

Н.А. Маслова, Е.А. Маслова

Донецкий национальный технический университет

МЕТОДОЛОГИЯ ПРИМЕНЕНИЯ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА В ЗАЩИТЕ ИНФОРМАЦИИ

Особливістю викладеної методології є комплексний підхід, застосування методів інтелектуального аналізу, швидкодіючих алгоритмів, уніфікація інструментарію захисту складних систем та оцінка його ефективності.

Особенностью изложенной методологии является комплексный подход, применение методов интеллектуального анализа, быстродействующих алгоритмов, унификация инструментария защиты сложных систем и оценка эффективности.

Feature of the methodology set out a comprehensive approach, the application of mining techniques, high-speed algorithm, unification of security tools of complex systems and an evaluation of its effectiveness

Ключевые слова: интеллектуальный анализ, защита информации, комплексный подход.

Вступлення. Інтелектуальний аналіз даних використовується в різних областях діяльності сучасного суспільства, допомагаючи вирішувати всілякі завдання. Серед таких областей можна виділити страхування, банківське діло, маркетинг, аналіз фінансових ризиків, моніторинг обладнання і технологічних процесів, телекомунікації, комп'ютерну безпеку і т.д.

Найбільш складними сучасними інформаційними структурами, орієнтованими на великі компанії, є корпоративні системи. Базы даних корпоративних систем містять величезні обсяги інформації і мають всі ознаки складної системної організації. Для цих систем характерні використання множини комп'ютерів, спеціалізація серверів і розгалужена система прийому-передачі даних. Особливістю багатьох корпоративних систем, експлуатуваних в даний час, є те, що системи безпеки вихідно в їх складі не входили і повинні підбиратися і придбуватися окремо. Це призводить до додаткових витрат, вимагає розробки узгоджень при інтеграції двох різноманітних систем.

Особливістю систем захисту інформації (СЗІ) в корпоративних системах є комбінація як мінімум трьох проблем: захист інформації в комп'ютерних мережах; забезпечення безпеки баз даних; гарантія спроможності вузлів вводу, обробки і зберігання інформації [1]. Організувати бесперебойну роботу складної структури, захистити її від зовнішніх і внутрішніх загроз в сучасному інформаційному просторі традиційними методами дуже складно. Одним з перспективних напрямків є застосування інтелектуальних алгоритмів, і в частині інтелектуального аналізу даних (ІАД). Підтвердженням застосування інтелектуальних систем в корпоративному управлінні є наступні приклади.

Microsoft SQL Server 2008 надає інтегровану середовище для створення моделей інтелектуального аналізу даних і роботи з ними. Ця середовище називається Microsoft SQL Server Analysis Services і складається з набору спеціальних інструментів (Business Intelligence Development Studio, SQL Server Management Studio, Microsoft SQL Server 2008 Integration Services, BI Development Studio). Середовище включає алгоритми інтелектуального аналізу даних і засоби, що спрощують розробку комплексного рішення, застосованого в межах різних проектів. Microsoft заявляє про застосування «технології активної захисту», заснованої на оцінці поведінки програм з точки зору їх потенційної небезпечності. В частині, СЗІ коректують засоби захисту комп'ютера при зміні його статусу або блокують його, якщо виникає підозра на зараженні вірусом або проникненні злоумисленника.

В квітні 2010 г. компанія IBM представила системи інтелектуального аналізу і обробки транзакцій, що допомагають витягувати знання з великих масивів даних. Ці системи дозволяють визначати приховані можливості і виконувати аналіз систем на поведінковому рівні.

Компанія ISS (Internet Security Systems) розробила модель адаптивного управління безпекою, отримавшу назву ANS (Adaptive Network Security). Адаптивний компонент ANS дозволяє модифікувати процес аналізу захищеності, надаючи актуальну інформацію про знову виникаючі вразимості систем. Він також модифікує компонент виявлення атак, доповнюючи його останніми інформацією про підозріливі дії і атаки.

Сімейство продуктів SAFEsuite, розроблене ISS, в даний час є найбільш потужним комплексом систем, що включає в себе такі компоненти моделі інтелектуального адаптивного управління безпекою мережі, як:

- систему аналізу захищеності на рівні мережі Internet Scanner;

- системы анализа защищенности на уровне операционной системы и прикладного ПО System Scanner и Online Scanner;
- систему анализа защищенности на уровне СУБД Database Scanner;
- системы обнаружения атак RealSecure (Network Sensor, Appliance, OS Sensor, Server Sensor);
- систему поддержки принятия решения и прогнозирования в области безопасности SAFESuite Decisions.

Российская компания «Информзащита» сообщает о разработке концепции и реализации подсистемы адаптивной безопасности и противодействия внешним атакам ОАО «Вымпелком». Примером интеллектуального компонента служит механизм обновления баз данных антивирусных программ, которые являются частным случаем систем обнаружения атак.

