

НАХОЖДЕНИЕ ОПТИМАЛЬНОГО ВЕКТОРА ПРИЗНАКОВ ДЛЯ ОПРЕДЕЛЕНИЯ КОНТЕКСТА ОКРУЖАЮЩЕЙ СРЕДЫ ПО ДАННЫМ ГЛОБАЛЬНЫХ НАВИГАЦИОННЫХ СПУТНИКОВЫХ СИСТЕМ

© 2024 г. А. И. Болкунов^а, В. В. Кульнев^а, Е. В. Кульнев^а,
Е. О. Наконечный^{а,*}, В. И. Яремчук^а

^аАО «ЦНИИмаш», Королёв, Россия

*e-mail: nakonechnyieo@tsniimash.ru

Поступила в редакцию 16.02.2024 г.

После доработки 10.03.2024 г.

Принята к публикации 13.05.2024 г.

В глобальных навигационных спутниковых системах показатели качества позиционирования зависят как от условий окружающей среды, так и от поведения потребителя. Окружающая среда влияет на качество приема радиосигналов, которые доступны для позиционирования. Для работы в различных условиях окружающей среды требуется адаптивное навигационное решение, которое будет определять тип окружающей среды и применять различные методы для навигационного решения. Рассматриваются признаки, формируемые по данным принимаемых навигационных сигналов, которые могут быть использованы для определения типа окружающей среды. Настоящая статья посвящена нахождению оптимального вектора признаков для определения типа окружающей среды по информации от глобальных навигационных спутниковых систем. Собраны экспериментальные навигационные данные для различных типов окружающей среды. Рассмотрены критерии и методы определения оптимального вектора признаков с помощью алгоритмов из математической статистики. Предложен оптимальный вектор признаков, который вносит наибольший вклад в определение различных типов окружающей среды.

Ключевые слова: ГНСС, навигационные сигналы, контекст окружающей среды, вектор признаков

DOI: 10.31857/S0002338824040118 EDN: UDUYPP

FINDING THE OPTIMAL FEATURE VECTOR FOR DETECTING THE ENVIRONMENTAL CONTEXT FROM GLOBAL NAVIGATION SATELLITE SYSTEMS DATA

A. I. Bolkunov^а, V. V. Kulnev^а, E. V. Kulnev^а,
E. O. Nakonechnyi^{а,*}, V. I. Yaremchuk^а

^аJSC «TsNIImash», Korolev, Russia

*e-mail: nakonechnyieo@tsniimash.ru

In global navigation satellite systems, positioning quality indicators depend on both environmental conditions and user behaviour. The environment affects the reception quality of the radio signals that are available for positioning. An adaptive navigation solution is required to operate in different environmental conditions, which will detect the type of environment and apply different methods for navigation solution. The features formed from the received navigation signal data that can be used to determine the type of environment are discussed. This paper is devoted to finding an optimal feature vector for determining the type of environment from information from global navigation satellite systems. Experimental navigation data for different types of environment are collected. Criteria and methods for determining the optimal feature vector using algorithms from mathematical statistics are considered. The optimal feature vector that contributes the most to the determination of different types of environment is proposed.

Keywords: GNSS, navigation signals, environmental context, feature vector

Введение. Глобальные навигационные спутниковые системы (ГНСС) являются основой современной навигации благодаря высокой точности, охвату и низкой стоимости абонентских терминалов. В последние годы спутниковое позиционирование стало наиболее эффективным в связи с увеличением количества космических аппаратов (КА), совершенствованием навигационной аппаратуры потребителя (НАП) и разработкой новых алгоритмов, которые повышают точность навигации [1].

Однако характеристики внешних условий и поведение потребителя обычно отличаются от средних показателей, учтенных в алгоритмах НАП, что может привести к ухудшению точности местоопределений в НАП, например из-за приема сигналов вне зоны прямой видимости или эффекта многолучевости [1]. Данные эффекты особенно актуальны в плотной городской застройке, в которой ошибки позиционирования могут составлять значения до десятков метров [1]. Основная проблема заключается в том, что большинство алгоритмов, направленных на уменьшение влияния указанных выше факторов, эффективно работают только в условиях определенной окружающей среды (контексте окружающей среды) и алгоритмы навигации требуется адаптировать под реальные условия окружающей среды [2]. В контексте окружающей среды обычно выделяют классы окружающей среды (например, открытая местность, закрытая и др.) и типы окружающей среды (морская поверхность, воздушное пространство и т.п.) [2]. В связи с этим важным становится понимание окружающего контекста (ОК) потребителя и выбор подходящего алгоритма для увеличения качества навигации [2]. ОК обычно определяется путем анализа степени соответствия характерных признаков анализируемой среды возможным классам и типам окружающей среды. Характерные признаки среды могут быть найдены и проанализированы различными методами.

В настоящее время часто применяемым подходом для распознавания реального ОК выступает техническое зрение (путем обработки изображений с камер или лидаров), которое широко используется в автономных транспортных средствах и роботизированной промышленности. Однако применение средств технического зрения совместно с ГНСС требует высоких затрат энергопотребления, вычислительных ресурсов, большой массы, габаритов и стоимости оборудования и не всегда возможно для массового потребителя. Другим подходом, который позволяет устранить указанные выше недостатки, является непосредственное использование данных от ГНСС [2]. Благодаря принимаемым с КА навигационным сигналам можно сформировать характерные признаки (вектор признаков), которые изменяются в зависимости от ОК.

