УДК 004.056.5 DOI: 10.14529/ctcr150303

КОЛИЧЕСТВЕННЫЙ АНАЛИЗ ПРОЦЕДУРЫ ОБЕЗЛИЧИВАНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ. МЕТОД ВВЕДЕНИЯ ИДЕНТИФИКАТОРОВ

Е.Ю. Мищенко, А.Н. Соколов

Южно-Уральский государственный университет, г. Челябинск

Обезличивание – способ обработки персональных данных, целью которого является приведение этих данных в защищенное состояние, которое не позволяет злоумышленнику использовать их во вред физическому лицу. Результат обезличивания персональных данных зависит от их содержания и применяемого метода обезличивания. Нормативные акты определяют несколько методов обезличивания, но все они описываются качественными критериями. В статье производится количественный анализ одного из методов обезличивания – метода введения идентификаторов. Предлагается вариант технической реализации данного метода, включая решение проблемы необходимого и достаточного идентификационного набора атрибутов таблицы соответствий, определение требований к связующему идентификатору, а также рассмотрение возможных способов связи таблицы соответствий с обезличенными данными. На основе реального примера производится оценка эффективности метода по различным критериям. В том числе по техническим критериям (невозможность идентификации, с одной стороны, и возможность деобезличивания с применением имеющихся дополнительных данных, с другой стороны), а также по экономическим критериям (окупаемость). На базе показателей вероятности идентификации и степени обезличивания персональных данных приводятся рекомендации по повышению эффективности данного метода обезличивания персональных данных.

Ключевые слова: персональные данные, обезличивание персональных данных, метод введения идентификаторов.

В статье [1] проанализирована схема идентификации персональных данных (ПД), как процедуры проверки эффективности обезличивания ПД, рассмотрено воздействие потенциального злоумышленника на область поиска (БД) с целью получения результата идентификации. Введены количественные критерии полученного результата – вероятность идентификации (ВИ) и степень обезличивания (СО). Целью данной статьи является определение значений количественных критериев для одного из методов обезличивания ПД – метода введения идентификаторов.

Независимо от метода обезличивания ПД, полученная в результате обезличенная база данных, будучи доступной без ограничений, должна нивелировать (обесценить) попытки злоумышленника использовать любые обезличенные данные для нанесения вреда (компрометации, обмана, шантажа) конкретному физическому лицу. Возможный вред от уничтожения обезличенных данных мы не рассматриваем, так как его можно компенсировать иными (технологическими) методами.

Для оператора (обработчика) ПД обезличивание будет иметь смысл, если затраты на обезличивание будут явно меньше затрат на средства и меры защиты ПД при их обработке в рамках автоматизированной ИСПДн. «Явно» – это не меньше, чем в 2 раза (из-за погрешности расчетов).

Для оценки затрат на обезличивание важно понимать, что обезличивание ПД не избавляет оператора от затрат на защиту ПД полностью. На каких-то рабочих местах необходимо эти ПД обезличивать и деобезличивать, где-то – обрабатывать ПД в явном виде (поиск, вывод документов на печать). Эти рабочие места должны быть защищены как составные части ИСПДн. Поэтому общие затраты на обезличивание сложатся, с одной стороны, из затрат на защиту указанных рабочих мест, и с другой стороны, из затрат на модернизацию структуры БД и технологического процесса обработки ПД.

Соотношение этих двух групп затрат зависит от технологической цели обезличивания:

1) обезличивание для дальнейшей передачи по каналам связи (в частном случае – для хранения на внешних носителях). В этом случае затраты на защиту рабочих мест остаются, но эконо-

мятся затраты на защиту каналов связи (до 50 % затрат на средства защиты информации!). Затраты на модернизацию структуры БД (на входе/выходе) от цели не зависят, затраты на модернизацию технологического процесса – минимальные;

2) обезличивание для обработки в таком виде на большей части рабочих мест. В этом случае затраты на защиту большей части (до 90 %) рабочих мест и на защиту каналов связи экономятся, затраты на модернизацию структуры БД и модернизацию технологического процесса – максимальные.

