РАЗРАБОТКА СИСТЕМЫ ЗАЩИТЫ ПЕРСОНАЛЬНЫХ ДАННЫХ С ПОМОЩЬЮ ДЕПЕРСОНАЛИЗАЦИИ

В.С. Трофимов, Н.Н.Минакова
Алтайский государственный университет, г. Барнаул

Введение.

Нарастающая автоматизация процессов сбора, обработки, хранения и передачи больших объемов информации, необходимых для полноценного функционирования как государственных, так и коммерческих организаций, рост киберпреступности, обусловленный повсеместным осваиванием новых технологий, вызывают необходимость в непрерывной комплексной защите информационных ресурсов. Такая защита обеспечивается различными методами: начиная с физической защиты периметра и заканчивая организационными мерами [1- 3]. Защита персональных данных, являющихся конфиденциальной информацией в соответствии с Федеральным законом «О персональных данных» (152-ФЗ), требует существенных расходов организаций. Согласно требованиям законодательства РФ, обезличенные (относящиеся к 4-ой категории) персональные данные не нуждаются в обеспечении конфиденциальности. Поэтому задача сокращения затрат на защиту персональных данных часто решается путем понижения их категории. В методических рекомендациях [4] перечислены следующие методы деперсонализации: введение идентификаторов; изменение состава или семантики; декомпозиции; перемешивание. В представленной работе выбор метода был реализован с помощью следующих критериев: обратимость; возможность обеспечения заданного уровня конфиденциальности.

1. Алгоритм обезличивания.

Исходя из выше изложенных критериев, применялось сочетание методов введения идентификаторов и перемешивания. Используемый метод перемешивания базируется на свойствах циклических перестановок множеств атрибутов исходного массива персональных данных.

Алгоритм обезличивания реализован следующим образом:

- исходный набор персональных данных разделяется на подмножества по определенным параметрам;
- в каждом подмножестве выполняется циклический сдвиг элементов вправо на заданное число позиций;
- подмножества циклически сдвигаются между собой;

- данный цикл повторяется определенное количество раз, при этом исходные параметры автоматически изменяются на установленное значение;
- данная процедура повторяется для каждого атрибута исходного массива с измененными параметрами;
- определенный атрибут перемешанных персональных данных заменяется заданными идентификаторами.

Пример распределения значений перемешанных обезличенных персональных данных приведен на рис.1. На оси абсцисс откладываются значения, соответствующие элементам исходного массива данных, на оси ординат - модуль разности исходных и деперсонализированных значений.

Предложенный подход, включающий комбинацию методов, закладывает следующие преимущества обезличивания:

- устойчивость обезличенных персональных данных к анализу повторов;
- возможность внесения изменений в массив данных без дешифрации;
- хорошая защищенность данных;
- приемлемая скорость работы алгоритма;
- возможность восстановления исходных данных;
- низкое потребление ресурсов.

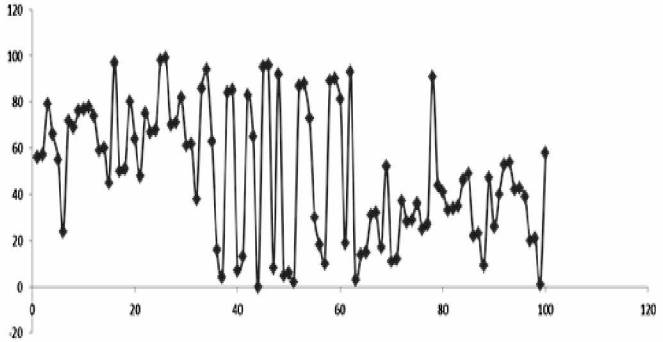

Рис.1. График распределения значений атрибута исходных данных

Для оценивания качества защищенности обезличенных персональных данных к анализу повторов была введена величина K:

$$K = \frac{\sum_{i=1}^{n-1}|(a_i - a_{i+1})|}{n},$$

где a_i - элемент обезличенного массива; n – размер исходного массива персональных данных. Величина K показывает среднее рассеивание массива относительно соседних элементов: при увеличении значения.

