Труды научно-технической конференции кластера пензенских предприятий,
обеспечивающих **БЕЗОПАСНОСТЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ.**
Том 10, Пенза-2016, с. 74-77 (http://пниэи.рф/activity/science/BIT/T10-p74.pdf)

УДК: 519.2, 612.087, 621.319.7

Павлов М.А., Секретов М.В.

**Обезличивание персональных данных пациентов с использованием «нечетких экстракторов» или нейросетевых преобразователей биометрии в длинный код перед их размещением в облачных хранилищах**

В настоящее время активно идут процессы информатизации современного общества, как итог, наша персональная информация постепенно перемещается в Интернет облака. Ярким примером общего вектора развития являются медицинские информационные системы. В 2006 году в России был введен в действие отечественный стандарт, регламентирующий требования к электронной истории болезни [1], это позволило разработать типовую медицинскую информационную систему [2]. Далее, встал вопрос о переходе к использованию типового электронного места врача [3]. В итоге, информационная технология уже позволяет создавать интегрированные электронные медицинские карты [4], которые могут размещаться как на локальном сервере медицинской информационной системы, так и на Интернет серверах поставщиков облачных услуг.

Естественно, что эта общая тенденция порождает новые угрозы информационной безопасности, которые должны быть устранены с учетом уже сложившейся технической практики [5] и национального законодательства [6]. За рубежом проблема решается через биометрическую аутентификацию личности человека с использованием, так называемых, «нечетких экстракторов» [7]. В России для этих же целей используются искусственные нейронные сети [8], применяя которые можно осуществлять обезличивание [9] медицинских электронных документов, в случае их размещения в облачных хранилищах.

Очевидно, что обезличивание биометрической информации может быть осуществлено как с использованием «нечетких экстракторов», так и с использованием нейросетевых преобразователей биометрия-код. В случае, когда используются «нечеткие экстракторы», алгоритм обезличивания представлен на рисунке 1.

Полученные персональные данные пациента обезличиваются с помощью одного или нескольких методов: метода введения идентификаторов, метода изменения состава или семантики, метода декомпозиции, метода перемешивания. В результате образуются идентифицирующие и обезличенные данные. Биометрические данные обезличиваются с использованием нечеткого экстрактора. Для чего пациент предъявляет несколько примеров биометрического образа «Свой». Затем вычисляются вектора двоичных биометрических параметров. Далее создается случайный ключ, который затем преобразуется в двоичный вектор с использованием помехоустойчивого кода, например, кода Боуза-Чоудхури-Хоквингема. Далее, двоичный вектор шифруется гаммированием, а в качестве гаммы используется вычисляемый двоичный вектор биометрических параметров. В результате создается защищенный биометрический контейнер, содержащий идентификатор (псевдоним) пациента, результат гаммирования и значение хеш-функции от личного ключа (кода доступа) пациента. Содержимое биометрического контейнера – это защищенные

персональные данные, которые прикрепляются к обезличенным иным персональным данным (например, к истории болезни пациента).