В [3] рассмотрен комплексный подход для контекстно-адаптивной навигации, заключающийся в применении ГНСС, Wi-Fi и инерциальных навигационных систем. Для определения ОК по данным ГНСС рассчитывалось среднее значение отношения мощности несущей к мощности шума на единицу полосы пропускания C/N_0 дБГц. В [4] демонстрируется определение ОК с помощью модуля ГНСС в смартфоне. Для определения внешней или внутренней среды из данных смартфона извлекается C/N_0 и количество видимых спутников, затем для классификации применяется схема обнаружения с помощью скрытой модели Маркова. В [5] для определения типа ОК по данным от НАП предлагается формировать многомерный вектор признаков, состоящий из среднего значения мощности сигнала C/N_0 , среднеквадратичного отклонения (СКО) мощности сигнала C/N_0 , коэффициента блокировки спутников, суммарного геометрического фактора, расширенного геометрического фактора, количества видимых спутников, которые извлекаются из принимаемых данных навигационного протокола NMEA. Предложенный алгоритм учитывает измерения C/N_0 для открытой среды в других рассматриваемых средах, а также скорость движения потребителя, тем самым дополняя вектор признаков. В [6] представлен метод определения ОК по следующим признакам: среднее значение C/N_0 , СКО мощности сигнала C/N_0 , медиана C/N_0 , верхние и нижние квартили C/N_0 , количество видимых спутников, геометрический фактор по местоположению, горизонталь и вертикали. В [7] описана мобильная наземная платформа для сбора необходимых данных с увеличенной частотой измерений информации, в которой используются следующие признаки: количество видимых спутников, среднее значение маски для количества видимых спутников по C/N_0 , среднее значение для C/N_0 для предыдущих измерений спутников, остаток ошибки псевдодалности и среднее значение угла места.

В рассмотренных выше статьях представлены основные признаки для определения ОК, однако во всех данных работах отсутствует комплексный подход для формирования оптимального вектора признаков, который бы содержал только наиболее значимые признаки. Кроме того, для НАП массового применения нахождение оптимального вектора признаков важно также для минимизации используемых вычислительных ресурсов и энергопотребления. Предлагаемый в работе оптимальный вектор признаков для определения ОК в НАП может быть рассчитан на этапе вторичной обработки сигналов в навигационном чипе или в постобработ-

ке данных на процессоре навигационного устройства, например, в приложении мобильного телефона, однако в таком случае оперативность определения ОК будет ниже.

В разд. 1 сформирован вектор признаков для последующей оптимизации, а также добавлены новые признаки, не рассмотренные ранее в перечисленных работах. В разд. 2 описаны критерии и методы определения оптимального вектора признаков на основе алгоритмов математической статистики. В разд. 3 рассмотрены различные классы и типы ОК в исследовании. В разд. 4 показаны результаты исследований и предложены оптимальные векторы признаков.

1. Формирование вектора признаков. В статье [5] для определения типа ОК были предложены следующие признаки:

$$\gamma(t) = [\mu(t), \sigma(t), \alpha(t), \lambda(t), n(t), GDOP(t)], \quad (1.1)$$

где $\mu(t)$ – среднее значение мощности сигнала C/No , $\sigma(t)$ – СКО средней мощности сигнала C/No , $\alpha(t)$ – коэффициент блокировки спутников, $\lambda(t)$ – расширенный геометрический фактор, $n(t)$ – количество видимых спутников, $GDOP(t)$ – геометрический фактор.

Значения указанных признаков могут быть вычислены по следующим зависимостям.

Коэффициент блокировки $\alpha(t)$ спутников [5]:

$$\alpha(t) = 1 - \frac{N_{\text{видимый}}(t)}{N_{\text{всего}}(t)}, \quad (1.2)$$

где $N_{\text{видимый}}(t)$ – количество видимых спутников, $N_{\text{всего}}(t)$ – общее количество всех спутников, которые должны быть видны потребителю.

Геометрический фактор точности определения местоположения с учетом поправок показаний часов потребителя (GDOP) рассчитывается по следующей формуле [1, 5]:

$$GDOP = \sqrt{\text{tr}(H)}, \quad (1.3)$$

$$H = (G^T G)^{-1}, \quad (1.4)$$

$$G = \begin{bmatrix} -\cos \theta^{(1)} \sin \varphi^{(1)} & \cos \theta^{(1)} \cos \varphi^{(1)} & -\sin \theta^{(1)} & 1 \\ -\cos \theta^{(2)} \sin \varphi^{(2)} & \cos \theta^{(2)} \cos \varphi^{(2)} & -\sin \theta^{(2)} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ -\cos \theta^{(N)} \sin \varphi^{(N)} & \cos \theta^{(N)} \cos \varphi^{(N)} & -\sin \theta^{(N)} & 1 \end{bmatrix}, \quad (1.5)$$

где θ – угол возвышения спутника, φ – азимут спутника.

Расширенный геометрический фактор $\lambda(t)$ [5]:

$$\lambda(t) = \frac{GDOP(t)}{GDOP^0(t)}, \quad (1.6)$$

где $GDOP^0(t)$ – геометрический фактор по всем видимым спутникам.

В [7] для определения типа ОК рассмотрен признак ограничения количества видимых спутников по C/No :

$$M(t) = \sum_{k=1}^N \text{Mask}_k(t), \quad (1.7)$$

$$\text{Mask}_k = \begin{cases} 1, & \frac{C}{No} > \varepsilon, \\ 0, & \text{иначе} \end{cases}, \quad (1.8)$$

где Mask_k – маска для спутников по граничному значению ε для C/No .

Подход среднего значения $\zeta(t)$ для всех видимых спутников в момент измерения с учетом сглаживающего окна C / No_t по каждому спутнику приведен в [7]:

$$\zeta(t) = \frac{1}{N} \sum_{t=1}^N C / No_t. \quad (1.9)$$

Здесь N – количество измерений по каждому КА в единицу времени, C / No_t – среднее значение C / No для k предыдущих измерений видимого спутника (сглаживающее окно для видимого спутника):

$$C / No_t = \frac{1}{W} \sum_{kt=0}^W C / No_{kt}. \quad (1.10)$$

где W – длительность сглаживающего окна для видимого спутника, C / No_{kt} – предыдущее k -е измерение значения C / No для видимого спутника относительно текущего момента измерения t .

