**Лаба 1**

**ФИО**

Непряхин А. C. (группа 6114-100503D)

**Topic**

Anonymization; K-Anonymity; Microaggregation

Описание предметной области

Данный топик охватывает исследования в области защиты конфиденциальности в интернете при помощи различных подходов, таких как K-Anonymity и Microaggregation. K-Anonymity измеряет вероятность использования конфиденциальных данных третьими лицами для извлечения личной информации. Microaggregation - метод статистического контроля раскрытия данных.

Недостаток (Gap)

В статьях подробно описывается получение методов защиты конфиденциальности(K-Anonymity и Microaggregation) при помощи различных формул и алгоритмов. Не хватает примеров применения данных методов на практике. Работают ли данные методы, есть ли преимущество данных методов от других существующих?

Идея

Применить данные методы для частной жизни, для крупных организаций(учебные, медицинские учреждения и т.д.); Проанализировать работоспособность и эффективность данных методов и сравнить с другими существующими методами.

**Краткий текст обзора**

В современном мире, в связи с интенсивным развитием технологий, всё больше интернет ресурсов запрашивают разрешение на использование конфиденциальной информации. Мошенники стремятся получить эти данные с помощью различных обманных способов. В связи с этим необходимо следовать правилам защиты конфиденциальности для предотвращения попадания ее к третьим лицам, так как они могут использовать информацию в личных целях с нанесением морального и физического вреда человеку, у которого были украдены данные. Существуют различные методы, с помощью которых можно безопасно делиться конфиденциальной информацией в интернет ресурсах. Среди них можно выделить методы K-Anonymity и Microaggregation, о которых рассказывается в текстах El Emam Khaled и Oganian. Наибольший всплеск публикаций по защите конфиденциальности был в период с 2001 по 2008 года. В статьях Meyerson и Sweeney, связанных с методами защиты личной информации подробно указываются формулы и алгоритмы, по которым создаются данные подходы. В статье автора Rubinstein рассказывается об анонимности в социальных сетях и возможном ущербе для личной жизни из-за неаккуратного взаимодействия с интернет ресурсами.

Несмотря на указание рисков автором Rubinstein, связанных с использованием интернет ресурсов и подробное описание получения методов защиты конфиденциальности, указанных авторами Bayardo и Agrawal, данным публикациям не хватает примеров использования данных подходов на практике. Литература не дает понять, эффективны ли указанные методы для предотвращения кражи личной информации. Для получения практического примера применения данных подходов, можно начать использовать их в частной жизни и в различных учреждениях, где необходимо обеспечивать защиту конфиденциальной информации, например в учебных и медицинских. Полученный опыт использования данных методов K-Anonymity и Microaggregation, о которых рассказывается в работах Friedman и Oganian, поможет в анализе их работоспособности и эффективности. Также, после использования указанных подходов на практике, появится возможность сравнить методы K-Anonymity и Microaggregation, о которых рассказывается в текстах El Emam Khaled и Domingo-Ferrer, с другими существующими методами защиты конфиденциальности, например с t-closeness или differential privacy.

Таким образом, благодаря проведенным мероприятиям, станет понятно, имеется ли смысл использовать методы защиты конфиденциальности K-Anonymity и Microaggregation, о которых рассказывается в текстах El Emam Khaled и Oganian.

References

[1] Bayardo, R.J., и R. Agrawal. «Data Privacy through Optimal K-Anonymization». In 21st International Conference on Data Engineering (ICDE’05), 217–28. Tokyo, Japan: IEEE, 2005. <https://doi.org/10.1109/ICDE.2005.42>. [2] Domingo-Ferrer, Josep. «Microaggregation for Database and Location Privacy». В Next Generation Information Technologies and Systems, under redaction Opher Etzion, Tsvi Kuflik, and Amihai Motro, 4032:106–16. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. <https://doi.org/10.1007/11780991_10>. [3] El Emam, Khaled, and Fida Kamal Dankar. «Protecting Privacy Using K-Anonymity». Journal of the American Medical Informatics Association 15, num. 5 (1 of the september 2008): 627–37. <https://doi.org/10.1197/jamia.M2716>. [4] Friedman, Arik, Ran Wolff, и Assaf Schuster. «Providing K-Anonymity in Data Mining». The VLDB Journal 17, num. 4 (July 2008): 789–804. <https://doi.org/10.1007/s00778-006-0039-5>. [5] Ghinita, Gabriel, Panagiotis Karras, Panos Kalnis, and Nikos Mamoulis. «Fast Data Anonymization with Low Information Loss», 28 of the september 2007, 758–69. [6] Meyerson, Adam, and Richard Williams. On the Complexity of Optimal K-Anonymity. Proceedings of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems. Т. 23, 2004. <https://doi.org/10.1145/1055558.1055591>. [7] Nergiz, Mehmet Ercan, Christopher Clifton, and Ahmet Erhan Nergiz. «Multirelational k-Anonymity». IEEE Transactions on Knowledge and Data Engineering 21, num. 8 (august 2009): 1104–17. <https://doi.org/10.1109/TKDE.2008.210>. [8] Oganian, Anna, and Josep Domingo-Ferrer. «On the Complexity of Optimal Microaggregation for Statistical Disclosure Control». Statistical Journal of the United Nations Economic Commission for Europe 18, num. 4 (28 of the december 2001): 345–53. <https://doi.org/10.3233/SJU-2001-18409>. [9] Rubinstein, Ira S, and Woodrow Hartzog. «Anonymization and Risk». WASHINGTON LAW REVIEW 91 (6 of the January 2016): 704–11. [10] Sweeney, Latanya. «ACHIEVING k-ANONYMITY PRIVACY PROTECTION USING GENERALIZATION AND SUPPRESSION». International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 10, num. 05 (october 2002): 571–88. <https://doi.org/10.1142/S021848850200165X>. [11] Zhou, Bin, Jian Pei, and WoShun Luk. «A Brief Survey on Anonymization Techniques for Privacy Preserving Publishing of Social Network Data». ACM SIGKDD Explorations Newsletter 10, num. 2 (20 of the december 2008): 12–22. <https://doi.org/10.1145/1540276.1540279>.