При детальном рассмотрении может оказаться, что в зависимости от технологической цели тот или иной метод обезличивания является более эффективным экономически.

1. Описание метода

В соответствии с Приказом Роскомнадзора [2] метод введения идентификаторов реализуется путем замены ПД, позволяющих идентифицировать субъекта, их идентификаторами и созданием таблицы соответствия. То есть после применения данного метода единая база (БД) распадется на две базы:

1) таблица соответствия, в которой некий набор идентифицирующих физическое лицо (ФЛ) атрибутов однозначно сопоставляется с неким абстрактным атрибутом. Выражаясь терминами статьи [1], для каждого ФЛ набор значимых атрибутов идентификации соответствует некоему служебному уникальному идентификатору. Причем объем этой базы (обозначим его ОБТ) равен количеству ФЛ;

2) база прочих данных, в которой некоему служебному идентификатору (не уникальному, так как встречается в этой базе много раз) однозначно сопоставляется набор прочих данных – не значимых с точки зрения идентификации, но определяющих суть обработки. Причем объем этой базы (ОБП) может намного превышать значение ОБТ.

В данном процессе разделения БД необходимо решить три проблемы:

1. Какие атрибуты включить в таблицу соответствий.
2. Какими свойствами должен обладать связующий идентификатор.
3. Как обеспечить связь между двумя базами.

1.1. Атрибуты таблицы соответствий

Набор атрибутов, включаемых в таблицу соответствия, в первую очередь должен быть достаточным для однозначной идентификации в нем конкретного ФЛ, то есть интегральный показатель ВИ для данного набора должен быть равен 1. В статье [1] подробно рассмотрены критерии формирования такого набора, и показали, что определяющим критерием является объем базы ОБ. Например, для объема 1 млн записей достаточным для идентификации является набор «фамилия» + «дата рождения».

Но достаточность для идентификации не решает главную задачу – надежное обезличивание прочих данных, оставшихся во второй базе. Например, если среди прочих данных окажутся такие атрибуты как «имя», «адрес проживания», «номер телефона», «место работы», то для некоторых ФЛ такой набор может дать ВИ = 1 даже без атрибута «фамилия».

Следовательно, в таблицу соответствия должны быть включены все атрибуты, по которым **возможно** идентифицировать хотя бы одно ФЛ.

1.2. Требования к связующему идентификатору

Главное требование к связующему идентификатору – уникальность для любого ФЛ. В статье [1] показано, что таким идентификатором не может быть ни один условно значимый (служебный) идентификатор ведомственного типа – ИНН, СНИЛС, номер паспорта, из-за их отсутствия у значительных групп ФЛ и по другим причинам.

Поэтому данный идентификатор должен быть не только уникальным, но и абстрактным по отношению к ФЛ. Кроме того, его длина должна быть одинаковой для всей базы данных.

Кстати, факт использования служебных идентификаторов для идентификации ФЛ внутри конкретных ведомств подтверждает указанные здесь требования. Кроме того, все используемые в БД служебные идентификаторы должны быть включены в таблицу соответствий, как атрибуты, по которым **возможно** идентифицировать ФЛ.

1.3. Связь таблицы соответствий с прочими данными

Любая обработка ПД автоматизированным способом производится в рамках базы данных реляционного типа, что подразумевает создание нескольких таблиц данных, одна часть которых является справочниками (условно постоянные), а другая часть – изменяемые данные функционального характера (переменные). Справочные данные связаны с функциональными данными посредством специальных служебных идентификаторов. Такая структура является технологически наиболее эффективной. То есть после решения вопроса с составом таблицы соответствий и видом связующего идентификатора мы приходим к типичной автоматизированной обработке ПД.

Означает ли это, что обработка ПД в рамках СУБД производится уже в обезличенном виде? Это, конечно, не так. Ведь пока пользователь (и злоумышленник тоже!) имеет доступ к таблице соответствий (справочнику ФЛ), он может связать прочие (функциональные) данные с конкретным ФЛ, значит ПД в этой общей базе не обезличены.