СКО $\xi(t)$ сглаживающего окна для среднего значения C / No [7]:

$$\xi(t) = \sqrt{\frac{1}{N} \sum_{t=1}^N (C / No_t - \zeta(t))^2}. \quad (1.11)$$

Средний угол возвышения для видимых спутников $E(t)$ [7]:

$$E(t) = \frac{1}{N} \sum_{k=1}^N elev_k(t), \quad (1.12)$$

где N – количество измерений по каждому КА в единицу времени, $elev_k(t)$ – угол возвышения по k КА.

Для увеличения состава определяемых типов ОК и повышения точности определения предложено расширить состав вектора и добавить следующие новые признаки: геометрический фактор точности определения местоположения потребителя по горизонтали, расширенный геометрический фактор точности по горизонтали $\chi(t)$, средний коэффициент многолучевости $\omega(t)$, средняя оценка ошибки псевдодальности $p(t)$ и СКО оценки ошибки псевдодальности $\phi(t)$. Выражения для расчета значений новых признаков приведены ниже.

Геометрический фактор точности определения местоположения потребителя по горизонтали рассчитывается по следующей формуле [1]:

$$HDOP = \sqrt{D_{11} + D_{22}}, \quad (1.13)$$

где D – диагональные элементы матрицы H , расширенный геометрический фактор точности $\chi(t)$ определения местоположения потребителя по горизонтали вычисляется аналогично из (1.6).

Средний коэффициент многолучевости $\omega(t)$ определяется как

$$\omega(t) = \frac{1}{N} \sum_{k=1}^N m_k(t), \quad (1.14)$$

где N – количество измерений по каждому КА в единицу времени, m_k – коэффициент многолучевости по k КА [8, 9].

Средняя оценка ошибки псевдодальности $p(t)$ равна

$$p(t) = \frac{1}{N} \sum_{k=1}^N R_k(t), \quad (1.15)$$

где N – количество измерений по каждому КА в единицу времени, R_k – оценка ошибки псевдодальности по k КА [8, 9].

СКО оценки ошибки псевдодальности $\phi(t)$ запишем как

$$\phi(t) = \sqrt{\frac{1}{N} \sum_{k=1}^N (R_k(t) - p(t))^2}, \quad (1.16)$$

где N – количество измерений по каждому КА в единицу времени, R_k – оценка ошибки псевдодальности по k КА [8], $p(t)$ – средняя оценка ошибки псевдодальности.

Таким образом итоговый расширенный вектор признаков для оптимизации с помощью алгоритмов из математической статистики может быть представлен в виде:

$$\gamma(t) = \begin{bmatrix} \mu(t), \sigma(t), \alpha(t), \lambda(t), n(t), GDOP(t), HDOP(t), \chi(t), \\ \omega(t), \varphi(t), M(t), \zeta(t), \xi(t), E(t) \end{bmatrix}. \quad (1.17)$$

Учитывая различную информационную ценность характерных признаков, их число и состав вектора может быть оптимизирован.

2. Критерии и методы оптимизации вектора признаков. Для отбора наиболее значимых признаков и нахождения оптимального вектора рассмотрим часто используемые критерии и методы в математической статистике: корреляционный анализ, оценка дисперсии, хи-квадрат, критерий Фишера, взаимную информацию, L1 – регуляризацию, метод главных компонент, экстремальный градиентный бустинг и случайный лес.

Помимо критериев, которые применяются непосредственно в методах, описанных выше, целесообразно добавить критерий точности определения типа ОК.

Дополнительный критерий точности определяется по формуле [10]:

$$A(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^{n-1} 1(y = \hat{y})_i, \quad (2.1)$$

где y – истинное значение ОК, $\hat{y}$ – предсказанное значение ОК, $1(x)$ – индикатор функции.

Для расчета точности определения ОК будем использовать метод логистической регрессии совместно с перекрестной энтропией как наиболее простой линейный классификатор [11]:

$$P(y = 1|x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)}}, \quad (2.2)$$

где y – вероятность наступления события, X_p – признаки модели, β_p – коэффициенты регрессии модели.

Для корреляционного анализа запишем коэффициент корреляции Пирсона и нелинейную корреляцию через расстояние. Коэффициент корреляции Пирсона [12]

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}, \quad (2.3)$$

где X, Y – случайные величины, $\sigma_X \sigma_Y$ – стандартные отклонения, cov – ковариация. Корреляция через расстояние определяется как [13]

$$v_n^2(x, y) = \frac{1}{n^2} \sum_{i,j=1}^n A_{i,j} B_{i,j}, \quad (2.4)$$

$$A_{i,j} = a_{i,j} - \frac{1}{n} \sum_{l=1}^n a_{i,l} - \frac{1}{n} \sum_{k=1}^k a_{i,k} - \frac{1}{n^2} \sum_{k,l=1}^k a_{k,l},$$

$$B_{i,j} = b_{i,j} - \frac{1}{n} \sum_{l=1}^n b_{i,l} - \frac{1}{n} \sum_{k=1}^k b_{i,k} - \frac{1}{n^2} \sum_{k,l=1}^k b_{k,l},$$

где $a_{i,j} = \|x_i - x_j\|$ и $b_{i,j} = \|y_i - y_j\|$ – попарные Евклидовы расстояния x и y .