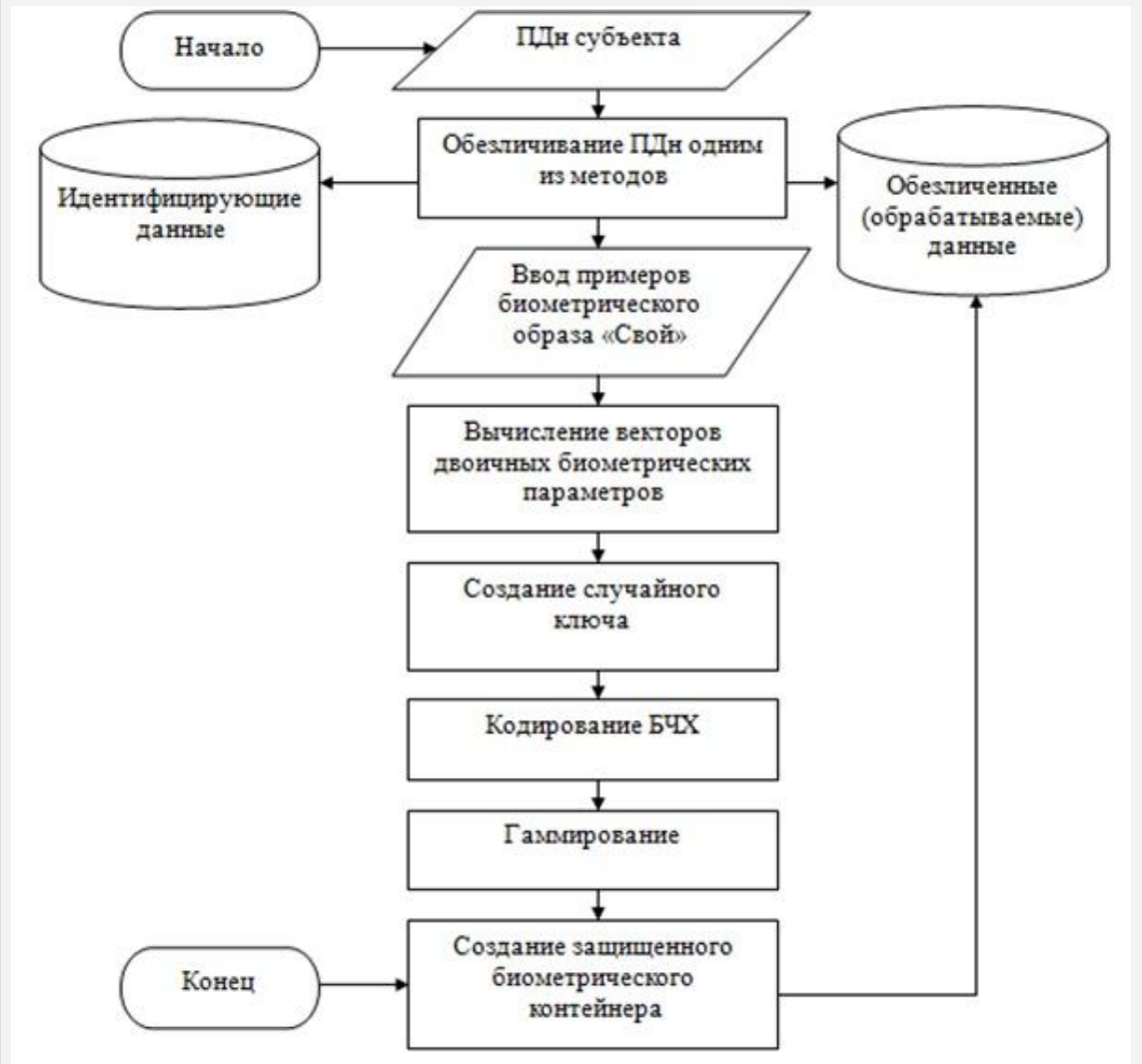


Рис. 1 Алгоритм обезличивания персональных данных с использованием нечеткого экстрактора

Алгоритм обезличивания персональных данных с использованием нейросетевого преобразователя биометрия-код представлен на рисунке 2.

Алгоритм отличается тем, что пациент предъявляет несколько примеров биометрического образа «Свой», тем самым подготавливая обучающую выборку для искусственной нейронной сети. Затем вычисляются вектора биометрических параметров и автоматически обучается искусственная нейронная сеть на случайно сгенерированном двоичном ключе. По окончании обучения создается защищенные биометрический контейнер, содержащий идентификатор (псевдоним) пациента, матрицу весовых коэффициентов и значение хеш-функции от личного ключа (кода доступа) пациента. Содержимое биометрического контейнера включается в обезличенные данные, обрабатываемые оператором информационной системы персональных данных.

Алгоритм отличается тем, что пациент предъявляет несколько примеров биометрического образа «Свой», тем самым подготавливая обучающую выборку для искусственной нейронной сети. Затем вычисляются вектора биометрических параметров и автоматически обучается искусственная нейронная сеть на случайно сгенерированном двоичном ключе. По окончании обучения создается

защищенные биометрический контейнер, содержащий идентификатор (псевдоним) пациента, матрицу весовых коэффициентов и значение хеш-функции от личного ключа (кода доступа) пациента. Содержимое биометрического контейнера включается в обезличенные данные, обрабатываемые оператором информационной системы персональных данных.

