К.Д. Михайлов
Научный руководитель – *В. Г. Жуков*
Сибирский государственный аэрокосмический университет
имени академика М. Ф. Решетнева, Красноярск

О ПРИМЕНЕНИИ САМООРГАНИЗУЮЩИХСЯ КАРТ КОХОНЕНА ДЛЯ РЕШЕНИЯ ЗАДАЧИ ОБНАРУЖЕНИЯ АНОМАЛИЙ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ В РЕЛЯЦИОННЫХ СУБД

В настоящее время любая организация использует автоматизированные системы (АС). Одним из компонентов АС является система управления базами данных (СУБД). СУБД нашли свое распространение в бухгалтерском учете, управлении предприятиями, органах государственной власти, создании Интернет ресурсов и многих других. Примером защищаемой информации, хранящейся в базе данных, может быть: коммерческая тайна, банковская тайна, персональные данные и т.п. Наибольшее распространение получили реляционные СУБД.

Сегодня на рынке представлено множество СУБД, как платные, так и свободно распространяемые. В составе широко распространенных СУБД есть встроенные механизмы защиты с разной степенью надежности, но, в основном, они основываются на парольной защите учетной записи пользователя. Тем не менее, по статистике, около 500 000 серверов баз данных в интернете содержат уязвимости[1]. Зачастую встроенные механизмы защиты также оказываются бессильными перед внутренними угрозами. Наблюдение за действиями, производимыми пользователями позволяет отслеживать необычные и подозрительные события и тенденции в поведении пользователя – аномалии информационной безопасности. Опасность аномального поведения заключается в том, что оно не является явным нарушением политики безопасности.

При распознавании аномалий возникает целый ряд затруднений, связанных, главным образом, с необходимостью учета и обнаружения ранее неизвестных типов атак и воздействий. Использование журналов системы протоколирования для наблюдения является наиболее подходящим способом для выполнения такой работы. Обнаружение аномалий использует модели предполагаемого поведения пользователей и приложений, интерпретируя отклонение от «нормального» поведения как потенциальное нарушение защиты.

Сегодня для обнаружения аномалий используются следующие методы:

– статистика Байеса. Вероятностный алгоритм, использующий в своей основе теорему Байеса;

– описательная статистика. Так же использует статистические методы оценки сравнения эталонного и текущего поведения;

– обнаружение аномалий на основе инвариантов подобия. Суть данного алгоритма заключается в распознавании аномалий вычислительных процессов с помощью сравнения значений инвариантов подобия реальных вычислительных процессов с эталонными значениями инвариантов;

– обнаружение аномалий с помощью прецедентов. В основе метода лежит оценка степени схожести двух прецедентов путем вычисления расстояния между всеми наблюдениями в n-мерном пространстве данных, например, с помощью функций Евклида, Миньковского, Миньковского с весами, максимума.

Целью работы является разработка алгоритма, позволяющего обнаружить аномальное поведение пользователя в процессе работы с базой данных. В качестве тестовой была выбрана СУБД Firebird 2.1. Выбор данной СУБД обусловлен тем, что она является одной из популярных среди свободно распространяемых СУБД. Минусом

[1] По данным аналитической компании “NGS Software”

данной СУБД является отсутствие средств протоколирования. Решение данной проблемы возможно применением средств языка SQL тремя способами:

– запись информации об операции в текстовый файл на стороне клиента;
– запись в таблицу базы данных;
– запись в отдельную базу данных;

Для решения поставленной задачи был выбран вариант записи в таблицу базы данных. Реализация была произведена при помощи триггеров, инициализируемых при операциях добавления, удаления или изменения элемента таблицы. Триггер записывает следующую информацию об операции: дату и время операции, имя пользователя, содержание операции (объект базы данных и применяемый к нему оператор SQL), IP-адрес пользователя. На основании этих данных будут производиться обнаружение и анализ аномалий информационной безопасности.

Следует отметить, что большинство статистических инструментов обнаружения сетевых аномалий имеют ряд недостатков, влияющих на эффективность их работы:

– отсутствие группировки исходных данных сразу по всем анализируемым параметрам;
– сложность формирования и представления общей картины активности пользователей БД для всех компьютеров одновременно с возможностью определения количественных величин анализируемых параметров для отдельно взятого компьютера.
– отсутствие механизмов нахождения «скрытой» информации, не содержащейся в явном виде в исходном массиве статистической информации.

Выходом из сложившейся ситуации является применение нейросетевых технологий анализа данных – самоорганизующихся карт Кохонена – в рамках автоматизированной среды интеллектуального анализа данных.

Самоорганизующаяся карта Кохонена (англ. Self-organizing map, SOM) – нейронная сеть с обучением без учителя, решающая задачу визуализации и кластеризации [1].

Самоорганизующаяся карта Кохонена – это модифицированный алгоритм *линейного векторного квантования данных,* то есть представления *N* точек данных с помощью меньшего числа точек-образцов.

В результате работы алгоритма строится карта, то есть двумерная сетка узлов, размещенных в многомерном пространстве. Для того чтобы изобразить их положение используются различные средства. Одно из них – такое раскрашивание карты, когда цвет отражает расстояние между узлами.

Таким образом, применение SOM для решения задачи обнаружения аномалий позволит создать эффективный механизм защиты от внутренних угроз информационной безопасности в целевой среде реляционной СУБД.

Библиографический список:

1. Teuvo Kohonen Self-organizing maps 3rd Edition. – Springer, Berlin – Heidelberg – New York, 2001.
2. Райх В.В. Исследование свойств нейронных сетей Кохонена и адаптивного резонанса применительно к задачам мониторинга информационной безопасности. // Информационная безопасность: Материалы VI Международной научно-практической конференции, 1-7 июля 2004 г. Таганрог: Изд-во ТРТУ, 2004. С. 193-195.
3. Information security [Электронный ресурс]: Электрон. журн. / учредитель ООО “Гротек” - 2007г., октябрь. – М.: Режим доступа к журн.: http://www.itsec.ru/newstext.php?news_id=37481
4. Пат. 20090292743 США, IPC8 Class: AG06F1200FI USPC Class: 707202. MODELING USER ACCESS TO COMPUTER RESOURCES [Электронный ресурс] / Christoph Lingenfelder Joseph P. Bigus Leon Gong; заявитель и

патентообладатель IBM CORPORATION, INTELLECTUAL PROPERTY LAW;DEPT 917, BLDG. 006-1; опубл. 26.11.09, Режим доступа: http://www.faqs.org/patents/app/20090292743