Из (2.4) получаем корреляцию расстояния:

$$R_n^2(x, y) = \begin{cases} \frac{v_n^2(x, y)}{\sqrt{v_n^2(x, x) v_n^2(y, y)}} , & \sqrt{v_n^2(x, x) v_n^2(y, y)} > 0 , \\ 0 , & \sqrt{v_n^2(x, x) v_n^2(y, y)} = 0 . \end{cases} \quad (2.5)$$

Наиболее простым методом отбора признаков является расчет дисперсии, который сводится к тому, что признаки, имеющие наименьшую дисперсию, отбрасываются:

$$D = E[(X - E(X))^2], \quad (2.6)$$

где X – случайная величина, E – математическое ожидание.

Критерий согласия Пирсона или хи-квадрат может быть вычислен следующим образом [14]:

$$\chi^2 = n \sum_{i=1}^k \frac{(\frac{n_i}{n} - P_i(\Theta))^2}{P_i(\Theta)}, \quad (2.7)$$

где n – общее количество измерений, n_i – количество измерений для выбранного интервала, $P_i(\Theta)$ – вероятность для выбранного интервала.

Критерий Фишера в дисперсионном анализе определяет, согласно [15]:

$$F = \frac{\sum_{i=1}^K \frac{n_i(\bar{Y}_i - \bar{Y})^2}{K-1}}{\sum_{i=1}^K \sum_{j=1}^{n_i} \frac{(Y_{ij} - \bar{Y}_i)^2}{N-K}}, \quad (2.8)$$

где $\bar{Y}_i$ – выборочное среднее выборки, $\bar{Y}$ – общее среднее выборки, K – количество групп, Y_{ij} – j наблюдение в i выборке, n_i – количество наблюдение в i выборке, N – размер выборки.

Запишем метод взаимной информации, который показывает количество информации, содержащееся в одной случайной величине относительно другой с помощью энтропии [16]:

$$I(X, Y) = H(X) - H(Y), \quad (2.9)$$

где $H(X)$ – энтропия события X , $H(Y)$ – энтропия события X при условии события Y . Распишем подробнее (2.9):

$$I(X, Y) = \sum_{y \in Y} \sum_{x \in X} P_{(X,Y)}(x, y) \log \frac{P_{(X,Y)}(x, y)}{P_{(X)}(x) P_{(Y)}(y)}, \quad (2.10)$$

где $P_{(X,Y)}$ – совместная вероятность событий X и Y , $P_{(X)}$ и $P_{(Y)}$ – вероятности событий X и Y .

Метод экстремального градиентного бустинга использует повышающиеся деревья решений, которые последовательно исправляют ошибки, в качестве оценщика применяется среднее усиление по всем выборкам [17]:

$$\text{Усиление} = \frac{1}{2} \left[\frac{G_L^2}{H_L + \alpha} + \frac{G_R^2}{H_R + \alpha} - \frac{(G_L + G_R)^2}{H_R + H_L + \alpha} \right] - \sigma, \quad (2.11)$$

где G_L , G_R – коэффициенты в функции потерь для 1-го члена ряда Тейлора, H_L , H_R – коэффициенты в функции потерь для 2-го члена ряда Тейлора, α, σ – коэффициенты регуляризации.

Метод случайного леса заключается в использовании ансамбля решающих деревьев [18], в качестве оценщика признаков применяется примесь Джини:

$$I_G(p) = 1 - \sum_{i=1}^J p_i^2, \quad (2.12)$$

где J – количество классов, P_i – вероятность элемента выбранного класса i .

Метод регуляризации L1 (совместно с логистической регрессией) [19] равен

$$L_1 = \sum_{i=1}^N (y_i - f(y_i))^2 + \lambda \sum_{i=1}^N |a_i|, \quad (2.13)$$

где y_i – целевой признак, $f(y_i)$ – предсказание модели, λ, a_i – весовые коэффициенты.

Метод главных компонент сводится к нахождению подпространства меньшей размерности, в ортогональной проекции в которых СКО максимально [20].

3. Условия проведения эксперимента и рассматриваемые классы ОК. В работах [3–9] число определяемых классов и типов ОК варьируется от трех до пяти (высотная городская застройка (каньон), средневысотная городская застройка, пригород, открытая местность и бульвар(лес)).

Учитывая более широкое разнообразие реальных сред функционирования массовых образцов НАП ГНСС, для расширения числа классов и типов ОК, рассмотренных ранее, добавляются типы сред внутри помещения и высотная городская застройка со зданиями, расположенными со всех сторон от НАП (колодец).

Для получения экспериментальных данных в реальных условиях ОК использовалась НАП массового применения: U-blox NEO-M8N.

При проведении эксперимента были собраны приблизительно двухчасовые измерения (табл. 1) с помощью НАП U-blox NEO-M8N (частота измерений 1 Гц) антенны Naigon НХ–CSX601А по открытым сигналам L1OF ГЛОНАСС и L1OC GPS для стационарного потребителя в разных ОК (помещение, полужакрытое пространство (окно, балкон), средневысотная городская застройка, лес (бульвар), высотная городская застройка (городской каньон), городской колодец и открытая местность). Из табл. 1 видно, что среднеквадратичная ошибка определения местоположения по системе ГЛОНАСС и GPS для рассматриваемых классов и типов ОК различная.

Таблица 1. Экспериментальные данные различных типов ОК

Типы ОК	Откры- тое небо	Средне- высотная Городская застройка	Полужакры- тое помещение (окно, балкон)	Внутри помеще- ния	Лесопарк (бульвар)	Город- ской каньон	Город- ской колодец
Общее количество измерений, %	6775 / 14.8	6306 / 13.7	7036 / 15.3	6619 / 14.4	6200 / 13.5	6779 / 14.8	6197 / 13.5
СКО ошибки определения местополо- жения по ГЛОНАСС, м	1.7	10.20	14.06	7.79	8.34	11.47	26.61
Количество навигационных решений от общего количества ГЛОНАСС, %	6775 / 100	5836 / 92.5	5433 / 77.2	336 / 5	5834 / 94	4430 / 65.3	4126 / 66.5
СКО ошибки определения местоположения по GPS, м	1	5.71	13.37	36.82	7.34	14.03	34.26
Количество навигационных решений от общего количества GPS, %	6775 / 100	4418 / 70	4136 / 58.7	1341 / 20.2	5602 / 90.3	5546 / 81.8	981 / 15.8

Примечание. СКО ошибки местоположения рассчитаны для угла места больше 15°, решение навигационной задачи весовым методом наименьших квадратов.