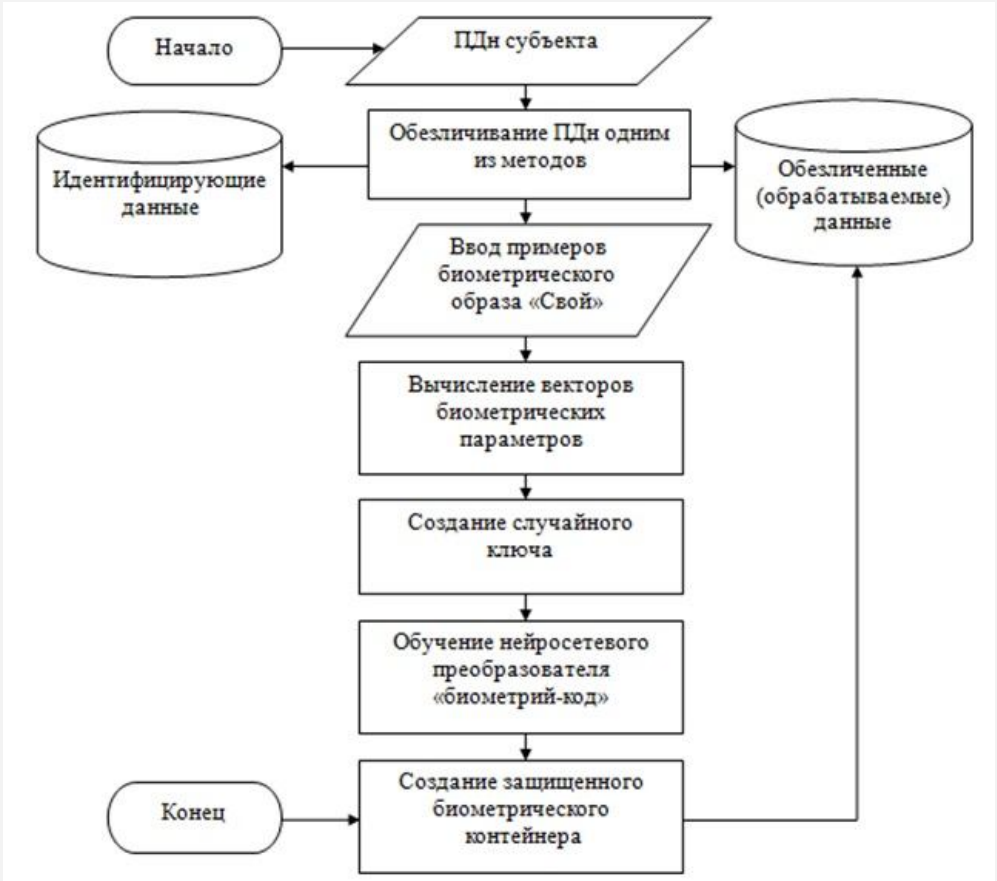


Рис. 2 - Алгоритм обезличивания персональных данных с использованием нейросетевого преобразователя

Так как идентифицирующие данные содержат сведения, по которым пациент может быть однозначно идентифицирован, их утечка может привести к разглашению информации о факте обращения за медицинской помощью, которая относится к врачебной тайне. Поэтому к идентифицирующим данным должны предъявляться более строгие требования по безопасности, в частности, полное ограничение доступа всех сотрудников оператора к такого рода информации.

Обезличенные данные не позволяют однозначно идентифицировать пациента, тем самым обеспечивается анонимность обрабатываемых сведений. Таким образом, возможная утечка обезличенных данных не приведет к разглашению информации о факте обращения за медицинской помощью пациента и его диагнозе. Тогда возникает вопрос: как изменять обезличенные данные и вносить новые сведения, касающиеся конкретного пациента?

Чтобы изменять обезличенные данные и вносить новые сведения необходимо определить принадлежность обезличенных данных конкретному пациенту. Это можно выполнить двумя способами. Первый способ заключается в использовании идентифицирующих данных и проведении процедуры деобезличивания. Второй способ заключается в использовании биометрического контейнера и проведении биометрической аутентификации (подтверждения личности). При использовании первого способа вновь возникает проблема,

связанная с возможной утечкой врачебной тайны. Однако, второй способ лишен подобного недостатка, так как позволяет определять принадлежность обезличенных данных конкретному пациенту без доступа к идентифицирующим данным, следовательно, становится возможным выполнение требования полного ограничения доступа сотрудников оператора к идентифицирующим данным.

В связи с тем, что интегрированная электронная медицинская карта [4] пациента может храниться в облачном хранилище, персональные данные в ней должны быть обезличены. Как показано в данной статье процедуры обезличивания персональных данных могут быть реализованы как при использовании «нечетких экстракторов», так и при использовании нейросетевых преобразователей биометрия-код. При выполнении защиты персональных данных обезличиванием, оператор, обрабатывающий их (в том числе оператор, предоставляющий облачный сервис по хранению данных), не может нанести ущерб пациентам. Злоумышленник, похитивший электронную базу обезличенных медицинских карт, не может воспользоваться этой информацией, так как он не знает, чья это персональная информация. При ограничении доступа к идентифицирующим пациентов данным, злоумышленник не сможет выполнить деобезличивание, даже являясь сотрудником оператора.

Используя биометрию пациент всегда может доказать врачу, что это именно его интегрированная электронная медицинская карта [4]. Кроме того, пациенту будет невозможно выдать себя за другого человека, так как биометрические параметры неразрывно связаны с лицом и их невозможно передать либо подделать.

ЛИТЕРАТУРА:


1. ГОСТ Р 52636-2006. Электронная история болезни. Общие положения.
2. Федеральная типовая медицинская информационная система (ФТМИС). Разработчик «Крокус Консалдинг» 2008 г., государственный контракт по ФЦП «Электронная Россия (2002-2010 годы).
3. Электронное рабочее место врача. Руководство пользователя. Москва-2014 г.
4. Зингерман Б.В., Шкловский-Корди Н.Е., Карп В.П., Воробъев А.И. Интегрированная электронная медицинская карта: задачи и проблемы. //Врач и информационные технологии. №1, 2015 г., с. 24-27.
5. Костков Д. Защита облачных вычислений: общие международные подходы. //Первая миля. №8, 2015 г.
6. Федеральный закон «О персональных данных» от 27.07.2006 № 152.
7. Dodis Y., Reyzin L., Smith A. Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy // Proc. EUROCRYPT, April 13, pages 523-540, 2004.
8. ГОСТ Р 52633.0-2006 «Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации».
9. Методические рекомендации по применению приказа Роскомнадзора от 5 сентября 2013 г. № 996 «Об утверждении требований и методов по обезличиванию персональных данных».



Статья поступила 09.12.2016, опубликована 12.12.2016
по положительной рецензии д.т.н. Иванова А.И.