Следовательно, для обезличивания базы прочих данных необходимо отделить ее от таблицы соответствий, после чего таблицу соответствий надо защитить согласно нормативным требованиям (она останется ИСПДн), а базу прочих данных можно открыть для свободного доступа (она станет обезличенной). Термин «отделить» означает либо физически отдельное хранение двух баз данных, либо установку между двумя базами данных межсетевого экрана (сертифицированного на соответствие нормативным требованиям). В первом случае связи между базами не будет совсем, и совместная их обработка возможна только с применением специальных внешних носителей. Во втором случае связь между базами будет односторонняя (со стороны таблицы соответствий), совместная обработка возможна тоже только с одной стороны.

2. Оценка эффективности метода

С точки зрения определения понятия «обезличивание» любой метод обезличивания эффективен настолько, насколько неэффективными окажутся попытки злоумышленника идентифицировать ФЛ в обезличенной базе данных. Данный критерий определен в Приказе[2] термином «анонимность», то есть его можно приравнять к значению степени обезличивания СО, которая связана с ВИ формулой СО = 1 – ВИмакс (для множества попыток идентификации).

Но есть еще один критерий эффективности, определенный Приказом [2], который зависит от метода обезличивания – это «применимость», то есть возможность обработки без предварительного деобезличивания, а в более широком смысле – возможность совместной защищенной обработки комплекса данных, состоящих из обезличенной базы и той самой «дополнительной информации», которая позволяет их деобезличивать. В рамках метода введения идентификаторов такой «дополнительной информацией» является таблица соответствий.

2.1. Эффективность анонимности

Произведем оценку значения СО на конкретном примере: злоумышленник хочет найти ФЛ на основании известной ему информации о его автомобиле (внешнем виде) при условии свободного доступа к обезличенной БД регистрации всех автомобилей нашей страны (предположим, что эта БД обезличена методом идентификаторов, и там есть вся информация об автомобилях, но нет ничего об их владельцах).

Для оценки ВИмакс примем следующие предварительные условия (информация, которую можно получить из открытых источников):

1. Злоумышленнику известен регион, в котором ФЛ эксплуатирует свой автомобиль (средний регион нашей страны с населением 2 млн человек, областной центр с населением 1 млн человек);

2. Возраст ФЛ – от 18 до 60 лет;

3. Количество ДТП в год по региону – 3 тыс. при количестве автомобилей – 800 тыс., по областному центру – 2 тыс. при количестве – 500 тыс.

4. Поскольку мы оцениваем ВИмакс, дадим злоумышленнику преимущество – на поиски у него есть срок 30 дней (назовем данный критерий «актуальность идентификации», он прямо пропорционален ВИ и в реальной жизни его значение – 3 дня).

Из первых двух условий следует, что по возрасту водителями в данном регионе могут быть 1 млн человек, а в областном центре – 500 тыс. (ВИ = 1 / 1 000 000 и ВИмакс = 1 / 500 000).

Третье и четвертое условия позволяют оценить вероятность того, что искомый автомобиль можно будет из-за ДТП обнаружить в ограниченном количестве известных мест (пункты регистрации ДТП, страховые компании, автосалоны по ремонту). Если принять, что в ДТП участвуют 2 автомобиля, то для региона вероятность попадания в ДТП конкретного автомобиля равна $(3/800 = 1/266)\cdot 2 = 1/133$ за год, а за 30 дней («актуальность идентификации») – $1/133/12 = 1/1596$. Для областного центра эта вероятность будет $1/1500$. Но для определения ВИ надо учесть К = «ограниченное количество известных мест». Если в областном центре 5 пунктов регистрации ДТП, то ВИмакс $= 1/1500/5 = 1/7500$. Может ли повысить ВИмакс информация из обезличенной базы?

В состав маркера поиска (МП) войдут атрибуты: НМ1 = «марка», НМ2 = «модель», НМ3 = «цвет кузова», НМ4 = «государственный номер» – их можно надежно определить по внешнему виду. В обезличенной базе регистрации автомобилей присутствуют все эти наименования атрибутов и еще некоторые другие атрибуты (например, «дата регистрации», «место регистрации», «наименование автосалона-продавца» (но не фамилия ФЛ-продавца!), реквизиты договора продажи, свидетельства о регистрации, полиса ОСАГО и др.). Цель злоумышленника будет достигнута при выполнении двух условий:

1. Поиск по заданному маркеру даст достаточно ограниченное количество автомобилей (идентификация автомобиля).