4. Результаты нахождения оптимального вектора признаков. На рис. 1, 2 (для удобства отображения обозначим $G - GDOP$, $H - HDOP$) построена карта корреляций Пирсона и расстояний между признаками сформированного вектора (1.17) для системы GPS. Корреляционный анализ вектора признаков (1.17) по системе GPS показывает наиболее сильную корреляцию ОК со следующими признаками: СКО мощности сигнала δ , количество видимых спутников n , средний коэффициент многолучевости ω , маска M по C/No для количества видимых КА, средний угол места E и сильную нелинейную связь с СКО мощности сигнала δ и маской M по C/No для количества видимых КА.

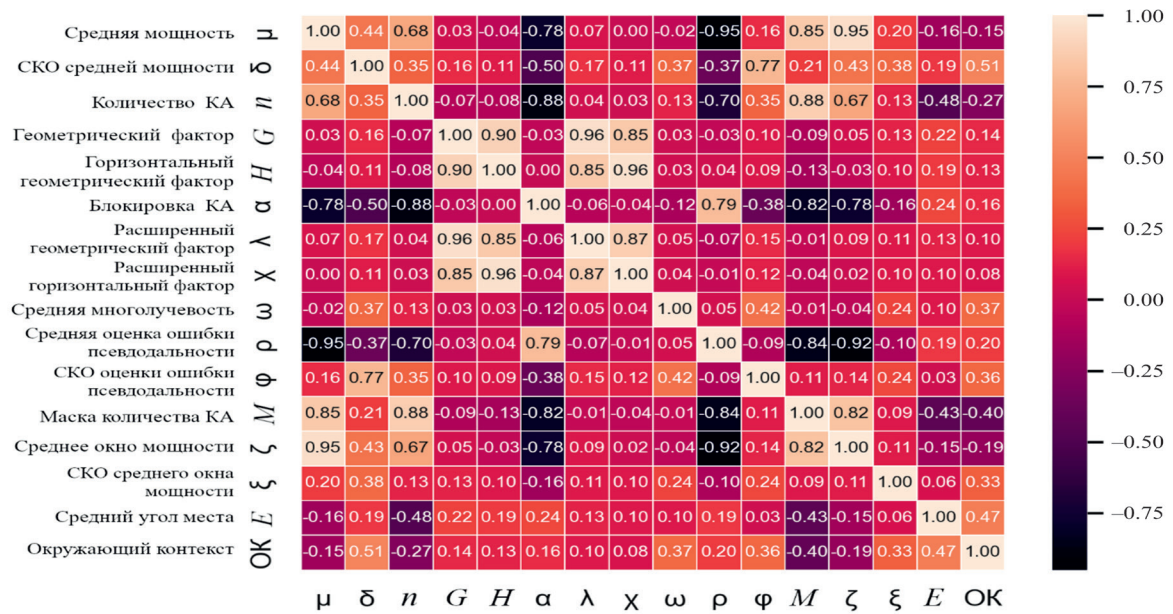


Рис. 1. Корреляция Пирсона для GPS

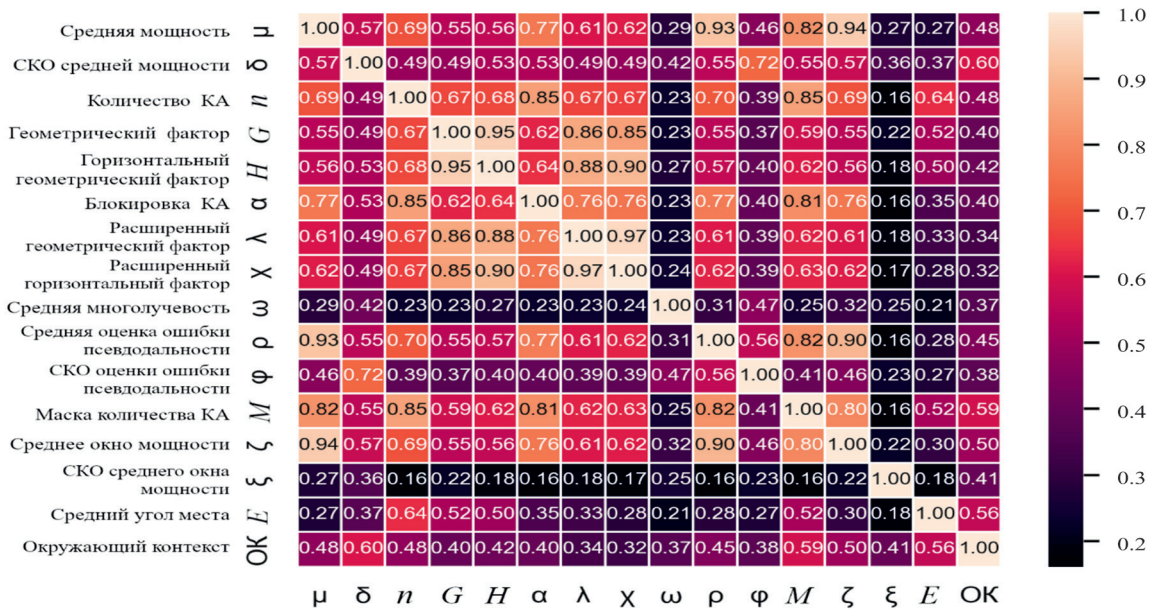


Рис. 2. Корреляция расстояния для GPS

На рис. 3, 4 построена карта корреляций Пирсона и расстояний между признаками сформированного вектора (1.17) для системы ГЛОНАСС. Корреляционный анализ вектора признаков (1.17) по системе ГЛОНАСС показывает наиболее сильную линейную и нелинейную корреляции ОК со следующими признаками: средний коэффициент многолучевости ω , СКО мощности сигнала δ , средний угол места E , маска M по C/N_0 для количества видимых КА и сглаживающего окна ζ для среднего значения C/N_0 , также