УДК 004.056.53
DOI:10.30533/scidata-2023-14-14

Проблема информационной безопасности пространственных данных в контексте понятия персональных данных и возможные пути ее решения

В.А. Пучин[1], В.А. Хлытина[1]

АФФИЛИАЦИИ

[1] Московский государственный университет геодезии и картографии, Москва, Россия
putch2016@inbox.ru

ЦИТИРОВАНИЕ

Пучин В.А., Хлытина В.А. Проблема информационной безопасности пространственных данных в контексте понятия персональных данных и возможные пути ее решения // Пространственные данные: наука и технологии. 2023. Т. 14. № 1. С. 18–29. DOI:10.30533/scidata-2023-14-14.

КЛЮЧЕВЫЕ СЛОВА

геомаскирование, геозонирование, API геолокации, обезличивание, агрегация точек, агрегация области, корректировка координат, замена координат, FOAM

АННОТАЦИЯ

Пространственные данные становятся персональными в тех случаях, когда в совокупности с другими данными позволяют однозначно идентифицировать субъект данных. В связи с этим необходимо четко разграничивать, в каких случаях следует считать такие данные персональными и защищать их как персональные, а в каких в этом нет необходимости. Целью исследования являлось определение набора необходимых требований по защите пространственных данных в контексте персональных при разработке систем, связанных с определением местоположения. Для достижения обозначенной цели проведена классификация систем, обрабатывающих пространственные данные, изучены подходы

к защите таких данных и обоснованы эффективные меры по их защите. В качестве способов защиты данных рассмотрены различные методы геомаскирования и обезличивания. В результате исследования обоснованы и рационально распределены методы защиты информации для выделенных классов систем. Выявлен набор необходимых требований по защите пространственных данных при разработке систем, связанных с определением местоположения, а также определены направления для дальнейших исследований в этой области.

1 Введение

В настоящее время многие приложения и сайты используют пространственные данные, определяющие местоположение людей, но защите таких данных не уделяется должного внимания. Как следствие, разделение пространственных и персональных данных сказывается на защите систем, обрабатывающих такие данные, что, в свою очередь, поднимает вопросы безопасности и данных как таковых, и данных каждого конкретного человека. При существовании такой проблемы необходимо четко разграничивать, в каких случаях следует считать такие данные персональными, а в каких нет. Авторы считают, что нужно четко обозначить требования, которые должны быть реализованы в системе, обрабатывающей пространственные данные[1], на этапе ее разработки.

В основу исследования положено утверждение, что пространственные данные становятся персональными в тех случаях, когда в совокупности с другими данными позволяют однозначно идентифицировать субъект данных. И, как следствие, эти данные необходимо защищать, поскольку пренебрежение такими свойствами информации, как конфиденциальность, целостность, доступность, может причинить вред владельцу данных.

Проблема недостаточной защищенности данных весьма актуальна и становится масштабнее в связи с развитием технологий геолокации в целях улучшения клиентоориентированных сервисов. Более того, нечетко сформулированное законодательное понятие «персональные данные» позволяет как компаниям, так и физическим лицам закладывать в него выгодную им информацию, что существенно обостряет проблему защиты персональных данных.

Цель статьи — определить набор необходимых требований по защите пространственных данных при разработке систем, связанных с определением местоположения.

Поставленная цель может быть достигнута путем решения следующих задач:

- выявление проблемы безопасности пространственных данных;
- определение рисков для пользователей при утечке таких данных;

1 Здесь и далее в статье под термином «пространственные данные» подразумеваются пространственные данные, относящиеся к персональным данным.

– проведение классификации систем, обрабатывающих пространственные данные;
– изучение подходов к защите пространственных данных;
– обоснование эффективных мер по защите пространственных данных.

В исследовании авторы сосредоточились на изучении существующих методов защиты пространственных данных и выборе наиболее подходящих для применения в современных реалиях.

