$  содержал идентификаторы точек доступа и соответствующие значения RSS. Было выбрано  $k$  наиболее сильных идентификаторов и взвешенное среднее их положений было использовано как оценка положения. При фильтрации взвешенное среднее использовалось как апостериорное среднее для фильтра Калмана, а апостериорная ковариационная матрица была задана  $10^2 m^2 I$ .

В методах области покрытия  $j$ -е измерение  $y_{k,j}$  моделировалось как  $y_{k,j} = m_{ID_{k,j}} + \epsilon$ , где  $m_{ID_{k,j}}$  – среднее областей покрытия для точки доступа с идентификатором  $ID_{k,j}$ , а  $\epsilon$  – центрированное гауссовское значение с такой же ковариацией, как и модель области покрытия точки доступа.

В методе ПНР измерение  $y_k$  включало значения RSS. RSS для точки доступа с идентификатором AP  $ID_{k,j}$  моделировалось как

$$y_{k,j} = A_{ID_{k,j}} - 10n_{ID_{k,j}} \log_{10} \left( \|\mathbf{x}_k - \mathbf{a}_{ID_{k,j}}\| \right) + v_{k,j}, \quad (12)$$

где  $\|\cdot\|$  – евклидово расстояние между ПУ в точке  $\mathbf{x}_k$  и идентификатором точки доступа в  $\mathbf{a}_{ID_{k,j}}$ . Эта модель измерения используется для статического и фильтрационного позиционирования.

В методе GGM для вывода ковариационных матриц из гауссовских компонентов использовалась модель ПНР (8). Каждое измеренное значение RSS моделировалось GGM с двумя компонентами. В методе GMEM подобие измерения  $j$  моделировалось как

$$p(y_{k,j} | \mathbf{x}_k) = p_N(f(\mathbf{x}_k) | y_{k,j}, \sigma^2), \quad (13)$$

где  $p_N(f(\mathbf{x}_k) | y_{k,j}, \sigma^2)$  – ФПРВ  $N(y_{k,j}, \sigma^2)$ , оцененная в точке  $f(\mathbf{x}_k)$ . Функция  $f(\mathbf{x}_k)$  дает значения RSS и определяется как

$$f(\mathbf{x}_k) = \sum_n \omega_{n, ID_{k,j}} p_N(\mathbf{x}_k | \boldsymbol{\mu}_{n, ID_{k,j}}, \boldsymbol{\Sigma}_{n, ID_{k,j}}) - 90\text{dBm}, \quad (14)$$

где  $\boldsymbol{\mu}_{n, ID_{k,j}}$  и  $\boldsymbol{\Sigma}_{n, ID_{k,j}}$  – среднее и матрица ковариации  $n$ -го гауссовской компоненты МГС для точки доступа с идентификатором  $ID_{k,j}$ ,  $\omega_{n, ID_{k,j}}$  – вес компонента.

Испытания методов проходили по четырем сценариям:

- Рис. 6: полные данные
- Рис. 7: для позиционирования используются только точки доступа с пятью сильнейшими сигналами
- Рис. 8: 90% точек доступа было проигнорировано псевдослучайным образом, чтобы проверить, как работают методы при низкой плотности точек доступа
- Рис. 9: данные для построения радиокарт и для позиционирования собирались с интервалом в несколько месяцев, чтобы оценить ухудшение качества работы методов со временем.

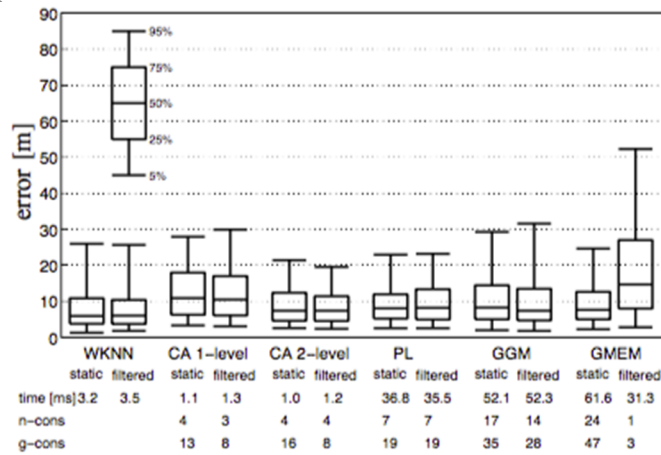


Рис. 6. Работа методов при использовании полного объема данных  
Статич/с фильтрацией, n-cons/g-cons – нормальная/общая устойчивость