видно отсутствие связи с геометрическими факторами $GDOP$, $HDOP$, λ , χ , что связано с отсутствием измерений из-за малого количества видимых КА. Оставшиеся признаки для ГЛОНАСС и GPS сильно коррелируют с упомянутыми выше признаками, поэтому их использование нецелесообразно и может привести к неправильному определению ОК.

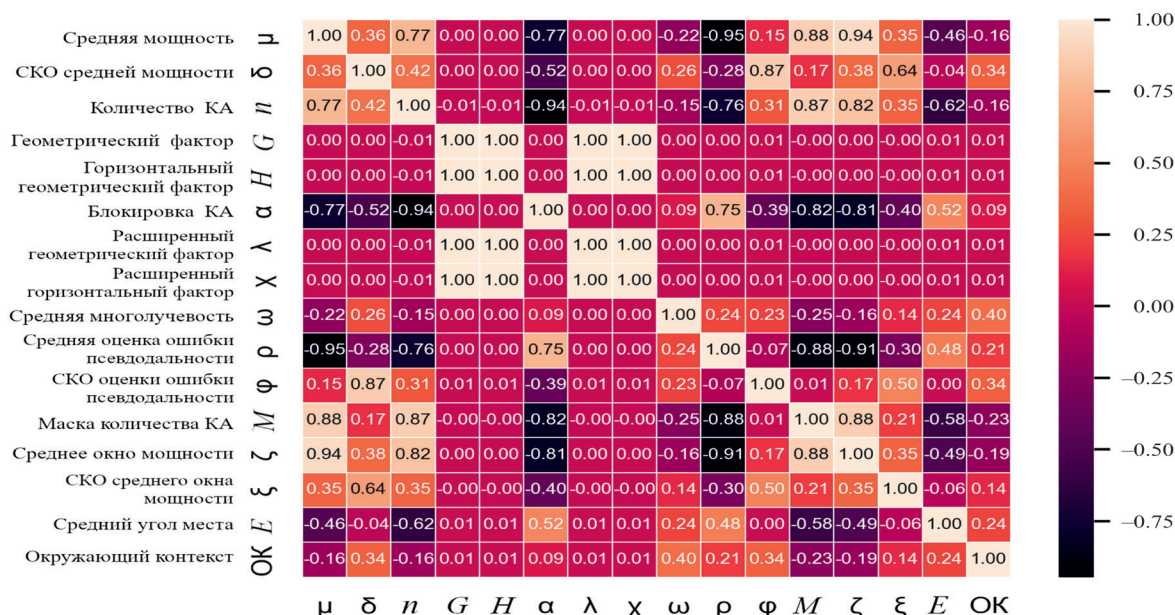


Рис. 3. Корреляция Пирсона для ГЛОНАСС

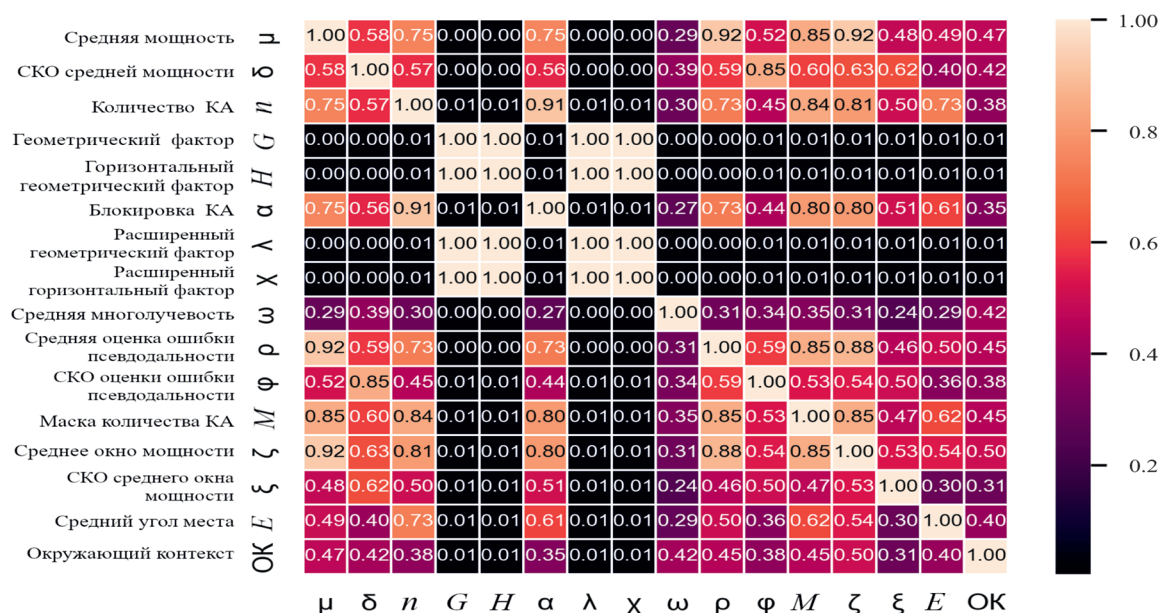


Рис. 4. Корреляция расстояния для ГЛОНАСС

В табл. 2 и 3 показаны результаты выбора наиболее важных признаков по критериям и методам, описанным в разд. 2 для навигационных систем ГЛОНАСС и GPS. Из табл. 2 и 3 видно, что каждый признак может вносить вклад (вплоть до 0.3–2% в методах экстремальный градиентный бустинг и случайный лес) в определение ОК, однако корреляционный анализ показал, что некоторые из этих признаков имеют сильную связь между друг другом и их использование может привести к неверной работе модели в реальных условиях навигации. Также видно, что сглаживающее окно для среднего значения мощности сигнала, СКО мощности сигнала и маска по C/N_0 для количества видимых КА влияют наиболее сильно, что соответствует выбранным в [7], однако эти признаки могут повлечь за собой потерю информации.

