

Н. В. МИХАЙЛОВ, А. В. НИКАНДРОВ, В. Ф. МИХАЙЛОВ

**ПРИМЕНЕНИЕ КЛАСТЕРНОГО АНАЛИЗА
ДЛЯ ИДЕНТИФИКАЦИИ И КОМПЕНСАЦИИ ВЛИЯНИЯ
ОТРАЖЕННОГО СИГНАЛА В ПРИЕМНИКАХ
СПУТНИКОВЫХ РАДИОНАВИГАЦИОННЫХ СИСТЕМ**

Проводится статистический анализ экспериментальных данных, полученных при работе навигационных приемников спутниковых радионавигационных систем (СРНС) в условиях многолучевости. В качестве исследуемых параметров выбраны отношение сигнал/шум сигнала и невязки измерения псевдодальности. Рассматривается возможность применения методов кластерного анализа для идентификации и компенсации отраженного луча. Предложена методика, повышающая точность местоопределения с использованием СРНС приемника. Приведены результаты экспериментальной проверки методики в статическом режиме.

Введение

На современном этапе развития СРНС на первый план выходят проблемы совершенствования функциональных характеристик навигационного приемника, имеющих наибольшую практическую ценность с точки зрения потребителя. К таким характеристикам несомненно относится точность местоопределения. При этом наибольший интерес представляет проблема ее повышения в условиях, неблагоприятных для приема сигналов навигационных космических аппаратов (КА), например в районах плотной городской застройки.

Как известно, в условиях городского каньона наибольшее влияние на характеристики точности СРНС приемника оказывает эффект деградации прямого сигнала (LOS, от англ. Line Of Sight), принимаемого от КА. Укажем две основные причины деградации. Прежде всего происходит подавление LOS-сигнала вплоть до полной невозможности его приема вследствие затенения городскими постройками. Сокращение по этой причине числа КА, от которых принимаются данные, означает увеличение геометрического фактора рабочего созвездия, что приводит к росту шумовой погрешности навигационного решения. Вторая причина – эффект многолучевости (т.е. вынужденный прием совокупности многократно переотраженных сигналов), который приводит к смещению в измерении псевдодальности до КА. Включение таких измерений в навигационное решение приводит к систематической ошибке местоопределения, изменяющейся со временем.

Эффект многолучевости проявляется обычно в двух основных вариантах.

Михайлов Николай Викторович. Кандидат физико-математических наук, глава представительства компании «Уайт Дварф Лимитед (С.-Петербург)». Действительный член Академии навигации и управления движением.

Никандров Александр Владимирович. Кандидат технических наук, ведущий специалист компании «Уайт Дварф Лимитед».

Михайлов Виктор Федорович. Доктор технических наук, профессор Санкт-Петербургского государственного университета аэрокосмического приборостроения.

Статья по докладу на XVII Санкт-Петербургской международной конференции по интегрированным навигационным системам.

Если одновременно с отраженным сигналом возможен прием и LOS-сигнала, то среднее смещение не столь велико и обычно не превышает десятка метров. Этот случай достаточно подробно рассмотрен в литературе, и существуют известные алгоритмы подавления такой многолучевости. К их числу относится применение стробового или узкого коррелятора на этапе обработки сигналов, а также методов контроля навигационного поля (RAIM, от англ. Receiver Autonomous Integrity Monitoring) на этапе решения навигационной задачи.

Другой случай возникает, когда LOS-сигнал затенен, т.е. подавлен до уровня шумов и не может быть выделен, а отраженный сигнал имеет мощность, сравнимую с мощностью ослабленного LOS-сигнала. Такой отраженный сигнал может вносить в измерения псевдодальности ошибку смещения порядка сотен метров. Идентификация такого сигнала как отраженного представляет собой нетривиальную проблему, поскольку его форма похожа на форму обычного LOS-сигнала. Кроме того, алгоритмы контроля навигационного поля могут быть неприменимы в условиях городского каньона из-за малого числа принимаемых сигналов.

Этот случай рассмотрен в литературе значительно менее подробно. Как правило, для адекватного описания отраженных сигналов используются многопараметрические статистические модели. При этом для оценки параметров сигнала используются методы максимального правдоподобия. Большинство таких алгоритмов приводится без раскрытия деталей, как, например, алгоритм АМТ (Anti-Multipath Triangulation) в [1].