2 Материалы и методы

Основными методами исследования являлись:

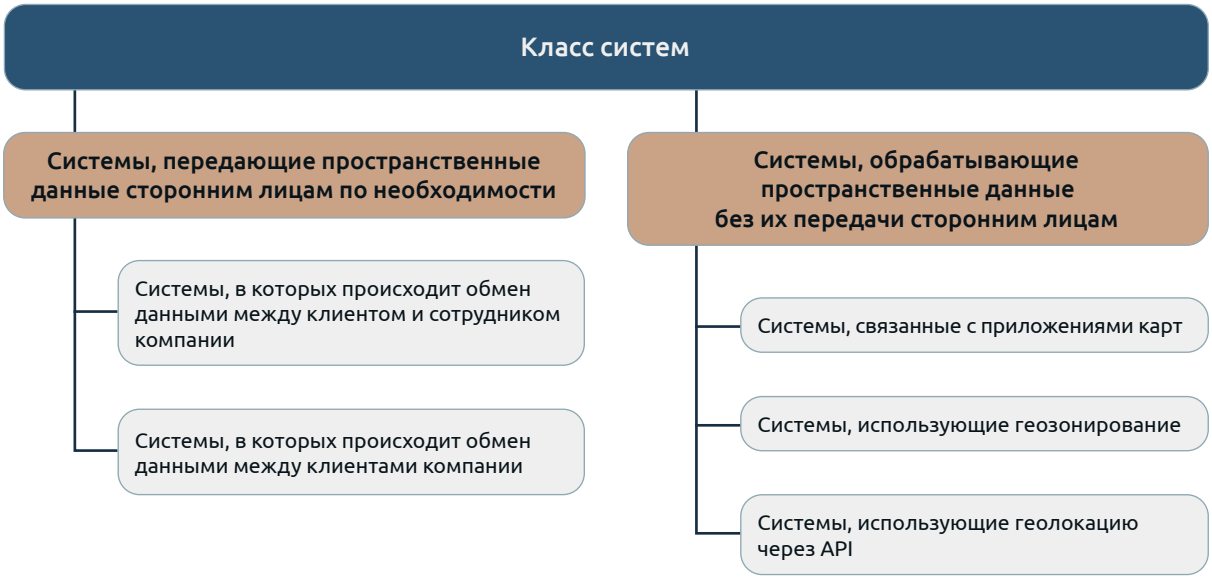
– метод анализа информации, с использованием которого были рассмотрены способы защиты систем, связанных с определением местоположения людей;
– метод классификации, с использованием которого были выделены классы систем, обрабатывающих пространственные данные пользователей приложений.

Объект исследования — способы защиты систем, связанных с обработкой пространственных данных о местоположении.

На первом этапе исследования были выделены следующие классы систем, обрабатывающих пространственные данные пользователей (**Рис. 1**):

Рис. 1 Классы систем, обрабатывающих пространственные данные пользователей.

Fig. 1 Classes of systems processing users' spatial data.



1. **Системы, передающие пространственные данные пользователей сторонним лицам по необходимости:**
   - системы, в которых происходит обмен данными между клиентом и сотрудником компании. В качестве примера таких систем можно назвать системы, где предусмотрена доставка товаров или оказание услуг, например заказ такси. В этом случае данные клиента сообщаются курьеру или водителю такси;
   - системы, в которых происходит обмен данными между клиентами компании. Примером подобной системы может быть приложение знакомств, которое определяет ближайших друг к другу пользователей и показывает их персональную информацию.
2. **Системы, обрабатывающие пространственные данные пользователей без их передачи сторонним лицам:**
   - системы, связанные с приложениями карт, которые обеспечивают передачу информации между пользовательским устройством и инфраструктурой компании;
   - системы, использующие геозонирование[2] (англ. — «geofencing») — технологию, с помощью которой можно задавать границы, при пересечении которых мобильным устройством происходит определенное событие, например приходит рекламное уведомление от компании, приложение которой установлено на телефоне;
   - системы, использующие геолокацию через API (англ. — «Geolocation Application Programming Interface», API), которые позволяют пользователю при его согласии передачу данных о его местоположении приложению. Например, в приложении банка по желанию пользователя могут быть показаны ближайшие к нему банкоматы или отделения банка.

Для определения требований по защите таких систем необходимо вначале рассмотреть методы защиты пространственных данных.