Таблица 2. Отбор признаков для системы GPS

Название признаков	Среднее значение мощности сигнала	СКО мощности сигнала	Количество спутников	Суммарный геометрический фактор	Горизонтальный геометрический фактор	Коэффициент блокировки спутников	Расширенный суммарный геометрический фактор	Расширенный горизонтальный геометрический фактор	Средний коэффициент многолучевости	Средняя оценка ошибки псевдоальности	СКО оценки ошибки псевдоальности	Маска количества спутников для $\varepsilon = 30$	Среднее значение мощности сигнала для окна $k = 3$	СКО мощности сигнала для окна	Средний угол мета	Точность на логистической регрессии
	μ	δ	n	G	H	α	λ	χ	ω	ρ	φ	M	ξ	ξ	E	A
Дисперсия	+	+	+	+	+	-	+	+	-	+	+	+	+	+	+	0.849
Хи-квадрат	1708	1823	1890	45	50	8511	17	13	933	3990	2020	3596	1910	324	259	0.839
Критерий Фишера	53412	14962	19452	252	153	23244	120	47	2467	30721	6245	35813	58314	3090	2928	0.882
Взаимная информация	1.23	0.87	0.7	1.57	1.61	0.65	0.99	0.99	0.41	0.93	0.91	0.85	1.29	0.67	1.1	0.885
Рекурсивное исключение логистической регрессией	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	0.903
L1-регуляризация	+	+	+	+	-	+	+	-	+	-	+	+	+	+	+	0.908
Главные компоненты	+	+	+	+	+	+	+	-	+	+	+	+	-	+	+	0.889
Экстремальный градиентный бустинг, %	5.2	4.8	4.8	1.7	1.4	0.7	0.7	0.6	4.4	0.3	1.9	53.4	9.6	5.2	5.3	0
Случайный лес, %	12.8	5.8	5.5	5.8	5.2	2.6	1.8	1.8	2.6	6.4	2.5	8.2	18.6	11.4	8.8	0

Примечание. 0 означает, что расчеты отсутствуют.

Таблица 3. Отбор признаков для системы ГЛОНАСС

Название признаков	Среднее значение мощности сигнала	СКО мощности сигнала	Количество спутников	Суммарный геометрический фактор	Горизонтальный геометрический фактор	Коэффициент блокировки спутников	Расширенный суммарный геометрический фактор	Расширенный горизонтальный геометрический фактор	Средний коэффициент многолучевости	Средняя оценка ошибки псевдодалыности	СКО оценки ошибки псевдодалыности	Маска количества спутников для $\varepsilon = 30$	Среднее значение мощности сигнала для окна $k = 3$	СКО мощности сигнала для окна	Средний угол мета	Точность на логистической регрессии
	μ	δ	n	G	H	α	λ	χ	ω	ρ	ϕ	М	ζ	ξ	E	A
Дисперсия	+	+	+	+	+	-	+	+	-	+	+	+	+	+	+	0.784
Хи-квадрат	1485	2096	2329	6	6	9194	6	6	485	3323	2420	3254	1950	526	245	0.803
Критерий Фишера	34837	13409	27440	0	0	23952	0	0	1320	21078	7648	27514	111629	5065	2716	0.857
Взаимная информация	0.98	0.8	0.69	1.03	1.28	0.66	0.8	0.79	0.56	0.82	0.81	0.84	1.26	0.88	1.21	0.832
Рекурсивное исключение логистической регрессией	+	+	+	-	-	+	-	-	+	+	+	+	+	+	+	0.869
L1-регуляризация	+	+	+	-	-	-	-	-	+	+	-	+	+	-	-	0.807
Главные компоненты	+	-	+	+	+	+	+	-	+	-	+	+	+	+	+	0.842
Экстремальный градиентный бустинг, %	4.1	10.8	9.8	4.5	9.7	1.4	1.5	2.2	6.1	1.4	1.1	4.2	28.3	9.6	5.2	0
Случайный лес, %	8.9	6.3	8.3	0	0	5.3	0	0	2.9	5.3	2.3	8.5	22.6	13.9	11.3	0

Примечание. 0 означает, что расчеты отсутствуют.

На основе анализа таблиц и карт корреляций можно сформировать следующий оптимальный вектор признаков для систем ГЛОНАСС и GPS:

$$\gamma_{opt1}(t) = [\sigma(t), HDOP(t), \omega(t), M(t), \zeta(t), \xi(t), E(t)], \quad (4.1)$$

где $\sigma(t)$ – СКО средней мощности сигнала C/N_0 , $HDOP(t)$ – геометрический фактор точности по горизонтали, $\omega(t)$ – средний коэффициент многолучевости, $M(t)$ – маска количества видимых КА по C/N_0 , $\zeta(t)$ – сглаживающее окно для среднего значения C/N_0 , $\xi(t)$ – СКО сглаживающее окно для среднего значения C/N_0 , $E(t)$ – среднее значение угла места.