Серьезным недостатком подобных методов является их высокая вычислительная сложность. Однако необходимо отметить, что постоянный технический прогресс привел к значительному повышению вычислительной мощности, доступной для СРНС приемников. При этом речь идет не только о мощных встраиваемых вычислителях, но и о возможности перенесения навигационной обработки на удаленные серверы, доступ к которым сильно упрощается с развитием беспроводных сетей. Все это позволяет применять более сложные алгоритмы, которые ранее даже не рассматривались по причине их высокой вычислительной сложности и/или значительного объема обрабатываемых данных. К их числу относится и метод, предлагаемый в настоящей статье.

Для описания отраженного сигнала была использована статистическая многофакторная модель, что позволяет проводить исследование экспериментальных данных методами кластерного анализа. В качестве основных факторов выбраны измерения отношения сигнал/шум (ОСШ) и невязки псевдодальности, что значительно упрощает модель и сводит ее к двумерной форме.

Поставим задачу разработки метода идентификации и компенсации отраженного сигнала, позволяющего увеличить точность навигационного решения в условиях сильной многолучевости.

Анализ экспериментальных данных

Для накопления экспериментальных данных были использованы два СРНС-приемника производства MStar Semiconductor и uBlox AG, работавших от одной антенны (использовалась антенна GPS-P1MAL фирмы ACC). Были выбраны статические условия с заранее известным местоположением, при этом антенна располагалась так, чтобы гарантировать постоянное присутствие и LOS-сигналов, и отраженных сигналов. Оба приемника действовали в режиме проведения измерений с известными координатами.

В качестве основных исследуемых параметров сигнала были выбраны невязки измерений псевдодальности и ОСШ для каждого принимаемого сигнала. Результаты обработки приведены на рис. 1–3. На всех рисунках ось ординат соответствует измерениям ОСШ в дБГц.

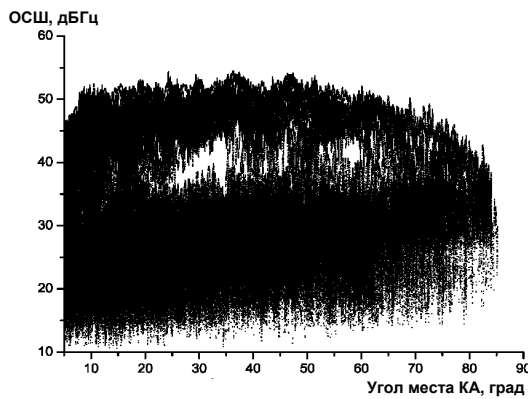
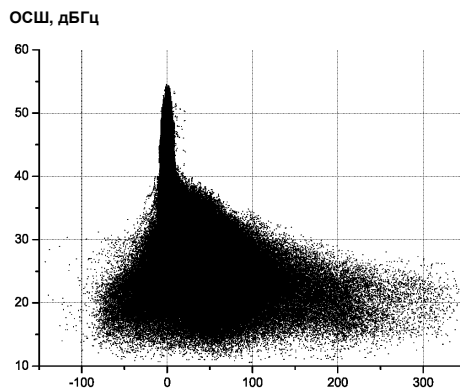


Рис. 1. Зависимость измеряемого ОСШ сигнала от угла места за 24 ч суммарно по всем видимым КА

На рис. 1 можно выделить две крупные области. Верхняя «дорожка» на уровнях 40–50 дБГц соответствует приему LOS-сигнала. Значительная ширина этой дорожки является следствием эффекта интерференционного сложения LOS-сигнала с отраженными сигналами от этого же спутника. Форма огибающей верхней «дорожки» соответствует диаграмме направленности стандартной серийной антенны, характеризующейся небольшим провалом в области зенитных углов.

Нижняя «дорожка» (на уровнях 15–35 дБГц) соответствует приему переотраженных (и поэтому ослабленных) сигналов в отсутствие LOS-сигнала. По визуальной ширине промежутка между дорожками можно оценить ослабление однократно отраженного сигнала величиной порядка 8–10 дБГц. Ширина нижней «дорожки» указывает, что современный СРНС приемник может принимать сигналы, переотраженные несколько раз. Так, в предположении, что каждое отражение ослабляет сигнал на 8 дБГц, возможные уровни отраженного сигнала могут составлять, например, 38, 30, 22 и 14 дБГц.