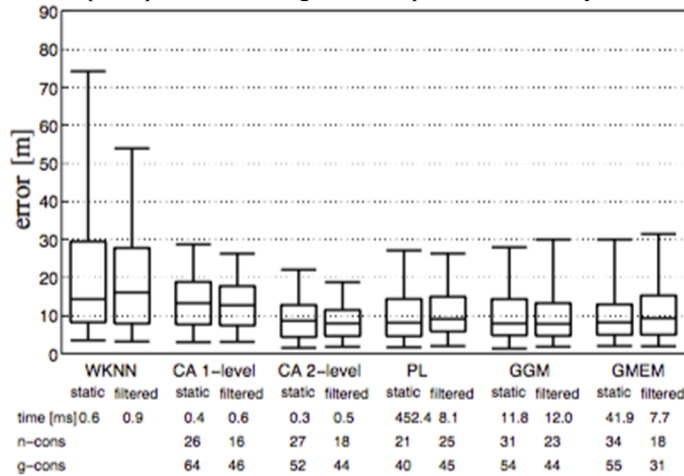


Рис. 7. Работа методов при использовании пяти сильнейших измерений



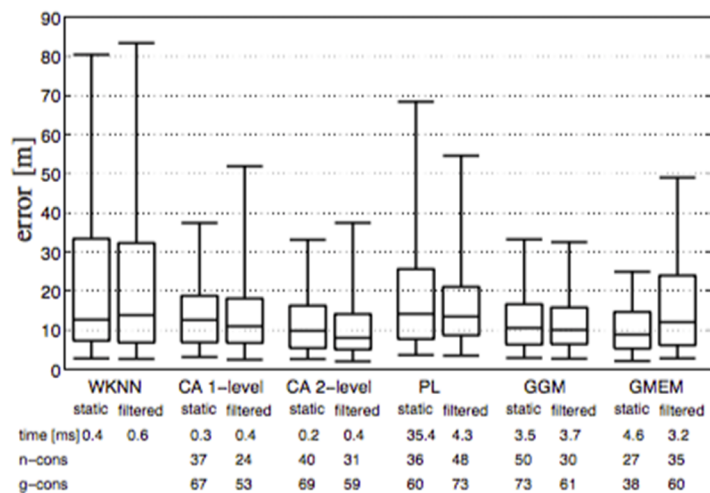


Рис. 8. Работа методов при опущении 90% точек доступа

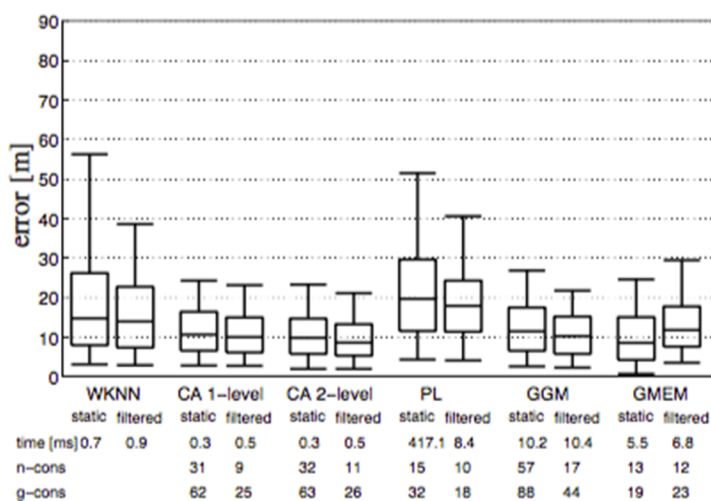


Рис. 9. Работа методов при проведении позиционирования и сбора данных для радиокарты с разрывом в несколько месяцев

На Рис. 6–9 представлены квантили с коробчатыми диаграммами, обозначающими ошибки позиционирования, абсолютное время одной оценки положения и показатели устойчивости, которые можно использовать для анализа точности оценки матрицы ковариации метода. В коробчатых диаграммах показаны квантили двухмерных ошибок позиционирования 5%, 25%, 50% (средняя линия), 75% и 95%. Для расчета параметра нормальной устойчивости [36, стр. 235] мы предположили гауссовское распределение ошибок позиционирования и вычислили, как часто ошибки попадали в 50% эллипс, т.е.,

$$(\hat{\mathbf{x}}_u - \mathbf{x}_u)^T \mathbf{P}_u^{-1} (\hat{\mathbf{x}}_u - \mathbf{x}_u) \leq \chi_2^2(0.5) = 1.3863, \quad (15)$$

Где  $\hat{\mathbf{x}}_u$  – оценка положения ПУ,  $\mathbf{P}_u$  – его матрица ковариации и  $\mathbf{x}_u$  – реальное положение. Этот параметр можно использовать для проверки оценки ошибки в обоих случаях (если она слишком мала или велика), если распределение близко к нормальному. Параметр общей устойчивости [84] подразумевает подсчет то-

го, насколько часто ошибки попадали в эллипс 50% при любом распределении при использовании модифицированного неравенства Чебышева:

$$(\hat{\mathbf{x}}_u - \mathbf{x}_u)^T \mathbf{P}_u^{-1} (\hat{\mathbf{x}}_u - \mathbf{x}_u) \leq \frac{2}{0.5} = 4. \quad (16)$$

При использовании полного набора данных все параметрические методы показали себя неустойчивыми, значения нормальной устойчивости были далеки от ожидаемых 50%, а общей устойчивости – от 50%, что может рассматриваться как минимальное требование (общая устойчивость на уровне 60% необязательно лучше/хуже, чем 55%), и все методы показали примерно одинаковую точность, кроме отфильтрованного метода GMEM. Время вычисления при статическом методе GMEM выше, чем при фильтрованном методе GMEM, т.к. расчет идет по сетке, а не с помощью парциального фильтра. Результаты показывают, что 300 частиц, предложенные в работе [44], это недостаточно. При использовании только пяти сильнейших измерений устойчивость всех методов повышалась и относительное время вычислений<sup>4</sup> сокращалось. Поскольку вычислительная сложность метода GGM экспоненциально зависит от количества измерений [81], сокращение времени вычислений для статического и фильтрованного методов GGM было ожидаемым, хотя в наших испытаниях зависимость не экспоненциальная вследствие свертки GGM после пяти измерений к одному гауссовскому распределению. В то же время точность позиционирования значительно снизилась только в методах WKNN и одноуровневом методе области покрытия.