Одним из методов сохранения конфиденциальности персональных данных является их обезличивание. **Методы обезличивания** описаны в приказе Роскомнадзора от 5 сентября 2013 г. № 996 «Об утверждении требований и методов по обезличиванию персональных данных»[3]:

- метод введения идентификаторов. Основан на замене части данных идентификаторами и составлении таблицы соответствия данных и идентификаторов;
- метод изменения состава и семантики. Заключается в обобщении или удалении части данных;

---

2 Геофенсинг (Geofencing) // Энциклопедия «Касперского». 2022. [Электронный ресурс]. Режим доступа: https://encyclopedia.kaspersky.ru/glossary/geofencing/ (дата обращения: 31.10.2022).

3 Приказ Федеральной службы по надзору в сфере связи, информационных технологий и массовых коммуникаций от 5 сентября 2013 г. № 996 «Об утверждении требований и методов по обезличиванию персональных данных». [Электронный ресурс]. Режим доступа: https://minjust.consultant.ru/files/7638 (дата обращения: 31.10.2022).

- метод декомпозиции. Реализуется путем разделения таблицы с персональными данными на несколько небольших таблиц и составления таблицы их соответствия;
- метод перемешивания. Основан на перемешивании данных в таблице.

Метод обезличивания эффективен, но недостаточен для использования в качестве единственного метода. В некоторых случаях при получении злоумышленником нескольких обезличенных баз данных у него появляется возможность персонализировать информацию пользователей систем. Следовательно, для обеспечения конфиденциальности, целостности и доступности информации необходимо использовать несколько способов защиты данных, например геомаскирование.

В работе [1] по ряду критериев проанализированы наиболее известные методы геомаскирования данных, которые авторы данной статьи считают основными методами защиты пространственных данных.

- **Агрегация точек**[4]. Суть метода заключается в том, что точки объединяются в кластеры заданного размера, которые преобразуются в одну точку в пределах кластера, ее размер зависит от количества исходных точек [2].
- **Агрегация области**[5]. Существует несколько вариантов реализации этого метода, но основная идея каждого из них заключается в том, что все точки объединяются по признаку (например, административное деление, ячейка системы координат) в полигоны, далее всем этим точкам присваивается одно значение (например, название района, идентификатор) или эта область закрашивается определенным цветом, который определяется количеством точек в выбранной области.
- **Корректировка координат**[6]. Основной принцип реализации данного метода заключается в смещении всех точек на карте по определенному правилу [3, 4].
- **Замена координат**[7]. Метод реализуется путем кодирования координат по определенному правилу или их замены определенными идентификаторами [5, 6].

На втором этапе исследования обоснованы и распределены рациональным образом методы защиты информации по выделенным ранее классам систем.

Прежде всего следует отметить, что в российском законодательстве закреплен перечень необходимых средств защиты персональных данных. Эти требования

4 Armstrong M.P., Rushton G., Zimmerman D.L. Geographically masking health data to preserve confidentiality // Statistics in Medicine. 1999. Vol. 18. No. 5. P. 497–525. DOI:10.1002/(sici)1097-0258(19990315)18:5%3C497::aid-sim45%3E3.0.co;2-#.

5 Там же.

6 Там же.

7 Distance Aware Address Encoding for Privacy-Preserving Record Linkage. [Электронный ресурс]. Режим доступа: https://medium.com/@wilko.henecka/distance-aware-address-encoding-for-privacy-preserving-record-linkage-a6cecdaadc22 (дата обращения: 10.11.2022).

не являются специфичными в отношении пространственных данных, поэтому их рассмотрение нецелесообразно.

При выборе способа защиты систем авторы использовали риск-ориентированный подход чтобы, с одной стороны, снизить риски при обработке и хранении пространственных данных компанией — владельцем сервиса, а с другой стороны — не ограничивать возможности сервиса, а также не замедлять его работу.

Во время обмена данными между клиентом приложения и представителем компании целесообразно рассмотреть несколько ситуаций, в которых следует использовать разные средства защиты пространственных данных. Особенностью такого вида систем является наличие человека, который будет доставлять товар к местонахождению клиента. До предоставления услуги, включая время на возможность отказа от нее, **авторы предлагают защищать данные** следующим образом:

- внедрять для входа в систему, имеющую доступ к клиентской базе данных, двухфакторную аутентификацию, а также после предоставления услуги запрещать доступ к данным клиента за исключением номера заказа, поскольку чаще всего лица, осуществляющие доставку услуги клиенту, для доступа к системе с данными клиента используют либо личное устройство, либо устройство, принадлежащее компании;
- в базе данных о клиентах должна храниться информация, обезличенная с помощью метода введения идентификаторов или метода декомпозиции, так как эти методы позволяют в случае необходимости персонализировать данные по таблицам соответствия. Важно отметить, что эти таблицы необходимо хранить отдельно от обезличенной базы данных;
- использование одного из методов геомаскирования — аффинных преобразований — для дополнительной защиты пространственных данных клиента. Суть метода аффинных преобразований[8] заключается в изменении координат путем смещения или поворота всех точек относительно какой-либо оси или начала координат.