Экспериментально установлено, что величина K возрастает:

- при разбиении множества атрибутов на более мелкие подмножества;
- при увеличении количества циклов перемешивания;
- при увеличении сдвига элементов в подгруппах.

2. Разработанная программа.

Реализация системы осуществлялась на языке C++ с использованием библиотеки Qt в среде разработки Qt Creator. Система построена по принципу трехуровневой архитектуры (клиент – сервер – база данных), имеет гибридную сеть и состоит из 3 базовых приложений:

- клиентское - обеспечивает интерактивное взаимодействие пользователя с системой;
- балансировочное - гарантирует заданное распределение нагрузки между доступными серверами;
- серверное - выполняет обслуживающие функции по запросу клиента, в том числе и обезличивание;

Архитектура приложения, основанная на гибридной сети, позволяет клиентам использовать вычислительные мощности удаленных хостов. Это увеличивает надежность системы в целом. При достаточном количестве хостов скорость работы ограничивается только величиной пропускных каналов. Каждый работающий в системе сервер способен выполнять параллельные вычисления позиций аргументов массива перемешиваемых данных. Перед началом взаимодействия сервер определяет узлы системы, готовые к выполнению операций. При пересылке открытых персональных данных сообщения шифруются симметричным алгоритмом шифрования.

Взаимодействие частей осуществляется следующим образом:

- клиент выполняет подключение к балансировочному приложению;
- балансировщик выполняет авторизацию пользователя и

соединяет клиента с доступным сервером;

- каждый из серверов синхронизирует данные между собой;
- сервер приложений обрабатывает запросы и производит передачу шифрованных данных клиенту.
- все данные управляющий сервер получает из баз данных, безотказность работы которых обеспечивается репликацией.

Панель программы (рис.2) по обезличиванию персональных данных разработана в Qt Designer. Испытания проводились для различных наборов входных данных и показали высокую эффективность алгоритма. Работа осуществлялась с объемами данных, содержащими квадратные матрицы с порядка 10^6 записей.

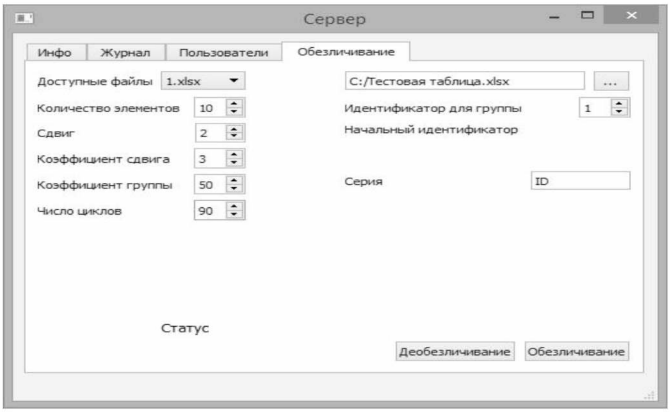

Рис.2. Панель программы.

Заключение.

На основании анализа существующих методов обезличивания персональных данных предложен алгоритм обезличивания персональных данных, разработано клиент-серверное программное обеспечение с использованием данного алгоритма. К особенностям разработанного программного обеспечения относятся: кроссплатформенность; широкий набор настроек.

Библиографический список

1. Минакова Н.Н. Методы технической и правовой защиты информации в сети Интернет // Н.Н. Минакова, В.В. Поляков, С.Н. Толстошеев – Барнаул, 2015.

2. Поляков В.В. Региональные аспекты технической и правовой защиты информации // В.В.Поляков, В.А. Трушин, В.В. Поляков и др. – Изд-во АлтГу, Барнаул, 2013.
3. Методы оценки несоответствия средств защиты информации //А. С. Марков, В. Л. Цирлов, А. В. Барабанов; под ред. А. С. Маркова. -М.: Радио и связь, 2012.
4. "Методические рекомендации по применению приказа Роскомнадзора от 5 сентября 2013 г. N 996 "Об утверждении требований и методов по обезличиванию персональных данных" (утв. Роскомнадзором 13.12.2013).