а)



б)

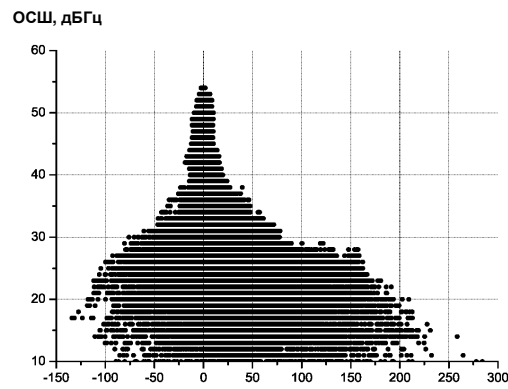


Рис. 2. Зависимость измеряемого ОСШ сигнала от невязки измерения псевдодальности суммарно за 24 ч по всем видимым КА :приемник MStar, (а), приемник uBloх-4, суммарно за 24 ч по всем видимым КА. Для обработки были доступны только целочисленные значения ОСШ (б)

Оба СРНС приемника показывают сравнимые результаты по невязкам (рис. 2). На графиках видны узкий высокий пик в области LOS-сигнала, а также длинный «хвост» в сотни метров для отраженных сигналов. Уровни сигналов для этих областей в целом соответствуют графику на рис. 1. Отметим, что появ-

ление отрицательных значений невязок (вплоть до минус 150 м) вызвано интерференцией слабого LOS-сигнала с более сильным и противофазным отраженным сигналом.

Итак, на рис. 2 можно предварительно выделить две структурно различимые области, которые соответствуют двум глобально различным условиям приема сигнала – наличию или отсутствию LOS. При этом форма проекции обеих областей на ось невязок существенно различается. Невязки для LOS-сигнала имеют незначительные среднее и дисперсию (единицы метров), в то время как невязки для отраженного сигнала при среднем и дисперсии в десятки метров имеют выбросы в сотни метров. Для отдельных спутников общая картина повторяется. На рис. 3 приведены суммарные графики для двух GPS спутников. Угол места для КА2 изменялся в диапазоне 5–51°, для КА10 – в диапазоне 5–55°.

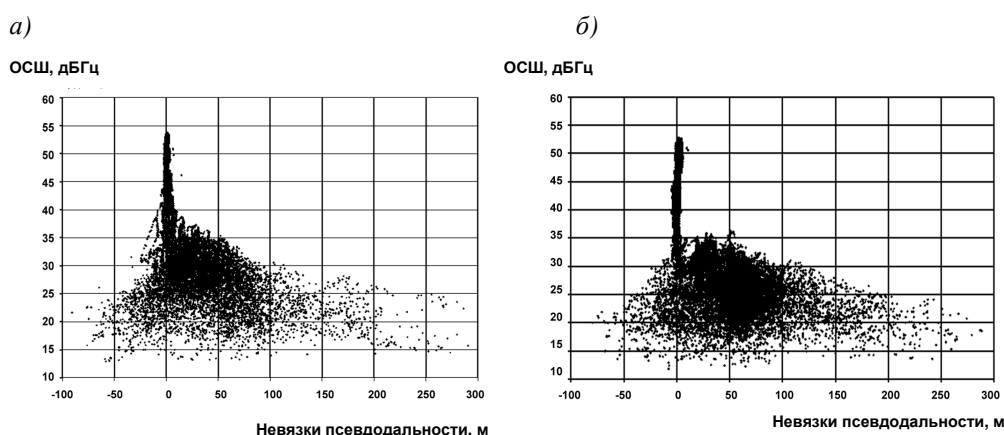


Рис. 4. Зависимость измеряемого ОСШ сигнала от невязки измерений псевдодальности суммарно за 24 ч для КА2 (а) и КА10 (б)

Увеличим масштаб графика на рис. 2, а внутри областей как LOS-сигнала, так и отраженного сигнала. В обоих случаях можно увидеть области сгущения точек (рис. 4). Это подтверждает возможность использования для решения нашей задачи методов кластерного анализа [2].