2. Другие реквизиты позволят еще уменьшить это количество, а в идеале помогут определить какие либо значимые реквизиты для идентификации искомого ФЛ.

Первое условие будет выполнено автоматически, если известен государственный номер автомобиля (автомобиль идентифицирован). Если же номер точно не известен (вычисляем интегральную вероятность идентификации для первых трех атрибутов), то для годовалой иномарки в БД записей нужного цвета (всего 5 цветов) будет найдено 600 записей, а для трехлетней отечественной марки (всего 12 цветов) будет найдено 50 000 записей.

Если автомобиль идентифицирован по номеру, то увеличить ВИ могут названия автосалона-продавца и страховой компании, выдавшей полис ОСАГО. Если автомобиль не идентифицирован по номеру, то прочие атрибуты могут значительно увеличить вероятность идентификации автомобиля, то есть определить его государственный номер, но сам по себе он ничего не дает.

Знание количества офисов страховой компании для известного по номеру автомобиля может увеличить ВИ (если такой офис в областном центре один – то ВИ $= 1/1500/2 = 1/3000$, так как есть вероятность 1/2, что ФЛ не является виновником в ДТП, тогда он в страховую компанию не придет). Знание количества салонов также может увеличить ВИ (если такой салон в областном центре один – то ВИ $= 1/1500$), хотя ФЛ может в салон и не обратиться.

Значение ВИмакс $= 1/1500$ (то есть СО $= 1 -$ ВИмакс $= 0{,}9993$) показывает, что реально злоумышленник не сможет идентифицировать ФЛ, а если автомобиль был зарегистрирован в одном регионе, а эксплуатируется в другом, найти ФЛ практически невозможно. Учет реальной актуальности уменьшает ВИ еще в 10 раз. То есть эффективность анонимности данного метода – вне сомнений.

2.2. Эффективность применимости

Решающим критерием применимости рассматриваемого метода является техническая возможность его реализации. А уже при наличии технической возможности определяющую роль играют стоимость и сроки реализации.

На рисунке приведена схема разделения базы ПД на таблицу соответствия и базу прочих данных, где в качестве связующего звена используется межсетевой экран.

Цифрами на рисунке обозначены:

1 – таблица соответствий (сервер ПД, входит в состав ИСПДн, защищается);

2 – обезличенные данные (сервер обезличенной базы, в свободном доступе, не защищается);

3 – рабочее место оператора ПД (входит в состав ИСПДн, защищается);

4 – рабочее место оператора обезличенной базы (в свободном доступе, не защищается);

5 – межсетевой экран, обеспечивает одностороннее движение информации, направление которого указано стрелками.

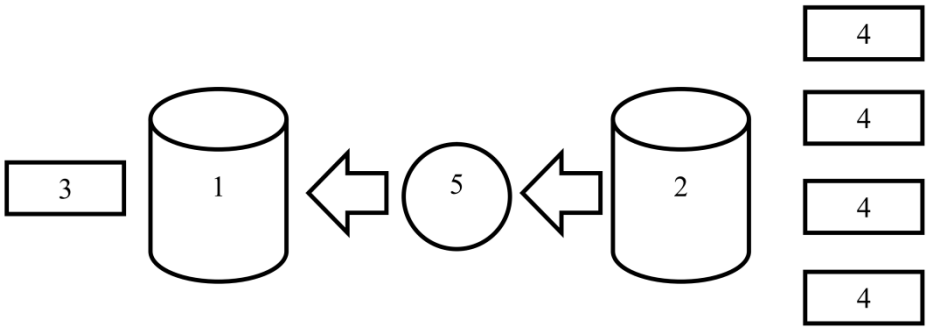

Схема разделения базы ПД

В нашем случае с БД регистрации автомобилей все чувствительные к идентификации ПД владельцев (фамилия, имя, отчество, дата рождения, место рождения, адрес проживания, телефон, номер паспорта) будут храниться и обрабатываться в базе 1, и доступ к ним будет разрешен только пользователям рабочего места 3 (регистратура). А в базе 2 будут храниться и обрабатываться данные о постановке и снятии с учета всех автомобилей любого ФЛ, и доступ к ним будет разрешен не только легальным пользователям рабочих мест 4, но и любым пользователям, которые смогут (в принципе без ограничений) получить физический доступ к базе 2 (либо с рабочего места 4 в отсутствие легального пользователя, либо подключив к базе 2 новое рабочее место).

