УДК 004.056:336.71(076)

ОБЕЗЛИЧИВАНИЕ ПЕРСОНАЛЬНЫХ ДАННЫХ НА ОСНОВЕ МЕТОДА ВВЕДЕНИЯ ИДЕНТИФИКАТОРОВ

А.М. Тимофеев[1], М.А. Тавгень[2], А.С. Янковец[2]
[1]Учреждение образования «Белорусский государственный университет информатики и радиоэлектроники», Минск, Беларусь
[2]Учреждение образования «Национальный детский технопарк», Минск, Беларусь

Аннотация. Применительно к информационным системам, в которых осуществляется автоматизированная обработка персональных данных с обеспечением информационной безопасности персональных данных за счет реализации методов их обезличивания, разработаны структурные схемы блоков обезличивания и деобезличивания персональных данных. Предложенные в работе схемы построены на базе симметричного блочного алгоритма ГОСТ 28147-89, являющегося одним из обязательных алгоритмов криптографического преобразования данных в соответствии с требованиями законодательства Республики Беларусь в сфере защиты информации. Особенности построения и функционирования разработанных схем характеризуются применением криптографических и криптоподобных операций, не требующих знания таблиц соответствия при выполнении процедуры деобезличивания персональных данных, что упрощает реализацию системы защиты информации, по сравнению с существующими схемами.

Ключевые слова: информационные системы; персональные данные; защита информации; обезличивание персональных данных; методы обезличивания персональных данных; метод введения идентификаторов.

DEPERSONALIZATION OF PERSONAL DATA BASED ON THE METHOD OF INTRODUCING IDENTIFIERS

[1]A. Timofeev, [2]M. Tavgen, [2]A Yankovets
[1]Education Institution "Belarusian State University of Informatics and Radioelectronics", Minsk, Belarus
[2] Education Institution "National Children's Technopark", Minsk, Belarus

Abstract. Structural diagrams of personal data depersonalization and depersonalization blocks have been developed. They can be used in information systems that perform automated processing of personal data while ensuring their information security based on depersonalization methods. The circuits are based on the symmetric block algorithms of GOST 28147-89. This standard is one of the mandatory algorithms for cryptographic data transformation in accordance with the requirements of the legislation of the Republic of Belarus in the field of information security. The basic principles of construction and operation of the developed diagrams are

described. They are characterized by the use of cryptographic and crypto-like operations that do not require knowledge of correspondence tables when performing the procedure of depersonalization of personal data. This simplifies the implementation of the information security system. compared to existing diagrams.
Keywords: information systems: personal data; information protection; depersonalization of personal data; methods of depersonalization of personal data; method of introducing identifiers.

Введение

Весьма важной задачей в сфере защиты информации является обеспечение информационной безопасности персональных данных, к которым относят любую информацию об идентифицированном физическом лице или физическом лице, которое может быть идентифицировано [1–3].

Под физическим лицом, которое может быть идентифицировано, понимают физическое лицо, которое может быть прямо или косвенно определено, в частности через фамилию, собственное имя, отчество, дату рождения, идентификационный номер либо через один или несколько признаков, характерных для его физической, психологической, умственной, экономической, культурной или социальной идентичности.

В соответствии с требованиями законодательства Республики Беларусь определены методы обезличивания персональных данных, которые предусматривают выполнение таких действий по отношению к персональным данным, в результате которых становится невозможным без использования дополнительной информации определить принадлежность персональных данных конкретному субъекту персональных данных – физическому лицу, в отношении которого осуществляется обработка персональных данных.