Это может служить доказательством зависимости измерений. В здании, где проводились испытания, было несколько многомерных точек доступа, которые и привели к зависимости измерений.

На Рис. 8 видно, что более сложные методы (ПНР, GGM и WKNN) работают хуже или так же, как относительно простые методы области покрытия для сценариев с низкой плотностью точек доступа. То же самое относится к сценарию с устаревшей радиокарты (ср. Рис. 9).

Одна из возможных причин плохой работы статических и фильтрованных методов GGM во всех четырех сценариях (в сравнении с их работой в [80] и [81]) может состоять в том, что мы использовали другой метод определения матриц ковариации двух гауссовских компонентов GGM, т.к. наши испытания проводились в сети WLAN, а не в сотовой сети. Более подробный анализ метода GGM см. в [81].

### Заключение

В статье рассмотрены различные параметрические методы позиционирования на основе фингерпринтинга, анализируются их достоинства и недостатки. Мы провели испытания нескольких методов с реальными данными WLAN при различных тестовых маршрутах и сценариях, затем сравнили точность позиционирования и согласованность друг с другом и с методом WKNN (как пример непараметрического МПФП).

В таблице 2 сведены данные по рассмотренным параметрическим методам; какие параметры содержатся в радиокarte вместо данных отпечатков, как получают эти параметры и как они используются в позиционировании.

<sup>4</sup> Долгое время вычислений при статическом методе ПНР объясняется тем, что в двух (из 308) точек сходжение было крайне медленным и тем, что в нашей реализации мы не ограничивали число итераций.

## Параметрические МПФП, рассмотренные в статье

Метод	Кол-во точек радиокарты на узел сети	Способ построения радиокарты	Метод позиционирования
Область покрытия (гауссовское распределение) (1 уровень) [20, 29]	1 двумерное гауссовское распределение (2 параметра для среднего, 3 для ковариации)	Правило Байеса с использованием местоположений, где получен сигнал узла сети	Правило Байеса с использованием центров областей покрытия наблюдаемых узлов сети
Область покрытия (гауссовское распределение) (2 уровня) [39]	2 двумерных гауссовских распределения (4 параметра для среднего, 6 для ковариации)	Правило Байеса с использованием местоположений, где получен сигнал узла сети	Правило Байеса с использованием центров областей покрытия наблюдаемых узлов сети
Область покрытия (распределение Стьюдента) (1 уровень) [37]	1 двумерное стьюдентово распределение (2 параметра для среднего, 3 для формы, 1 для степени свободы)	Максимизация ожидания (для MAP) или выборка Гиббса (апостериорное) с использованием местоположений, где получен сигнал узла сети	Правило Байеса с использованием центров областей покрытия наблюдаемых узлов сети
Область покрытия (распределение Стьюдента) (2 уровня) [37]	2 двумерных стьюдентова распределения (4 параметра для среднего, 6 для формы, 2 для степени свободы)	Максимизация ожидания (для MAP) или семплирование Гиббса (апостериорное) с использованием местоположений, где получен сигнал узла сети	Правило Байеса с использованием центров областей покрытия наблюдаемых узлов сети
Модель потерь на распространение Nurminen и др. [50]	2 двумерных гауссовских распределения (1 для положения узла, 1 для параметров модели ПНР)	Итеративный метод взвешенных наименьших квадратов (IRLS)	Метод сеток с использованием стандартного метода Монте-Карло, Метрополиса-Гастингса или IRLS
Модель потерь на распространение Nap и др. [18]	8 параметров (2 для положения узла, 6 для модели ПНР)	Решение линейного уравнения методом наименьших квадратов	Восстановление полной радиокарты отпечатков и WKNN
Модель потерь на распространение Shrestha и др. [19]	3 параметра для положения узла (трехмерного), плюс 1) 2 параметра для модели ПНР (8) 2) 3 для модели ПНР (8) с параметром этажа 3) $M$ для многотактной модели ПНР порядка $M$ 4) $M+1$ для многотактной модели с параметром этажа	Минимизация среднеквадратической ошибки а) методом наименьших квадратов б) взвешенным методом наименьших квадратов в) минимизация среднеквадратической ошибки	Восстановление полной радиокарты отпечатков и KNN ( $k=4$ )
Общая смесь нормальных распределений [80, 81]	5 параметров (2 для узла сети, 2 для модели ПНР и 1 дополнительный)	Итеративный метод взвешенных наименьших квадратов (IRLS)	Правило Байеса с использованием центров областей покрытия наблюдаемых узлов сети и значения RSS
МГС Kajı и Kawaguchi [44]	$N$ двумерных гауссовских распределений для компонентов МГС	Преобразование данных отпечатков, параметры компонентов МГС, подобранных по методу максимизации ожидания	Статическая задача: метод сеток. Задача фильтрации: фильтр частиц
Байесовский подход на основе МГС с максимизацией регуляризованного ожидания (GMB-REM)	1 регуляризованная МГС для каждой точки сетки	Максимизация ожидания для нахождения максимального логарифмического правдоподобия данных отпечатков	Умножение подобия каждой точки сетки на априорную ФПРВ для получения апостериорной ФПРВ