Также большинству компаний необходимо анализировать данные о заказах для улучшения своего сервиса и развития бизнеса. Чтобы компании не обрабатывали персональные данные в ситуациях, когда они могут получить нужную информацию без них, предлагается обезличивать персональные данные клиента методом изменения их состава и семантики. Если компания не может обезличить персональные данные этим методом, рекомендуется использовать такие методы геомаскирования, как агрегация точек или агрегация области.

Для систем, в которых пользователи взаимодействуют между собой, авторы советуют разработчикам приложения по умолчанию отключить публикацию геолокации. В случаях, когда пользователи сами хотят поделиться геолокацией, следует предупреждать их, почему не следует этого делать, и давать им возможность выбрать, насколько полной геолокацией они хотят поделиться.

8 Loudon T.V., Andrew K.P. Affine transformations for digitized spatial data in geology // Computers & Geosciences. 1980. Vol. 6. No. 4. P. 397–412.

В приложениях карт пользователи чаще всего авторизуются через свои аккаунты в других системах, поэтому пространственные данные пользователя можно связать с конкретным человеком. Если приложениям карт необходимо анализировать данные о перемещениях пользователей, рекомендуется внедрять в системы метод агрегации точек или области. В ином случае пользовательские данные нужно обезличить путем изменения состава и семантики данных.

Для безопасности данных пользователя приложения, в котором реализовано геозонирование, владельцам не следует обрабатывать и хранить никакие данные, кроме времени и зоны события.

При использовании геолокации через API рекомендуется всегда исключать из запроса все персональные данные клиента, кроме геолокации. В базе данных, куда поступают запросы, следует хранить их количество из каждой области, откуда выполнялся запрос, а также его дату.

3 Результаты

В ходе исследования выделены классы систем, обрабатывающих пространственные данные пользователей:

1. **Системы, передающие пространственные данные пользователей сторонним лицам по необходимости:**
   - системы, в которых происходит обмен данными между клиентом и сотрудником компании;
   - системы, в которых происходит обмен данными между клиентами компании.
2. **Системы, обрабатывающие пространственные данные пользователей без их передачи сторонним лицам:**
   - системы, связанные с приложениями карт;
   - системы, использующие геозонирование;
   - системы, использующие геолокацию через API.

Также авторы статьи перечислили основные, по их мнению, методы защиты пространственных данных, а именно следующие виды **геомаскирования**:

1) агрегация точек;
2) агрегация области;
3) корректировка координат;
4) замена координат.

Основной результат работы — обоснование и распределение рациональным образом методов защиты информации по выделенным классам систем (**Табл. 1**).

Таблица 1 Методы защиты информации по подклассам систем.

Table 1 Information protection methods by system subclasses.

Подкласс системы	Метод защиты
Обмен данными между клиентом и сотрудником компании	**До предоставления услуги:**
	двухфакторная аутентификация для сотрудников компании и ограничение их доступа к данным клиента
	обезличивание с помощью метода введения идентификаторов или метода декомпозиции
	геомаскирование методом аффинных преобразований
	После предоставления услуги:
	обезличивание методом изменения состава и семантики
	геомаскирование методом агрегации точек или агрегации области
Обмен данными между клиентами компании	Отключение публикации геолокации по умолчанию, предупреждение пользователей об опасности публикации своего местоположения
Приложения карт	Геомаскирование методом агрегации точек или области, обезличивание методом изменения состава и семантики
Геозонирование	Ограничение сбора собираемых и обрабатываемых данных
Геолокация через API	Исключение из запроса всех персональных данных, кроме геолокации

4 Обсуждение

В перспективе планируется также рассмотреть такой способ защиты пространственных данных, как метод геомаскирования под названием «Сетка официальной статистики»[9]. Для официальной статистики в странах Европейского союза используется специальный метод агрегирования координат по районам. Сетка накладывается на данные, а маскируемые координаты заменяются идентификатором, сформированным из координаты нижнего левого угла или координаты центра ячейки, в которую попадают маскируемые координаты. Для реализации данного метода в РФ необходимо адаптировать его к российской системе координат, в связи с чем требуется разработка с нуля алгоритма автоматического агрегирования координат по районам.