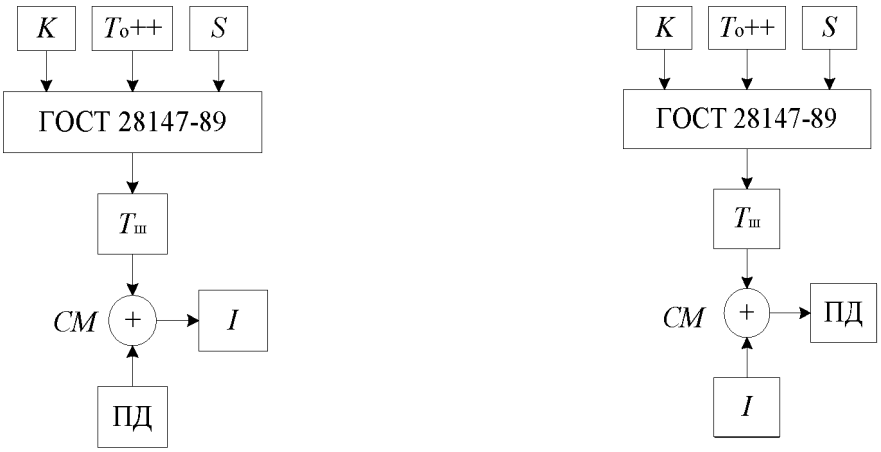
Одним из методов обезличивания персональных данных является метод введения идентификаторов, сущность которого заключается в замене персональных данных или части персональных данных, позволяющих идентифицировать субъекта персональных данных, их идентификаторами и создании таблицы соответствия с последующим раздельным хранением идентификаторов и таблиц. Однако известные реализации процедур обезличивания персональных данных на основе метода введения идентификаторов не позволяют выполнить деобезличивание персональных данных без знания указанных выше таблиц [1–3]. Это усложняет существующие реализации, поскольку таблицы соответствия являются ключевыми компонентами информационных систем и должны сохраняться в секрете. В связи с этим целью данной работы являлось разработать структурные схемы, позволяющие обезличивать и деобезличивать персональные данные на базе метода введения идентификаторов, которые упрощают известные схемы за счет выполнения процедуры деобезличивания персональных данных без знания таблиц соответствия.

Объектом исследования являлся стандарт ГОСТ 28147-89 – блочный шифр с 256-битным ключом и 32 циклами преобразования, оперирующий 64-битными блоками. ГОСТ 28147-89 выбран в качестве объекта исследования, поскольку является одним из обязательных стандартов для предприятий и организаций, осуществляющих криптографическую защиту информации в соответствии с требованиями законодательства Республики Беларусь.

Предметом исследования являлось использование ГОСТ 28147-89 в режиме гаммирования для решения задач обезличивания и деобезличивания персональных данных на базе метода введения идентификаторов.

Структурные схемы блоков обезличивания и обезличивания персональных данных

На рисунке приведены структурные схемы блоков обезличивания и обезличивания персональных данных.



a *b*

Структурные схемы блоков обезличивания (a) и деобезличивания (b) персональных данных
Block diagrams of the depersonalization (a) and depersonalization (b) blocks of personal data

Схема блока обезличивания персональных данных функционирует следующим образом (см. часть *a* рисунка). Персональные данные ПД, закодированные на базе кодировочной таблицы UTF-16, разбивают на блоки длиной 64 бита. Затем первый блок ПД длиной 64 бита подают на первый вход сумматора СМ, на второй вход которого поступает первый блок $T_{ш}$ длиной 64 бита, выработанный в соответствии с требованиями, определяемыми ГОСТ 28147-89 в режиме гаммирования с использованием секретного криптографического ключа K, блоков открытого текста T_o и синхропосылки S. Результат с выхода сумматор СМ образует первый блок идентификатора. Процедуру выработки др. блоков идентификаторов выполняют аналогичным образом, за исключением того, что для каждого последующего идентификатора предыдущий блок T_o инкрементируют. Полученные таким образом блоки идентификаторов конкатенируют в общую последовательность, являющуюся выработанным идентификатором.