Все предлагаемые методы позволяют значительно сократить объем радиокарты, используемой на этапе позиционирования в сравнении с непараметрическими методами. Также, как показали испытания, имеется возможность обновления радиокарт, используемых в методах области покрытия, модели ПНР, GGM и GMEM по мере появления новых отпечатков.

Испытания продемонстрировали, что все параметрические методы, кроме одноуровневого метода области покрытия и фильтрованного метода GMEM, обеспечивают примерно одинаковую точность позиционирования, как и непараметрический метод WKNN при большой плотности узлов сети и при использовании всех доступных измерений. Однако в методах модели ПНР, GGM и GMEM при этом значительно увеличивается время вычислений. При использовании только пяти сильнейших измерений время вычислений резко сокращается. Более того, все параметрические методы еще продолжают обеспечивать аналогичную точность позиционирования, в то время как качество работы метода WKNN серьезно падает. Это означает, что параметрическим методам требуется меньшее количество наблюдаемых узлов, чем непараметрическим, чтобы обеспечить удовлетворительную точность позиционирования. Если плотность узлов низкая, или данные карты устарели, более простой двухуровневый метод области покрытия обеспечивает такую же точность позиционирования, как более сложные параметрические методы и WKNN. Т.о., метод области покрытия обеспечивает наилучшее сочетание точности и вычислительной сложности. Другие параметрические методы, как например WKNN, более уязвимы в сложных условиях. Однако, по мнению авторов и в соответствии с работами, представленными в данной статье, методы модели ПНР и GGM могут превзойти по качеству работы методы области покрытия при условии необходимого обучения для специфических приложений, что было исключено из наших испытаний.

Важно отметить, что полученная точность позиционирования всех методов достаточна для многих практических приложений (прогноз погоды, реклама), но недостаточна для навигации. Для улучшения качества работы методов можно использовать данные карты и дополнительные измерения (от инерциальных измерительных модулей, измерения угла наклона траектории сигналов или временных задержек). По мнению авторов, использование карт этажей может повысить точность позиционирования, поскольку при этом, к примеру, будет запрещено пересечение стен. В работе Nurminen и др. [85] показывается, что применение карты этажей значительно повышает точность позиционирования с использованием фильтра частиц. Однако вопрос совмещения этих данных с некоторыми другими методами, представленными в статье, до сих пор остается открытым. Также, картографические данные можно использовать на этапе офлайн для построения более точной радиокарты.

### **Благодарности**

Работа Ф. Мюллера осуществлялась при финансовой поддержке Программы докторантуры Университета Тампере, работа М. Райтохарью – при поддержке Программы докторантуры Тампере по информационной науке и технологии, Nokia Inc., Nokia Foundation и Фонда Дженни и Антти Вихури. Данная работа частично финансировалась компанией HERE.

Авторы выражают благодарность Г. Нурминену за советы и предоставление Matlab кода для метода [50].

Авторы отрицают наличие конфликта интересов в связи с публикацией данной статьи.