9 Geographical Grids for Germany: GeoGitter. [Электронный ресурс]. Режим доступа: https://gdz.bkg.bund.de/index.php/ default/geographische-gitter-fur-deutschland-in-lambert-projektiongeogitter-inspire.html (дата обращения: 30.11.2022).

Изучение технологии FOAM[10] и ее адаптация с учетом российской специфики — одна из перспектив развития защиты пространственных данных. FOAM — открытый протокол для децентрализованных рынков пространственных данных. Он предназначен для того, чтобы пользователи могли создавать на основе консенсуса карту мира, которой может доверять любое приложение. Составные элементы протокола FOAM предназначены для предоставления пространственных протоколов, стандартов и приложений, которые передают пространственные данные в блокчейны и расширяют возможности карты мира, основанной на консенсусе. FOAM защищает физическое пространство в блокчейне, используя вычислительную мощность Ethereum, и позволяют распределенным пользователям координировать свои действия и взаимодействовать децентрализованно и без разрешения. FOAM — технология, которая находится на стадии разработки, в связи с чем в настоящее время не представляется возможным говорить о ее полноценной работоспособности. Кроме того, она не получила особого распространения нигде, кроме США.

5 Выводы

Разработка сервисов, использующих пространственные данные для предоставления услуг клиентам, — динамично развивающаяся область, которая приобретает все большее значение в жизни современных людей. Технология защиты таких сервисов является важной и перспективной областью исследования, поскольку позволяет сохранять в безопасности данные клиентов и обеспечивать конфиденциальность, целостность и доступность информации.

Авторами обоснован набор необходимых требований по защите пространственных данных в контексте персональных при разработке систем, связанных с определением местоположения. Также были **решены поставленные задачи исследования**:

- выявлена проблема безопасности пространственных данных;
- определены риски для пользователей при утечке таких данных;
- разделены на классы системы, обрабатывающие пространственные данные;
- изучены подходы к защите пространственных данных;
- предложены меры по защите пространственных данных.

Кроме того, авторы отмечают, что помимо технологического стека следует обратить внимание на правовой аспект данного вопроса, например конкретизировать определение термина «персональные данные».

10 FOAM Whitepaper // Foamspace Corp. 2018. [Электронный ресурс]. Режим доступа: https://foam.space/publicAssets/FOAM_Whitepaper.pdf (дата обращения: 22.11.2022).

БИБЛИОГРАФИЯ

1. Redlich S. Quantitative Analysis of Geomasking Methods. Dr. of Sci. thesis. Essen: 2022. 318 p.
2. Kounadi O., Leitner M. Adaptive Areal Elimination (AAE): A tansparent way of disclosing protected spatial datasets // Computers, Environment and Urban Systems. 2016. Vol. 57. P. 59–67. DOI:10.1016/j.compenvurbsys.2016.01.004.
3. Leitner M., Curtis A. Cartographic Guidelines for Geographically Masking the Locations of Confidential Point Data // Cartographic Perspectives. 2004. Vol. 49. No. 7. P. 22–39. DOI:10.14714/CP49.439.
4. Allshouse W.B., Fitch M.K., Hampton K.H., et al. Geomasking sensitive health data and privacy protection: an evaluation using an E911 database // Geocarto International. 2010. Vol. 25. No. 6. P. 443–452. DOI:10.1080%2F10106049.2010.496496.
5. Kroll M., Schnell R. Anonymisation of geographical distance matrices via Lipschitz embedding // International Journal of Health Geographics. 2016. Vol. 15. No. 1. P. 1–14. DOI:10.1186/s12942-015-0031-7.
6. Schnell R., Klingwort J., Farrow J.M. Locational privacy-preserving distance computations with intersecting sets of randomly labeled grid points // International Journal of Health Geographics. 2021. Vol. 20. No. 14. P. 1–16. DOI:10.1186/s12942-021-00268-y.