Отметим, что в качестве блочного шифра может быть использован любой другой алгоритм блочного типа, например, AES, СТБ 34.101.31-2020 и пр. При этом принципы, изложенные выше, останутся неизменными. Аналогичным образом кодировочная таблица UTF-16, применяемая в разработанной схеме, может быть заменена на любую другую кодировочную таблицу, обеспечивающую однозначный перевод символов персональных данных в двоичный код. Важно отметить, что алгоритм блочного типа и кодировочная таблица должны быть выбраны однотипными для блоков обезличивания и деобезличивания персональных данных.

Схема блока деобезличивания персональных данных (см. часть *b* рисунка) функционирует схожим образом, как и схема обезличивания персональных данных, за исключением следующего. В схеме деобезличивания персональных данных на первый вход сумматора СМ подают блоки идентификаторов, а на выходе СМ получают двоичную последовательность, которая с помощью кодировочной таблицы UTF-16

преобразуется в блоки персональных данных ПД. Это позволяет, реализовав один пакет программного обеспечения для обезличивания персональных данных, использовать этот же пакет программного обеспечения для деобезличивания персональных данных. Причем при выполнении обезличивания персональных данных входной информацией будут является непосредственно персональные данные, а при деобезличивании – идентификаторы, что упрощает реализацию системы защиты информации.

Важно отметить, что схема деобезличивания персональных данных, предложенная в настоящей работе, не требует использования и обязательного хранения таблиц соответствия, что выгодно отличает данную схему, по сравнению с существующими.

Заключение

При реализации метода введения идентификаторов чрезвычайно важно определить алгоритмы, отдельные параметры алгоритмов и используемые математические вычисления, которые позволят не только преобразовывать персональные данные к виду идентификаторов (обезличивать их), но и вычислительно выполнить обратную процедуру, т.е. на основе идентификаторов рассчитывать персональные данные.

Предложены структурные схемы блоков обезличивания и деобезличивания персональных данных, построенные на базе симметричного блочного шифра ГОСТ 28147-89 в режиме гаммирования. Схема деобезличивания персональных данных не требует хранения и использования таблиц соответствия при осуществлении процедуры деобезличивания, что упрощает ее реализацию, по сравнению с существующими.

Список использованных источников

1. Ворона. В. А. (2023) *Биометрическая идентификация личности.* Москва. Горячая линия-Телеком.
2. Коллинз. М. (2020) *Защита сетей. Подход на основе анализа данных.* Москва. ДМК Пресс.
3. Остапенко, Г. А. (2020) *Информационные операции и атаки в социотехнических системах: организационно-правовые аспекты противодействия.* Москва, Горячая линия-Телеком.

References

1. Vorona V. A. (2023) *Biometric Identification of Personality*. Moscow. Goryachaya Liniya-Telecom (in Russian).
2. Collins M. (2020) *A Data-Based Approach*. Moscow. DMK Press (in Russian).
3. Ostapenko G. A. (2020) *Information Operations and Attacks in Socio-Technical Systems: Organizational and Legal Aspects of Counteraction*. Moscow. DMK Press (in Russian).

Сведения об авторах

Тимофеев А.М., канд. техн. наук. доц.,
доц. каф. защиты информации. учреждение образования «Белорусский государственный университет информатики и радиоэлектроники»,
tamvks@mail.ru.
Тавгень М.А, учащийся. учреждение образования «Национальный детский технопарк»,
makdisp2@gmail.com
Янковец А.С., учащийся. учреждение образования «Национальный детский технопарк»,
tomraid3301@gmail.com

Information about the authors

Timofeev A., Cand. Sci. (Tech.). Assosiate Professor. Associate Professor of the Department of Information Protection. Educational Institution "Belarusian State University of Informatics and Radioelectronics",
tamvks@mail.ru.
Tavgen M., Student. Educational Institution "National Children's Technopark".
makdisp2@gmail.com.
Yankovets A., Student. Educational Institution "National Children's Technopark".
tomraid3301@gmail.com.