Анонимность обеспечивается межсетевым экраном, который запрещает потоки информации (ПД) от таблицы соответствий в сторону обезличенной базы, то есть злоумышленник (как и любой пользователь рабочего места 4) не имеет доступа к ПД.

Для определения эффекта применимости показанной на рис.1 системы в целом необходимо оценить применимость ее обеих частей – защищаемой (состоит из сервера 1 и малого количества рабочих мест 3) и не защищаемой (состоит из сервера 2 и большого количества рабочих мест 4).

Пользователь рабочего места 3 может решать как частную, так и общую задачу идентификации, и другие функции:

1) указав значение связующего идентификатора ФЛ, выдать запрос через экран 5, получить все данные конкретного ФЛ из обезличенной базы (сервер 2) и сопоставить их с этим ФЛ (на сервере 1);

2) указав значения атрибутов из обезличенной базы, получить список идентификаторов с заданными значениями атрибутов и сопоставить им ПД из таблицы соответствий;

3) кроме того, только этот пользователь может добавлять записи о новых ФЛ в таблицу соответствий и генерировать для них идентификаторы, а также удалять ПД из таблицы соответствий.

Таким образом применимость (техническая реализация функций) с точки зрения защищаемой системы не вызывает сомнений.

Пользователь рабочего места 4 может выполнять аналогичные рабочему месту 3 функции 1) и 2), но не может сопоставить эту информацию конкретным ФЛ, а функцию 3) он выполнять не может в принципе. Зато функцию ввода существенной информации (ради которой и создана вся система обработки) может выполнять только этот пользователь.

Для оценки эффекта применимости с точки зрения рабочего места 4 надо ответить на три вопроса:

1. На каком основании пользователь 4 при вводе данных ФЛ в базу 2 будет сопоставлять их с идентификатором этого ФЛ (откуда пользователь узнает этот идентификатор, особенно при вводе первой записи об этом ФЛ)?

2. Как снизить вероятность ошибки ввода идентификатора (абстрактного вида!) для получения доступа к обезличенной информации?

3. Как пользователь 4 убедится в соответствии идентификатора данному ФЛ (проблема подлинности)?

Решение вопроса 1 подразумевает предварительную запись ПД ФЛ в базу 1 (регистрация) с генерацией связующего идентификатора, но этого еще не достаточно, поскольку не понятно, как

этот идентификатор попадет из базы 1 в базу 2. В рамках нашего примера при первом посещении (постановка автомобиля на учет) владелец должен прийти к рабочему месту 3, представить свои ПД (показать паспорт) и получить идентификатор. Чтобы ФЛ мог представить идентификатор кому-либо (пользователю рабочего места 4), этот идентификатор должен быть записан на какомнибудь носителе (номер на листке бумаги). После получения идентификатора ФЛ-владелец идет к рабочему месту 4 (к любому из многих), представляет идентификатор и все данные об автомобиле. При повторном посещении (снятие с учета или постановка на учет другого автомобиля) владелец идет с идентификатором сразу к рабочему месту 4. Описанный способ настолько очевиден, что был запатентован в качестве полезной модели (независимо от вида внешнего носителя для идентификатора). Патент № 103414 «Система взаимодействия разделенных баз персональных данных информационной системы».

Решение вопроса 2 полностью определяется видом внешнего носителя для идентификатора. Преимущество листка бумаги с номером только одно – техническая простота реализации. Но для увеличения скорости и устранения возможной ошибки ввода идентификатора-номера пользователем 4 с бумажного листа желательно процедуру ввода автоматизировать. Для этого вместо символьного номера можно использовать штрих-код. Но помимо возникающих технических сложностей есть еще один важный недостаток – ненадежность самого носителя (бумага боится воды и механических повреждений). Устранить этот недостаток можно ламинированием.