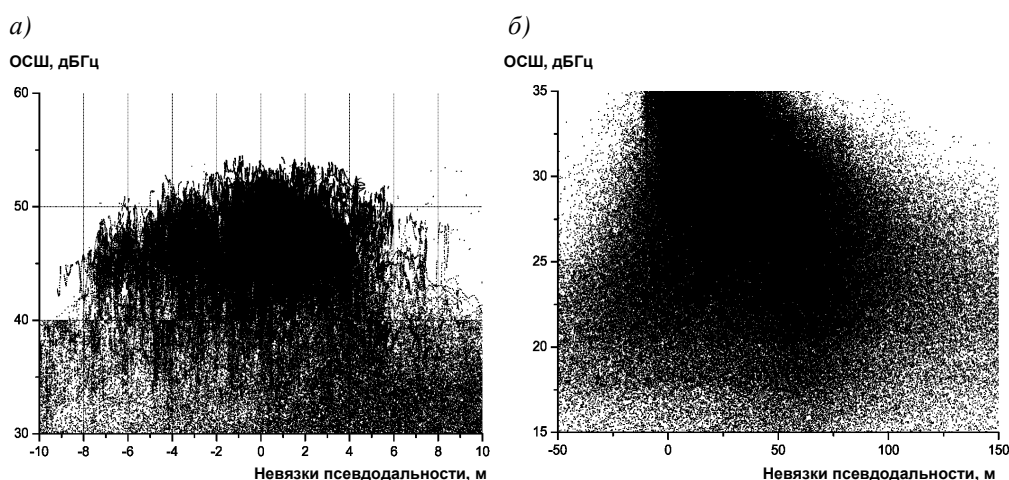


Рис. 4. Увеличение масштаба для LOS области (а) и для области отраженного сигнала (б)

В кластерном анализе каждый объект характеризуется некоторым набором параметров и может быть рассмотрен как точка в многомерном пространстве. Главной целью кластерного анализа является нахождение в таком пространстве отдельных групп (кластеров). Внутри каждого кластера объекты должны быть схожи между собой по некоторому критерию. В качестве такого критерия (или меры близости точек) в кластерном анализе применяется одна из метрик – линейное (или «манхеттенское») расстояние, евклидово расстояние, метрика Минковского, расстояние Махаланобиса. К наиболее важным свойствам кластера относятся его плотность, размер и отличимость от других кластеров. Наиболее распространенным показателем, характеризующим плотность кластера, является дисперсия расстояния от центра кластера до его отдельных точек. Чем меньше дисперсия, тем ближе к центру кластера находятся его точки. Важным свойством кластера является его размер, или «радиус». Это свойство наиболее полно отображает фактический размер кластера, если он имеет круглую форму и является гиперсферой в многомерном пространстве. Однако если кластер имеет удлиненную форму, то для оценки его размера лучше применять радиусы соответствующего многомерного эллипсоида. Локальность, или отличимость, кластера характеризует степень перекрытия и взаимной удаленности кластеров друг от друга в многомерном пространстве.

Рассмотрим набор из двух параметров (ОСШ и невязка измерений псевдодальности) и соответственно двумерную плоскость «ОСШ – невязки». Выберем для наших целей евклидову метрику расстояния. Однако поскольку указанные параметры измерены в разных единицах (ОСШ в дБГц, а невязки в метрах), то требуется их перенормировка в безразмерные величины. После этой процедуры евклидова метрика расстояния между точками нашей плоскости становится осмысленной.

Для оси ОСШ производится нормировка вида $y_{\text{норм}} = (y - y_{\text{мин}}) / (y_{\text{макс}} - y_{\text{мин}})$, где для нашего случая можно принять $y_{\text{мин}} = 10$ дБГц, $y_{\text{макс}} = 55$ дБГц. Для оси невязок измерений псевдодальности нормировка имеет вид $x_{\text{норм}} = (x - x_{\text{ср}}) / x_{\text{СКО}}$, переводящая множество значений невязок к нулевому среднему и единичному СКО.

Выделение кластеров производится с использованием метода К-средних [3], который минимизирует дисперсию точек внутри кластеров и максимизирует межкластерное расстояние. Перечислим кратко основные шаги этого метода:

- 1) по некоторому правилу определяется число искомых кластеров k ;
- 2) по некоторому правилу устанавливаются k начальных центров кластеров;
- 3) каждая точка приписывается к ближайшему кластеру;
- 4) производится пересчет положения центра каждого кластера как «центра тяжести» принадлежащих ему точек;
- 5) итерации пп. 3 и 4 повторяются до тех пор, пока изменения координат кластерных центров не станут достаточно малыми.

