

A NEW APPROACH TO DEPERSONALIZATION OF PERSONAL DATA

*A. Bogdanov*¹, *N. Shchegoleva*^{1,2}, *G. Dik*¹, *V. Khvatov*³,
*J. Kiyamov*¹, *A. Dik*^{1,*}, *V. Tregubov*¹

¹ Saint Petersburg State University, Saint Petersburg, Russia

² Saint Petersburg State Marine Technical University, Saint Petersburg, Russia

³ DGT Technologies AG., Toronto, Canada

Ensuring the confidentiality and protection of personal information in Big Data is an important aspect in data processing. One of the effective methods to achieve a high level of protection is depersonalization of data. An overview of modern methods of preserving personal data when conducting various kinds of research, in business analytics, etc., is presented. To reduce the probability of data de-identification, a hashing method based on the use of the Kessak-256 hash function and the addition of a dynamic random string for each dataset element is proposed. This method allows you to significantly increase the time of hacking and the amount of resources required by the attacker. This approach can be used for secure data transmission, exchange and storage.

Обеспечение конфиденциальности и защиты личной информации в больших данных является важным аспектом при обработке данных. Одним из эффективных методов, позволяющих достичь высокого уровня защиты, является обезличивание данных. Представлен анализ современных методов сохранения персональных данных (ПДн), полученных в результате различных исследований. Для уменьшения вероятности деобезличивания данных предложен метод хеширования, основанный на использовании хеш-функции «Кессак-256», а также добавления динамической случайной строки для каждого элемента датасета. Данный метод позволяет существенно увеличить время взлома и количество требуемых злоумышленнику ресурсов, что в итоге целесообразно использовать для защищенной передачи, обмена и хранения ПДн.

PACS: 89.20.Ff; 07.05.Tr

* E-mail: a.dik@spbu.ru