В случае если ОК изменяется быстро и скорость движения потребителя высокая, а также вычислительные ресурсы ограничены, можно сформировать следующий оптимальный вектор признаков для систем ГЛОНАСС и GPS:

$$\gamma_{opt2}(t) = [\mu(t), \sigma(t), n(t), HDOP(t), \omega(t), E(t)], \quad (4.2)$$

где $\mu(t)$ – среднее значение мощности сигнала C/N_0 , $\sigma(t)$ – СКО средней мощности сигнала C/N_0 , $n(t)$ – количество видимых спутников, $HDOP(t)$ – геометрический фактор точности по горизонтали, $\omega(t)$ – средний коэффициент многолучевости, $E(t)$ – среднее значение угла места.

Закключение. При использовании ГНСС позиционирование зависит как от условий окружающей среды, так и от поведения потребителя. Окружающая среда влияет на качество приема радиосигналов, которые доступны для позиционирования. Для работы в различных условиях навигации требуется адаптивное навигационное решение, которое будет определять и учитывать контекст окружающей среды в НАП. Для этого можно формировать признаки из информации от навигационных сигналов, приходящей с каждого КА. Такие исследования активно ведутся в последние годы, однако во всех известных работах отсутствует комплексный подход для формирования оптимального вектора признаков.

В ходе проведенной работы получены следующие результаты: добавлены новые классы и типы ОК (внутри помещения и городской колодец), новые признаки для определения ОК (средний коэффициент многолучевости $\omega(t)$, средняя оценка ошибки псевдодалности $p(t)$, СКО оценки ошибки псевдодалности $\varphi(t)$), проведен линейный и нелинейный корреляционный анализ вектора признаков для системы ГЛОНАСС и GPS, отобраны признаки с помощью различных методов и критериев для системы ГЛОНАСС и GPS, найден оптимальный вектор признаков $\gamma_{opt1}(t)$ для системы ГЛОНАСС и GPS, включающий: СКО средней мощности сигнала, средний коэффициент многолучевости, геометрический фактор точности по горизонтали, маску количества видимых КА по C/N_0 , сглаживающее окно для среднего значения C/N_0 , СКО сглаживающего окна для среднего значения C/N_0 и среднее значение угла места. Предложен вектор признаков $\gamma_{opt2}(t)$ для системы ГЛОНАСС и GPS в случае высокودинамичного потребителя и быстрого изменения ОК.

СПИСОК ЛИТЕРАТУРЫ

1. Перов А.И. ГЛОНАСС. Модернизация и перспективы развития. Монография. М.: Радиотехника, 2020.
2. Кульнев В.В., Кульнев Е.В., Наконечный Е.О. Методы определения контекста по данным навигационной аппаратуры глобальных навигационных спутниковых систем // Космонавтика и ракетостроение. 2023. № 4. С. 43–53.
3. Groves P.D., Martin H., Voutsis K., Walter D., Wang L. Context Detection, Categorization and Connectivity for Advanced Adaptive Integrated Navigation. London UK: University College London, 2013.
4. Gao H., Groves P.D. Environmental Context Detection for Adaptive Navigation using GNSS Measurements from a Smartphone. London UK: University College London, 2018.
5. Wang Y., Liu P., Liu Q., Adeel M., Qian J., Jin X., Ying R. Urban Environment Recognition Based on the GNSS Signal Characteristics // Wiley ION. 2018. V. 66.
6. Liu H., Zhang M., Pei L., Wang W., L. Li, Pan C., Li Z. Environment Classification for Global Navigation Satellite Systems Using Attention // Shanghai Key Laboratory of Navigation and Location Based Services. Shanghai: Shanghai Jiao Tong University, 2021. February.
7. Feriol F., Watanabe Y., Vivet D. GNSS-based Environmental Context Detection for Navigation // IEEE Intelligent Vehicles Symposium (IV). Aachen. Germany, 2022. Jun.
8. U-blox M8 Receiver description. UBX-13003221. 2022. August.
9. Digital Cellular Telecommunications System (Phase 2+) (GSM) Location Services (LCS) Mobile Station (MS). Serving Mobile Location Centre (SMLC) Radio Resource LCS Protocol (RRLP) (3GPP TS 44.031 Version 17.0.0 Release 17). ETSI TS 144 031 V17.0.0. 2022.

10. BS ISO 5725-1. Accuracy (Trueness and Precision) of Measurement Methods and Results – Part 1: General Principles and Definitions. Geneva, 1994. P.1.
11. *Cramer J.S.* The Origins of Logistic Regression (PDF) // Tinbergen Institute. 2002. V. 119. P. 167–178.
12. *Pearson K.* Notes on Regression and Inheritance in the Case of Two parents. London: Proceedings of the Royal Society of London, 1895. June.
13. *Székely G.J., Rizzo M.L., Bakirov N.K.* Measuring and Testing Dependence by Correlation of Distances // The Annals of Statistics. 2007. V. 35. № 6. P. 2769–2794.
14. *Никитин М.С.* Критерий хи-квадрат для непрерывных распределений с параметрами сдвига и масштаба // Теория вероятностей и ее применение. 1973. Т. XVIII. Вып. 3.
15. *Box G.E.P.* Non-Normality and Tests on Variances // Biometrika. 1953. V. 40. P. 318–335.
16. *Kraskov A., Stogbauer H., Grassberger P.* Estimating Mutual Information // Physical Review E. 2004. V. 69.
17. *Friedman J.* Greedy Function Approximation: A Gradient Boosting // The Annals of Statistics. 2001. V. 29.
18. *Hastie T., Tibshirani R., Friedman J.* Chapter 15. Random Forests // The Elements of Statistical Learning. 2nd ed. Springer-Verlag, 2009.
19. *Boyd S., Vandenberghe L.* Convex Optimization. London UK: Cambridge University Press, 2004. 716 p.
20. *Tipping M.E., Bishop C.M.* Probabilistic Principal Component Analysis // J. Royal Statistical Society: Series B (Statistical Methodology). 1999. V. 61(3). P. 611–622.