АВТОРЫ

Пучин Вячеслав Александрович
ФГБОУ ВО «Московский государственный университет геодезии и картографии» (МИИГАиК), Москва, Россия
кафедра информационно-измерительных систем,
факультет геоинформатики и информационной безопасности
бакалавр
0000-0003-0444-9502

Хлытина Влада Андреевна
ФГБОУ ВО «Московский государственный университет геодезии и картографии» (МИИГАиК), Москва, Россия
кафедра информационно-измерительных систем, факультет геоинформатики и информационной безопасности
бакалавр
vladakh0702@mail.ru
0000-0002-6096-7104

Поступила 05.06.2023. Принята к публикации 23.06.2023. Опубликована 30.06.2023.

UDC 004.62
DOI:10.30533/scidata-2023-14-14

The problem of information security of spatial data as personal data and possible ways to solve it

Vyacheslav A. Puchin[1], Vlada A. Khlytina[1]

AFFILIATIONS

[1] Moscow State University of Geodesy and Cartography, Moscow, Russia
putch2016@inbox.ru

CITATION

Puchin VA, Khlytina VA. The problem of information security of spatial data as personal data and possible ways to solve it. *Spatial Data: science, research and technology.* 2023;14(1): 18–29. DOI:10.30533/scidata-2023-14-14.

KEYWORDS

geomasking, geofencing, geolocation API, depersonalization, point aggregation, areal aggregation, adjusting coordinates, coordinate replacement, FOAM

ABSTRACT

Spatial data becomes personal in cases where, in combination with other data, they allow the data subject to be uniquely identified. In this regard, it is necessary to clearly distinguish in which cases such data should be considered personal and protected as personal, and in which this is not necessary. The purpose of the research was to determine a set of necessary requirements for the protection of spatial data in the context of personal data when developing systems related to location determination. To achieve this goal, a classification of systems that process spatial data has been carried out, approaches to protecting such data have been studied, and effective measures to protect them have been justified. Various methods of geomasking and depersonalization are considered as methods of data protection. As a result of the research, information security methods for selected classes of systems are justified and rationally

distributed. A set of necessary requirements for the protection of spatial data when developing systems related to location determination has been identified, and directions for further research in this area have been identified.

REFERENCES

1. Redlich S. Quantitative Analysis of Geomasking Methods. Dr. of Sci. thesis. Essen: 2022. 318 p.
2. Kounadi O, Leitner M. Adaptive Areal Elimination (AAE): A tansparent way of disclosing protected spatial datasets. ***Computers, Environment and Urban Systems.*** 2016;57: 59–67. DOI:10.1016/j.compenvurbsys.2016.01.004.
3. Leitner M, Curtis A. Cartographic Guidelines for Geographically Masking the Locations of Confidential Point Data. ***Cartographic Perspectives.*** 2004;49(7): 22–39. DOI:10.14714/CP49.439.
4. Allshouse WB., Fitch MK, Hampton KH, et al. Geomasking sensitive health data and privacy protection: an evaluation using an E911 database. ***Geocarto International.*** 2010;25(6): 443–452. DOI:10.1080%2F10106049.2010.496496.
5. Kroll M, Schnell R. Anonymisation of geographical distance matrices via Lipschitz embedding. ***International Journal of Health Geographics.*** 2016;15(1): 1–14. DOI:10.1186/s12942-015-0031-7.
6. Schnell R, Klingwort J, Farrow JM. Locational privacy-preserving distance computations with intersecting sets of randomly labeled grid points. **International Journal of Health Geographics.** 2021;20(14): 1–16. DOI:10.1186/s12942-021-00268-y.

AUTHORS

Vyacheslav A. Puchin
Moscow State University of Geodesy and Cartography, Moscow, Russia
Department of Information and Measuring Systems, Faculty of Geoinformatics and Information Security
Bachelor
0000-0003-0444-9502

Vlada A. Khlytina
Moscow State University of Geodesy and Cartography, Moscow, Russia
Department of Information and Measuring Systems, Faculty of Geoinformatics and Information Security
Bachelor
vladakh0702@mail.ru
0000-0002-6096-7104

Submitted: June 05, 2023. Accepted: June 23, 2023. Published: June 30, 2023.