В качестве альтернативы номеру на бумажном носителе можно предложить пин-код на пластиковой карте, криптографический ключ на флеш-накопителе или отпечаток на пальце. Использование этих носителей не является более сложным технически, чем чтение-запись штрих-кода, но по надежности, пожалуй, выигрывает пластиковая карта. В рамках нашего примера необходимо в качестве альтернативы упомянуть также один из атрибутов ФЛ, являющийся идентификатором служебного типа, – номер водительского удостоверения. Причем само удостоверение – достаточно качественный носитель для этого идентификатора, но недостаток данного варианта уже обсуждался в предыдущей статье – не все владельцы имеют водительские удостоверения. Кроме того, ввод этого символьного идентификатора на рабочем месте 4 сложно автоматизировать, то есть велика вероятность ошибки ввода.

Есть еще один критерий эффективности для внешнего носителя – он определяется человеческим фактором: легко или сложно такой идентификатор спутать с другим аналогичным (своим или чужим), легко или сложно его потерять. По этому критерию преимуществом обладает только отпечаток пальца – ни спутать, ни потерять его нельзя, но, к сожалению, опыт показывает, что до 10 % ФЛ имеют проблемы со считыванием отпечатка.

Необходимость решения вопроса 3 вытекает из возможности использования чужого идентификатора (ошибочного или преднамеренного). Для решения проблемы пользователь рабочего места 4 должен иметь возможность идентифицировать ФЛ, предъявляющего носитель с идентификатором. Причем идентификация должна проводиться вне рамок автоматизированной обработки. Есть два способа реализации этой процедуры:

1. На внешнем носителе кроме идентификатора должен быть помещен некий идентифицирующий атрибут ФЛ, который может быть сравнен пользователем рабочего места 4 с тем же атрибутом на официальном документе. В данном варианте эффективность применимости будет в идеале равна значению ВИ для данного атрибута.

2. ФЛ будет предъявлять пользователю рабочего места 4 два носителя: один с постоянным идентификатором, а второй – одноразовый, который ФЛ должен при каждом посещении получать на рабочем месте 3, где владелец автомобиля будет идентифицирован. Одноразовый носитель может содержать сеансовый идентификатор для данного посещения ФЛ.

2.3. Экономическая эффективность

Для всех описанных выше вариантов использования идентификатора на внешнем носителе, наибольшие финансовые затраты потребуются не для реализации организационных и технических решений, а на модернизацию программного обеспечения (ПО) ИСПДн. К сожалению, необходимость модернизации ПО может стать принципиально непреодолимым препятствием. Это может произойти в случае, когда производителем ПО является не оператор ПД, а некая сторонняя организация.

Для оценки экономической эффективности определим стоимость создания системы защиты ИСПДн, состоящей из такого же количества рабочих мест, как в обезличенной базе, а затем определим стоимость модернизации и сравним два полученных значения.

Стоимость средства защиты от несанкционированного доступа для одного рабочего места составляет в среднем 5000 руб., его установка и настройка – 2000 руб. Остальные затраты (стоимость средств межсетевого экранирования и обнаружения вторжений, стоимость разработки эксплуатационных документов) и для защищенной ИСПДн, и для обезличенной базы будут одинаковыми, их можно не учитывать.

Стоимость модернизации структуры БД и ПО зависит не от общего количества рабочих мест, а от количества различных режимов работы, и может составлять от 0 (ПО собственного производства) до 300000 руб.

Для 10 рабочих мест (70 000 руб.) обезличивание вряд ли целесообразно, для 100 рабочих мест (700 000 руб.) целесообразно однозначно.

Таким образом, рассмотренный метод обезличивания характеризуется эффективной анонимностью, простотой технической реализации, но экономическая эффективность зависит от количества рабочих мест (как и для любого другого метода обезличивания).