ЛИТЕРАТУРА

1. **S. Gezici, Z. Tian, G. B. Biannakis, H. Kobayashi, A. F. Molisch, H. V. Poor, and Z. Sahinoglu**, "Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 70–84, July 2005.
2. **F. Gustafsson and F. Gunnarsson**, "Mobile positioning using wireless networks: possibilities and fundamental limitations based on available wireless network measurements," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 41–53, July 2005.
3. **E. Trevisani and A. Vitaletti**, "Cell-id location technique, limits and benefits: an experimental study," in *Proceedings of Sixth IEEE Workshop on Mobile Computing Systems and Applications (WMCSA 2004)*. Windermere, Cumbria, United Kingdom, December 2004.
4. **L. Wirola**, "Studies on location technology standards evolution in wireless networks," Ph.D. dissertation, Tampere University of Technology, March 2010. [Online]. Available: <http://URN.fi/URN:NBN:fi:tty-201002121065>
5. **D. Dardari, A. Conti, U. Ferner, A. Giorgetti, and M. Z. Win**, "Ranging with ultrawide bandwidth signals in multipath environments," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 404–426, February 2009.
6. **R. Mautz**, "Indoor positioning technologies," Habilitation Thesis, <http://e-collection.library.ethz.ch/eserv/eth:5659/eth-5659-01.pdf>, February 2012.
7. **H. Liu, H. Darabi, P. Banerjee, and J. Liu**, "Survey of wireless indoor positioning techniques and systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 37, no. 6, pp. 1067–1080, November 2007.
8. **M. I. Silventoinen and T. Rantalainen**, "Mobile station emergency locating in GSM," in *Proceedings of IEEE International Conference on Personal Wireless Communications*. New Delhi, India, February 1996, pp. 232–238.
9. **I. Constandache, S. Gaonkar, M. Sayler, R. R. Choudhury, and L. Cox**, "EnLoc: Energy-efficient localization for mobile phones," in *2009 Proceedings IEEE INFOCOM*. Rio de Janeiro, Brazil, April 2009, pp. 2716–2720.
10. **M. A. Spirito, S. Pöykkö, and O. Knuutila**, "Experimental performance of methods to estimate the location of legacy handsets in GSM," in *IEEE VTS 54th Vehicular Technology Conference (VTC)*, 2001. Atlantic City, NJ, USA, October 2001, pp. 2716–2720.
11. **J. A. del Peral-Rosado, J. A. López-Salcedo, G. Seco-Granados, F. Zanier, and M. Crisci**, "Analysis of positioning capabilities of 3GPP LTE," in *Proceedings of 25th International Technical Meeting of the Satellite Division of ION*, Nashville, TN, USA, September 2012.
12. **L. Wirola, T. A. Laine, and J. Syrjärinne**, "Mass-market requirements for indoor positioning and indoor navigation," in *2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. Zurich, Switzerland, September 2010.
13. **J. Park, D. Curtis, S. Teller, and J. Ledlie**, "Implications of device diversity for organic localization," in *2011 Proceedings IEEE INFOCOM*, April 2011, pp. 3182–3190.
14. **A. F. Molisch**, *Wireless Communications*, 2nd ed. Wiley - IEEE, January 2011.
15. **V. Honkavirta, T. Perälä, S. Ali-Löyty, and R. Piché**, "A comparative survey of WLAN location fingerprinting methods," in *Proceedings of the 6th Workshop on Positioning, Navigation and Communication 2009 (WPNC'09)*. Hannover, Germany, March 2009, pp. 243–251.
16. **C. Figuera, I. Mora-Jiménez, A. Guerrero-Curieses, J. L. Rojo-Álvarez, E. Everess, M. Wilby, and J. Ramos-López**, "Nonparametric model comparison and uncertainty evaluation for signal strength indoor location," *IEEE Transactions on Mobile Computing*, vol. 8, no. 9, pp. 1250–1264, September 2009.
17. **P. Müller, M. Raitoharju, and R. Piché**, "A field test of parametric WLAN-fingerprint-positioning methods," in *Proceedings of the 17th International Conference on Information Fusion (Fusion 2014)*. Salamanca, Spain, July 2014, pp. 1 – 8.
18. **S. Han, Z. Gong, W. Meng, and C. Li**, "An indoor radio propagation model considering angles for WLAN infrastructures," *Wireless Communications and Mobile Computing*, 2015 (in press).
19. **S. Shrestha, J. Talvitie, and E. S. Lohan**, "Deconvolution-based indoor localization with WLAN signals and unknown access point locations," in *2013 International Conference on Localization and GNSS (ICL-GNSS)*. Torino, Italy, June 2013, pp. 1–6.
20. **L. Koski, R. Piché, V. Kaseva, S. Ali-Löyty, and M. Hännikäinen**, "Positioning with coverage area estimates generated from location fingerprints," in *Proceedings of the 7th Workshop on Positioning, Navigation and Communication 2010 (WPNC'10)*. Dresden, Germany, March 2010, pp. 99–106.