Будем выбирать начальные центры кластеров, исходя из критерия максимального расстояния между ними. Далее, метод требует априорного задания числа искомых кластеров. Это число сначала оценивалось вручную, а потом тестировалось с помощью алгоритма V-кратной кросс-проверки (V-fold cross-validation [4]). Критерием этого алгоритма является усредненная по различным выборкам ошибка классификации (CV cost), рассчитываемая для различных значений числа кластеров. С ростом числа кластеров значение CV cost сначала достаточно быстро убывает, а затем либо стабилизируется, либо начинает медленно возрастать. На практике помимо минимизации указанного критерия ста-

раются минимизировать еще и число кластеров, что приводит к снижению вычислительной сложности, поэтому первое же существенное замедление скорости убывания ошибки (error rate) объявляется точкой квазиоптимальной классификации. При обработке данных было установлено, что к наилучшим (в указанном смысле) результатам приводит пяти- или шестикластерное разбиение.

Результаты применения кластерного анализа к данным для КА2 и КА10 приведены на рис. 5. Центры пятикластерного разбиения помечены кружками, 6-кластерного разбиения – квадратами. Эллипсы с радиусами, равными СКО, дают представление о форме кластеров (для пятикластерного разбиения).

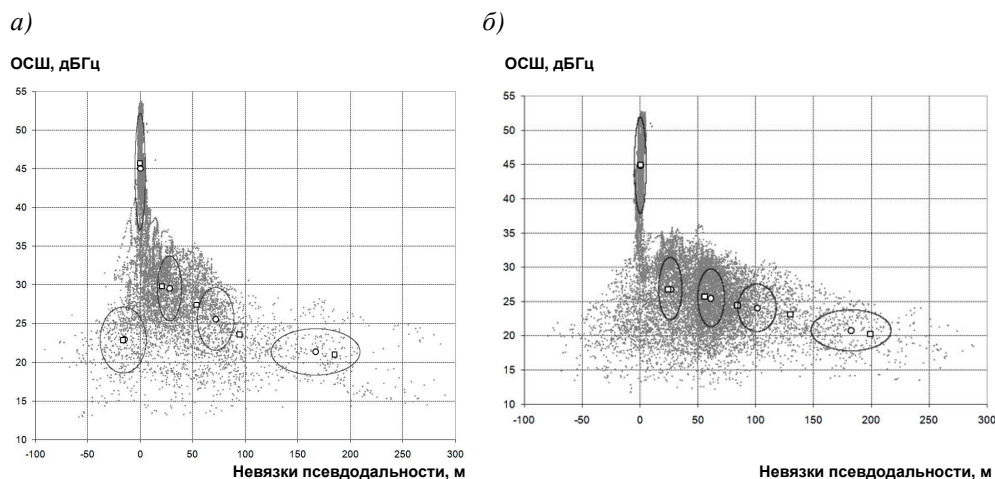


Рис. 5. Результат применения кластерного анализа:
для КА2 (обработано 23049 точек) (а), для КА10 (обработано 23224 точек) (б)

В табл. 1–4 приведены числовые параметры вычисленных кластеров.

Т а б л и ц а 1

Характеристики кластеров для КА2 при пятикластерном разбиении

Параметр	Кластер № 1		Кластер № 2	
	Среднее	СКО	Среднее	СКО
Невязки (м)	-14.87	22.34	0.21	4.92
ОСШ (дБГц)	22.99	4.12	45.03	7.33
Число точек	619		14783	

Параметр	Кластер № 3		Кластер № 4		Кластер № 5	
	Среднее	СКО	Среднее	СКО	Среднее	СКО
Невязки (м)	27.97	11.13	71.90	16.71	167.12	38.98
ОСШ (дБГц)	29.57	4.04	25.54	4.35	21.39	3.19
Число точек	4693		2510		444	

Т а б л и ц а 2

Характеристики кластеров для КА2 при шестикластерном разбиении

Параметр	Кластер № 1		Кластер № 2		Кластер № 3	
	Среднее	СКО	Среднее	СКО	Среднее	СКО
Невязки (м)	-16.39	22.15	-0.14	4.72	20.70	9.29
ОСШ (дБГц)	22.87	4.13	45.74	6.66	29.84	4.02
Число точек	580		14123		4295	