Литература

1. Мищенко, Е.Ю. Количественные критерии идентификации физического лица при обезличивании персональных данных / Е.Ю. Мищенко, А.Н. Соколов // Вестник УрФО. Безопасность в информационной сфере. – Челябинск: Издат. центр ЮУрГУ, 2014. — № 1 (11). – С. 27–33.

2. Об утверждении требований и методов по обезличиванию персональных данных: приказ Федеральной службы по надзору в сфере связи, информационных технологий и массовых коммуникаций от 5 сентября 2013 г. № 996. – http://www.garant.ru. – Заглавие с экрана (дата обращения: 01.06.2015).

Мищенко Евгений Юрьевич, старший преподаватель кафедры безопасности информационных систем, Южно-Уральский государственный университет, г. Челябинск; Eug6303@mail.ru.

Соколов Александр Николаевич, канд. техн. наук, доцент, заведующий кафедрой безопасности информационных систем, Южно-Уральский государственный университет, г. Челябинск; ANSokolov@inbox.ru.

Поступила в редакцию 27 мая 2015 г.

DOI: 10.14529/ctcr150303

QUANTITATIVE ANALYSIS OF THE DEPERSONALIZATION PROCEDURE. METHOD OF IDENTIFIERS

***E.Yu. Mishchenko**, South Ural State University, Chelyabinsk, Russian Federation, Eug6303@mail.ru,*
***A.N. Sokolov**, South Ural State University, Chelyabinsk, Russian Federation, ANSokolov@inbox.ru*

Depersonalization is the way of personal data processing for the purpose of transforming data to protected status, in order to prevent disturber use it to damage the person. The result of depersonalization depends on content of the personal data and the depersonalization method also. Standard acts define some methods of depersonalization, but all of them are describing by qualitative criteria. This article makes the quantitative analysis of one of the depersonalization methods – the method of identifiers (ID Method). Proposed variant of technical realization of ID Method solves the problem of the identification attribute set, which is necessary and sufficient for the cross-reference table, define the linking identifier requirements, and also describes the methods of communication between the cross-

reference table and depersonalized (impersonal) data. On practical example the performance evaluation of ID Method is made by using some criteria, including the technical criteria (identification impossibility on the one side and reconstruction possibility by means of complementary data on the other side), and commercial criteria (economic return). On the base of identification probability and depersonalization degree some recommendations of efficiency enhancement are proposed.

Keywords: personal data, depersonalization, method of identifiers.

References

1. Mishchenko E.Yu., Sokolov A.N. [Quantitative Criteria of Identification of Person at a Depersonalization of Personal Data]. *Bulletin of Ural Federal District. Safety in the Information Sphere*, Chelyabinsk, South Ural St. Univ. Publ., 2014, no. 1 (11), pp. 27–33.

2. Ob utverzhdenii trebovaniy i metodob po obezlichivaniyu personal'nykh dannykh: prikaz Federal'noy sluzhby po nadzoru v sfere svyazi, informatsionnykh tekhnologiy i massovykh kommunikatsiy ot 5 sentyabrya 2013 № 996 [About the approval of requirements and methods on a depersonalization of personal information: the order of Federal Service for Supervision in the Sphere of Telecom, Information Technologies and Mass Communications of September 5, 2013 no. 996]. Available at: http://www.garant.ru (accessed 01.06.2015).

Received 27 May 2015

ОБРАЗЕЦ ЦИТИРОВАНИЯ

Мищенко, Е.Ю. Количественный анализ процедуры обезличивания персональных данных. Метод введения идентификаторов / Е.Ю. Мищенко, А.Н. Соколов // Вестник ЮУрГУ. Серия «Компьютерные технологии, управление, радиоэлектроника». – 2015. – Т. 15, № 3. – С. 18–25. DOI: 10.14529/ctcr150303

FOR CITATION

Mishchenko E.Yu., Sokolov A.N. Quantitative Analysis of the Depersonalization Procedure. Method of Identifiers. *Bulletin of the South Ural State University. Ser. Computer Technologies, Automatic Control, Radio Electronics*, 2015, vol. 15, no. 3, pp. 18–25. (in Russ.) DOI: 10.14529/ctcr150303