21. **J. Machaj, R. Piché, and P. Brida**, “Rank based fingerprinting algorithm for indoor positioning,” in 2011 International Conference on Indoor Positioning and Indoor Navigation (IPIN). Guimaraes, Portugal, September 2011, pp. 6–11.
22. **L. Wirola, L. Wirola, and R. Piché**, “Bandwidth and storage reduction of radio maps for offline WLAN positioning,” in 2013 International Conference on Indoor Positioning and Indoor Navigation (IPIN). Montbeliard-Belfort, France, October 2013.
23. **S. Eisa, J. Peixoto, F. Meneses, and A. Moreira**, “Removing useless APs and fingerprints from WiFi indoor positioning radio maps,” in 2013 International Conference on Indoor Positioning and Indoor Navigation (IPIN). Montbeliard-Belfort, France, 2013, pp. 739–745.
24. **M. Raitoharju, T. Fadjukoff, S. Ali-Löytty, and R. Piché**, “Using unlocated fingerprints in generation of WLAN maps for indoor positioning,” in Proceedings of PLANS 2012 IEEE/ION Position Location and Navigation Symposium. Myrtle Beach, SC, USA, April 2012, pp. 576–583.
25. **L. Wirola, 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.
26. **K. Kaemarungsi and P. Krishnamurthy**, “Properties of indoor received signal strength for WLAN location fingerprinting,” in The First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MOBIQUITOUS 2004). Boston, MA, USA, August 2004, pp. 14–23.
27. **T. Vaupel, J. Seitz, F. Kiefer, S. Haimerl, and J. Thielecke**, “Wi-Fi positioning: System considerations and device calibration,” in 2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN). Zurich, Switzerland, September 2010.
28. **C. Laoudias, R. Piché, and C. G. Panayiotou**, “Device signal strength self-calibration using histograms,” in 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN). Sydney, NSW, Australia, November 2012.
29. **L. Koski, T. Perälä, and R. Piché**, “Indoor positioning using WLAN coverage area estimates,” in 2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN). Zurich, Switzerland, September 2010.
30. **A. Haeberlen, E. Flannery, A. M. Ladd, A. Rudys, D. S. Wallach, and L. E. Kavraki**, “Practical robust localization over large-scale 802.11 wireless networks,” in MobileCom’04. Philadelphia, PA, USA, September–October 2004.
31. **M. B. Kjaergaard**, “Indoor location fingerprinting with heterogeneous clients,” *Pervasive and Mobile Computing*, vol. 7, no. 1, pp. 31–43, 2011.
32. **Y.-C. Cheng, Y. Chawathe, A. LaMarca, and J. Krumm**, “Accuracy characterization for metropolitan-scale Wi-Fi localization,” in Proceedings of the 3rd International Conference on Mobile Systems, Applications and Services (MobiSys 2005). Seattle, WA, USA, June 2005.
33. **A. M. Hossain and W.-S. Soh**, “Cramér-Rao bound analysis of localization using signal strength difference as location fingerprint,” in 2010 Proceedings IEEE INFOCOM. San Diego, CA, USA, March 2010.
34. **A. H. Jazwinski**, *Stochastic Processes and Filtering Theory*, ser. Mathematics in Science and Engineering. Academic Press, 1970, vol. 64.
35. **S. Särkkä**, *Bayesian Filtering and Smoothing*. Cambridge University Press, 2013.
36. **Y. Bar-Shalom, R. X. Li, and T. Kirubarajan**, *Estimation with Applications to Tracking and Navigation, Theory Algorithms and Software*. John Wiley & Sons, 2001.
37. **R. Piché**, “Robust estimation of a reception region from location fingerprints,” in 2011 International Conference on Localization and GNSS (ICL-GNSS). Tampere, Finland, June 2011, pp. 31–35.
38. **L. Koski**, “Positioning with Bayesian coverage area estimates and location fingerprints,” Master’s thesis, Tampere University of Technology, [http://math.tut.fi/posgroup/koski\\_mscth.pdf](http://math.tut.fi/posgroup/koski_mscth.pdf), March 2010.
39. **M. Raitoharju, S. Ali-Löytty, R. Piché, and M. Dashti**, “Positioning with multilevel coverage area models,” in 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN). Sydney, NSW, Australia, November 2012.
40. **S. Geman and D. Geman**, “Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, no. 6, pp. 721–741, 1984.
41. **A. P. Dempster, N. M. Laird, and D. B. Rubin**, “Maximum likelihood from incomplete data via the EM algorithm,” *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.