Параметр	Кластер № 4		Кластер № 5		Кластер № 6	
	Среднее	СКО	Среднее	СКО	Среднее	СКО
Невязки (м)	53.63	10.21	94.68	17.04	185.27	34.26
ОСШ (дБГц)	27.44	4.32	23.60	3.60	21.01	3.32
Число точек	2646		1102		303	

Т а б л и ц а 3

Характеристики кластеров для КА10 при пятикластерном разбиении

Параметр	Кластер № 1		Кластер № 2	
	Среднее	СКО	Среднее	СКО
Невязки (м)	0.38	5.81	26.54	10.15
ОСШ (дБГц)	44.82	7.18	26.74	4.84
Число точек	13849		3113	

Параметр	Кластер № 3		Кластер № 4		Кластер № 5	
	Среднее	СКО	Среднее	СКО	Среднее	СКО
Невязки (м)	61.21	9.83	101.48	16.42	182.27	33.07
ОСШ (дБГц)	25.49	4.09	24.04	3.50	20.78	2.96
Число точек	4191		1578		493	

Т а б л и ц а 4

Характеристики кластеров для КА10 при шестикластерном разбиении

Параметр	Кластер № 1		Кластер № 2		Кластер № 3	
	Среднее	СКО	Среднее	СКО	Среднее	СКО
Невязки (м)	0.36	5.81	24.25	9.62	56.00	7.99
ОСШ (дБГц)	44.94	7.04	26.77	4.88	25.71	4.21
Число точек	13774		2871		3618	

Параметр	Кластер № 4		Кластер № 5		Кластер № 6	
	Среднее	СКО	Среднее	СКО	Среднее	СКО
Невязки (м)	84.24	9.71	129.92	16.05	198.91	30.16
ОСШ (дБГц)	24.46	3.56	23.12	3.36	20.28	2.88
Число точек	1954		691		316	

Анализируя полученные результаты, можно сделать ряд выводов. Все LOS-сигналы ожидаемо попадают в один кластер, причем рассчитанные значения внутригрупповой дисперсии (или среднеквадратичного отклонения, СКО) достаточно точно отражают его линейные размеры. Для КА2 выделен кластер с отрицательными невязками (отсутствующий для КА10), что связано с особенностями геометрической картины отражений для КА2 в эксперименте. Увеличение числа кластеров с пяти до шести не изменяет общий характер разбиения и затрагивает относительно небольшое (до 10% измерений) множество переотраженных сигналов.

Описание метода

Учитывая одинаковый характер распределения кластеров для различных КА СРНС, можно предложить следующий метод последовательной обработки измерений в реальном времени для идентификации и компенсации многолучевости.

1. Обозначим значение параметра невязки псевдодальности как НПД. Расчет и хранение статистик кластеров в двумерном пространстве (ОСШ, НПД) для каждого КА СРНС производится в виде множества значений $\{\text{ОСШ}_{\text{макс}}, \text{ОСШ}_{\text{ср0}}, \text{ОСШ}_{\text{ср1}}, \text{НПД}_{\text{ср1}}, \dots, \text{ОСШ}_{\text{срN}}, \text{НПД}_{\text{срN}}, \text{ОСШ}_{\text{мин}}\}$, где N – максимальный учитываемый порядок отражения (1 – однократное, 2 – двукратное и т. д.). Индекс 0 соответствует кластеру LOS-сигнала, и для него нет необходимости хранить среднее НПД, так как оно нулевое (в предположении, что невязки центрированы).

2. В самом начале работы метода предполагается $N=1$ (т.е. моделируются только два кластера – один, соответствующий LOS-сигналу, и второй, соответствующий наиболее сильному из отраженных сигналов), затем по мере накопления данных N может увеличиваться. Для наиболее эффективной работы метода все измерения должны начинаться в условиях прямой видимости большинства сигналов. В этом случае легко формируется первый LOS кластер.

3. Для контроля работы метода предполагается, что центры ($\text{ОСШ}_{\text{срi}}, \text{НПД}_{\text{срi}}$) кластеров отстоят друг от друга не менее чем на 6-8 дБГц по оси ОСШ. Это ограничивает общее число кластеров и препятствует кластеризации случайных выбросов. Также разброс мощностей LOS-сигнала от разных спутников не должен превышать 8-10 дБГц (см. верхнюю «дорожку» на рис. 1).