42. **V. Honkavirta**, "Location fingerprinting methods in wireless local area networks," Master's thesis, Tampere University of Technology, <http://urn.fi/URN:NBN:fi:ty-200911137109>, November 2008.
43. **C. Laoudias, C. G. Panayiotou, and P. Kempfi**, "On the RBF-based positioning using WLAN signal strength fingerprints," in Proceedings of the 7th Workshop on Positioning, Navigation and Communication 2010 (WPNC'10). Dresden, Germany, March 2010, pp. 93–98.
44. **K. Kaji and N. Kawaguchi**, "Design and implementation of WiFi indoor localization based on Gaussian mixture model and particle filter," in 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN). Sydney, NSW, Australia, November 2012.
45. **H. Hashemi**, "The indoor radio propagation channel," Proceedings of the IEEE, vol. 81, no. 7, pp. 943–968, July 1993.
46. **N. Patwari, I. Alfred O. Hero, M. Perkins, N. S. Correal, and R. J. O'Dea**, "Relative location estimation in wireless sensor networks," IEEE Transactions on Signal Processing, vol. 51, no. 8, pp. 2137–2148, August 2003.
47. **X. Li**, "RSS-based location estimation with unknown pathloss model," IEEE Transactions on Wireless Communications, vol. 5, no. 12, pp. 3626–3633, December 2006.
48. **T. Roos, P. Myllymäki, and H. Tirri**, "A statistical modeling approach to location estimation," IEEE Transactions on Mobile Computing, vol. 1, no. 1, pp. 59–69, Jan-Mar 2002.
49. **S. M. Bo-Chieh Liu, K.-H. Lin, and J.-C. Wu**, "Analysis of hyperbolic and circular positioning algorithms using stationary signal-strength-difference measurements in wireless communications," IEEE Transactions on Vehicular Technology, vol. 55, no. 2, pp. 499–509, March 2006.
50. **H. Nurminen, J. Talvitie, S. Ali-Löytty, P. Müller, E. S. Lohan, R. Piché, and M. Renfors**, "Statistical path loss parameter estimation and positioning using RSS measurements in indoor wireless networks," in 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN). Sydney, NSW, Australia, November 2012.
51. **J. Salo, L. Vuokko, H. M. El-Sallabi, and P. Vainikainen**, "An additive model as a physical basis for shadow fading," IEEE Transactions on Vehicular Technology, vol. 56, no. 1, pp. 13–26, January 2007.
52. **V. Erceg, L. J. Greenstein, S. Y. Tjandra, S. R. Parkoff, A. Gupta, B. Kulic, A. A. Julius, and R. Bianchi**, "An empirically based path loss model for wireless channels in suburban environments," IEEE Journal on Selected Areas in Communications, vol. 17, no. 7, pp. 1205–1211, 1999.
53. **T. S. Rappaport, S. Y. Seidel, and K. Takamizawa**, "Statistical channel impulse response models for factory and open plan building radio communication system design," IEEE Transactions on Communications, vol. 39, no. 5, pp. 794–807, 1991.
54. **S. S. Ghassenizadeh, R. Jana, C. W. Rice, W. Turin, and V. Tarokh**, "A statistical path loss model for in-home UWB channels," in 2002 IEEE Conference on Ultra Wideband Systems and Technologies. Baltimore, MD, USA, May 2002, pp. 59–64.
55. **B. Dil and P. Havinga**, "RSS-based localization with different antenna orientations," in 2010 Australasian Telecommunication Networks and Applications Conference (ATNAC). Auckland, New Zealand, 2010, pp. 13–18.
56. **S. Shrestha, E. Laitinen, J. Talvitie, and E. S. Lohan**, "RSSI channel effects in cellular and WLAN positioning," in Proceedings of the 9th Workshop on Positioning, Navigation and Communication 2012 (WPNC'12). Dresden, Germany, March 2012, pp. 187–192.
57. **T. S. Rappaport**, Wireless Communications - Principles and Practice, 2nd ed. Prentice-Hall, 2010.
58. **S. Y. Seidel and T. S. Rappaport**, "914 MHz path loss prediction models for indoor wireless communications in multifloored buildings," IEEE Transactions on Antennas and Propagation, vol. 40, no. 2, pp. 207–217, February 1992.
59. **B. Dil and P. Havinga**, "On the calibration and performance of RSS-based localization methods," in Internet of Things 2010 Conference. Tokyo, Japan, November/December 2010.
60. **M. Hata**, "Empirical formula for propagation loss in land mobile radio services," IEEE Transactions on Vehicular Technology, vol. 29, no. 3, pp. 317–325, August 1980.
61. **K. Shujaee, A. Ebaid, R. George, and M. A. Sazegarnejad**, "Channel modeling and range extension for UWB communications using directional antenna in LOS and NLOS path loss models," in World Automation Congress (WAC 2008). Hawaii, HI, USA, September-October 2008.
62. **H. W. Sorenson and D. L. Alspach**, "Recursive Bayesian estimation using Gaussian sums," Automatica, vol. 7, no. 4, pp. 465–479, July 1971.
63. **S. Ali-Löytty**, "Gaussian mixture filters in hybrid positioning," Ph.D. dissertation, Tampere University of Technology, August 2009. [Online]. Available: <http://URN.fi/URN:NBN:fi:ty-200905191055>
64. **J. T.-H. Lo**, "Finite-dimensional sensor orbits and optimal nonlinear filtering," IEEE Transactions on Information Theory, vol. 18, no. 5, pp. 583–588, September 1972.

65. **D. L. Alspach and H. W. Sorenson**, “Nonlinear Bayesian estimation using Gaussian sum approximations,” *IEEE Transactions on Automatic Control*, vol. 17, no. 4, pp. 439–448, August 1972.
66. **T. Koshizen**, “The sensor selection task of the Gaussians mixture Bayes’ with regularized EM (GMB-REM) technique in robot position estimation,” in *Proceedings of the 1999 IEEE International Conference on Robotics & Automation*. Detroit, MI Detroit, MI, USA, May 1999, pp. 2620–2625.
67. **T. Koshizen**, “The evolved Gaussian mixture Bayes’ technique using sensor selection task integrated with sensor fusion scheme in mobile robot position estimation,” in *Proceedings of 1999 IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA ’99)*. Monterey, CA, USA, November 1999, pp. 202–207.
68. **T. Koshizen, P. Bartlett, and A. Zelinsky**, “Sensor fusion of odometry and sonar sensors by the Gaussian mixture Bayes’ technique in mobile robot position estimation,” in *IEEE SMC ’99 Conference Proceedings*, vol. 4. Tokyo, Japan, 1999, pp. 742–747.
69. **S. Ali-Löyty and N. Sirola**, “Gaussian mixture filter in hybrid navigation,” in *Proceedings of The European Navigation Conference GNSS 2007*. Geneva, Switzerland, May 2007, pp. 831–837.
70. **M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp**, “A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking,” *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.
71. **S. Ali-Löyty**, “Efficient Gaussian mixture filter for hybrid positioning,” in *2008 IEEE/ION Position Location and Navigation Symposium*, Monterey, California, May 2008, pp. 60–66.
72. **S. Ali-Löyty and N. Sirola**, “Gaussian mixture filter and hybrid positioning,” in *Proceedings of ION GNSS 2007*, Fort Worth, Texas, September 2007, pp. 562–570.
73. **M. Raitoharju and S. Ali-Löyty**, “An adaptive derivative free method for Bayesian posterior approximation,” *IEEE Signal Processing Letters*, vol. 19, no. 2, pp. 87–90, February 2012.
74. **B. D. O. Anderson and J. B. Moore**, *Optimal Filtering*, ser. Prentice-Hall information and system sciences. Prentice-Hall, 1979.
75. **S. Ali-Löyty**, “On the convergence of the Gaussian mixture filter,” *Tampere University of Technology, Institute of Mathematics*, <http://urn.fi/URN:NBN:fi:tyy-2011041510704>, Tech. Rep. Research report 89, 2008.
76. **D. J. Salmond**, “Mixture reduction algorithms for target tracking,” in *IEE Colloquium on State Estimation in Aerospace and Tracking Applications*. London, England, 1989, pp. 7/1–7/4.
77. **D. Schieferdecker and M. F. Huber**, “Gaussian mixture reduction via clustering,” in *Proceedings of the 12th International Conference on Information Fusion (FUSION ’09)*. Seattle, WA, USA, 2009, pp. 1536–1543.
78. **D. F. Crouse, P. Willett, K. Pattipati, and L. Svensson**, “A look at Gaussian mixture reduction algorithms,” in *Proceedings of the 14th International Conference on Information Fusion (FUSION ’11)*. Chicago, IL, USA, 2011.
79. **M. Kristan, D. Skočaj, and A. Leonardis**, “Incremental learning with Gaussian mixture models,” in *Computer Vision Winter Workshop 2008*. Moravske Toplice, Slovenia, February 2008, pp. 25–32.
80. **P. Müller, S. Ali-Löyty, M. Dashti, H. Nurminen, and R. Piché**, “Gaussian mixture filter allowing negative weights and its application to positioning using signal strength measurements,” in *Proceedings of the 9th Workshop on Positioning, Navigation and Communication 2012 (WPNC’12)*. Dresden, Germany, March 2012, pp. 71–76.
81. **P. Müller, H. Wymeersch, and R. Piché**, “UWB positioning with generalized Gaussian mixture filters,” *IEEE Transactions on Mobile Computing*, vol. 13, no. 10, pp. 2406–2414, October 2014.
82. **C. M. Bishop and M. Svensén**, “Robust Bayesian mixture modeling,” in *Proceedings of 12th European Symposium on Artificial Neural Networks (ESANN 2004)*. Bruges, Belgium, April 2004, pp. 69–74.
83. **G. J. McLachlan and D. Peel**, “Robust cluster analysis via mixtures of multivariate  $t$ -distributions,” in *Advances in Pattern Recognition*, ser. Lecture Notes in Computer Science, A. Amin, D. Dori, P. Pudil, and H. Freeman, Eds. Springer Berlin Heidelberg, 1998, vol. 1451, pp. 658–666.
84. **S. Ali-Löyty, N. Sirola, and R. Piché**, “Consistency of three Kalman filter extensions in hybrid navigation,” in *Proceedings of The European Navigation Conference GNSS 2005*. Munich, Germany, Jul. 2005.
85. **H. Nurminen, A. Ristimäki, S. Ali-Löyty, and R. Piché**, “Particle filter and smoother for indoor localization,” in *2013 International Conference on Indoor Positioning and Indoor Navigation (IPIN2013)*. Montbeliard-Belfort, France, October 2013, pp. 137–146.
86. **Степанов О. А.** *Применение теории нелинейной фильтрации в задачах обработки навигационной информации*. СПб., 2003.



87. Степанов О. А., Торопов А. Б. Применение последовательных методов Монте-Карло с использованием процедур аналитического интегрирования при обработке навигационной информации. XII Всероссийское совещание по проблемам управления ВСПУ-2014. Институт проблем управления им. В. А. Трапезникова РАН. М., 2014. С. 3324-3337.

**Müller P., Raitoharju M., Ali-Löyty S., Wirola L., Piché R.** (Tampere University of Technology, Tampere, Finland)

A survey of fingerprinting and parametric fingerprint-positioning methods // *Giroskopiya i Navigatsiya*. 2016, Vol.24, No. 1 (92), pp. 3-35.

**Abstract.** The term fingerprint-based (FP) positioning includes a wide variety of methods for determining a receiver's position using a database of radio signal strength measurements that were collected earlier at known locations. Nonparametric methods such as the weighted  $k$ -nearest neighbor (WKNN) method are infeasible for large-scale mobile device services because of the large data storage and transmission requirements. In this work we present an overview of parametric FP methods that use model-based representations of the survey data. We look at three different groups of parametric methods: methods that use coverage areas, methods that use path loss models, and methods that use Gaussian mixtures. Within each group we study different approaches and discuss their pros and cons. Furthermore, we test the positioning performance of several of the analyzed approaches in different scenarios using real-world WLAN indoor data and compare the results to those of the WKNN method.

**Key words:** fingerprint-based (FP) positioning, parametric FP methods, comparison.

Материал поступил 31.08.2015