4. Если новое измерение (невязка и ОСШ) для каждого сигнала находится не далее выбранного порога расстояния от центра i -го кластера, то оно считается попавшим в этот кластер. Величины $\text{ОСШ}_{\text{срi}}$ и $\text{НПД}_{\text{срi}}$, определяющие центр i -го кластера, пересчитываются с учетом нового измерения. При этом измерение псевдодальности, соответствующее отраженному сигналу (т.е. $i > 0$), компенсируется на значение поправки $\text{НПД}_{\text{срi}}$. Если точка не попала в какой-либо кластер, то измерение отбраковывается.

5. В целях коррекции одновременного изменения уровня сигнала для всех КА вследствие общего изменения условий приема сигналов для антенны (которое возникает, например, при движении автомобиля под листвой деревьев) все значения ОСШ нормируются относительно максимального значения $\text{ОСШ}_{\text{макс}}$ по всем сигналам.

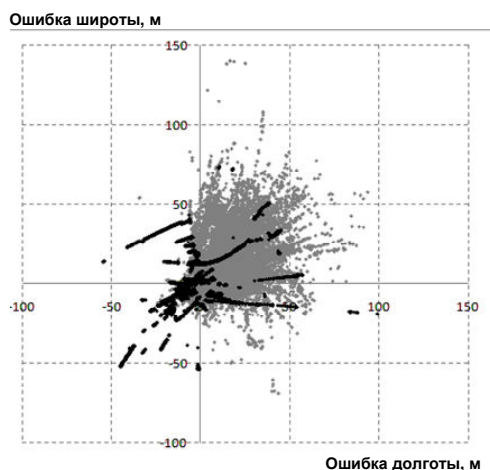


Рис. 6. Сравнение 2D-навигационного решения по скомпенсированным измерениям (черный цвет) и по исходным данным (серый цвет) для статического теста (24 ч)

Апробация метода

Предложенный метод был проверен в статике с помощью вспомогательной программы постобработки измерений. На рис. 6 серым цветом показан результат навигационного 2D-решения по исходным данным, черным – результат применения метода с коррекцией измерений псевдодальности. Точность местоопределения увеличилась за счет значительного уменьшения среднего смещения (с 14 до 4 м). Вместе с тем сохраняется значительный разброс оценки положения (СКО), что может быть вызвано и недостатками самого алгоритма, и неучтенными факторами постобработки.

Заключение

В данной статье произведен статистический анализ экспериментальных данных, полученных при работе СРНС приемников в статических условиях многолучевости. Показано, что для идентификации и компенсации отраженного луча в статических условиях возможно применение методов кластерного анализа. Предложен метод, демонстрирующий улучшение точности местоопределения за счет уменьшения смещения. Для практического применения требуется доработка метода в целях уменьшения дисперсии оценки. Дальнейшее развитие рассмотренного метода предполагает увеличение числа факторов модели, а также его модификацию для динамического потребителя.

ЛИТЕРАТУРА

1. **J. Stone, Dr. M. Chansarkar.** Anti-Multipath Triangulation (AMT) for Positioning in Dense Urban Environments, ION GNSS 17th International Technical Meeting of the Satellite Division, 21-24 Sept. 2004, Long Beach, CA.
2. **Факторный**, дискриминантный и кластерный анализ // Дж.-О.Ким [и др.] / Пер с англ./ Под ред. И.С.Енюкова. - М.: Финансы и статистика, 1989. - 215с.
3. **A. Jain, R. Dubes** - Algorithms for clustering data, 1988. – 320p.
4. **T. Hill, P. Lewicki** - STATISTICS Methods and Applications. StatSoft, 2007, Tulsa, OK. (Electronic Version: <http://www.statsoft.com/textbook/classification-trees/>).

Abstract. The paper presents the statistic analysis of experimental data obtained during operation of navigational receivers of satellite radio-navigation systems (SRNS) in multipathing conditions. A signal-to-noise ratio and pseudorange measurement discrepancies are selected as parameters to be examined. The possibility of applying the methods of cluster analysis for identification and compensation of indirect ray is considered. The procedure is suggested for increasing the SRNS receiver positioning accuracy. The results of experimental check of this procedure in static mode are presented.