В Eset NOD32 используется патентованная технология ThreatSense, предназначенная для выявления возникающих угроз в реальном времени путём анализа выполняемых программ на наличие вредоносного кода, что позволяет предупреждать действия вредоносных программ. Наряду с использованием баз вирусов NOD32 применяет эвристические методы, что позволяет обнаруживать новые неизвестные вирусы и проводить их нейтрализацию. Большая часть кода антивируса написана на языке ассемблера, что дает возможность эффективно использовать системные ресурсы и обеспечивать высокую скорость проверки настроек по умолчанию.

Постановка задачи. Перечисленные технологии будут активно работать, если продукт приобретается в полной комплектации. А что делать тем предприятиям, которые имеют налаженную систему, эксплуатируемую длительный период, в которой накоплены значительные объемы данных и которые считают переход на новую платформу нерентабельным? Им приходится применять совокупность методов защиты, не объединенных единой концепцией безопасности, и продолжать поиски лучших с точки зрения соотношения эффективность – затраты решений. В этом случае необходима методология применения интеллектуального анализа для обеспечения безопасного функционирования сложной распределенной структуры, единая методика, которая обеспечит защиту сетей, баз данных и систем обработки от постоянно совершенствующихся угроз.

Необходимость использования инструментария интеллектуального анализа данных в СЗИ корпоративных систем проистекает из разнородности структур информационных пространств этих систем; сложности получения аналитической информации из баз данных значительного объема; большого числа пользователей, одновременно работающих в системе; требований постоянного контроля функционирования и принятия обоснованных управленческих решений, зависящих от множества факторов.

Предпосылками использования ИАД является клиент-серверная технология, распределенные базы данных, наличие хранилищ информации, применение современных сетевых технологий и инструментария, используемого для сбора, анализа и визуализации данных. ИАД может работать с различными источниками и типами данных, помогает реализовать в системе защиты эволюционные свойства адаптации, самоорганизации, обучения, наследования и представления опыта экспертов информационной безопасности в виде доступной для анализа системы нечетких правил.

Сегодня администратору системной безопасности недостаточно иметь средства управления учетными записями и ресурсами либо механизмы защиты от какой-то конкретной выявленной угрозы. Ему необходим механизм прослеживания тенденций и прогнозирования событий в сфере безопасности. Поэтому решение проблем безопасности корпоративных информационных систем требует поиска эффективных механизмов, которые должны работать в режиме реального времени, обладать высокой чувствительностью к изменениям в информационной инфраструктуре, включать базу знаний, содержать элементы ИАД и позволять генерировать решение в соответствии с заданной целевой функцией в постоянно изменяющихся условиях внешней среды.

Ввиду вышесказанного разработка методологии применения интеллектуального анализа в защите информации актуальна и перспективна.

Целью работы является описание методологии построения интеллектуальной адаптивной системы защиты информации, основанной на применении принципов интеллектуального анализа данных.

В качестве математической модели, с помощью которой можно описать процессы, происходящие в интеллектуальной системе, можно принять следующие соотношения:

$$T \times X \times S \xrightarrow{a_1} M \times T; \quad T \times M \times X \xrightarrow{a_2} C \times T;$$

$$C \times T \times X \times S \xrightarrow{a_3} R \times T; \quad T \times X = \{A \times T\} X \times T + \{B \times T\} U \times T;$$

$$T \times Y = \{D \times T\} X \times T; \quad T \times R \times Y \xrightarrow{a_4} C \times T,$$

где T – множество моментов времени; $\{A\}$, $\{B\}$ и $\{D\}$ – матрицы параметров. X , S , M , R – это множества, характеризующие систему, среду, мотивации и цели соответственно. Интеллектуальные операторы преобразования обозначены a_1 – a_4 .

Современные компьютерные системы и сети находятся в состоянии постоянного развития и модификации, а объемы анализируемых данных в мире удваиваются ежегодно. Поэтому для обеспечения требуемого уровня защиты информации необходимо гибко и оперативно реагировать на изменяющиеся условия, обеспечивать надежную защиту с учетом постоянного изменения входных воздействий, предупреждать действия злоумышленников, т.е. иметь систему защиты, построенную с использованием элементов интеллектуальности, оперативно реагирующую на постоянно совершенствующиеся внешние угрозы.

Особенностью систем защиты информации в корпоративных системах является комбинация как минимум трех проблем: защита информации в компьютерных сетях; обеспечение безопасности баз данных; обеспечение безопасной работы систем автоматической обработки информации.

Наиболее актуальными угрозами информационной и сетевой безопасности корпоративных систем являются:

- угрозы, связанные с злонамеренной модификацией параметров функционирования системы внутренними нарушителями;
- угрозы несанкционированного доступа к информации с целью ознакомления, модификации или блокирования;
- угрозы, связанные с разграничением прав доступа и сложностью администрирования в распределенной системе;
- угрозы, связанные с вирусной атакой на рабочую станцию пользователя, локальную или корпоративную сеть предприятия;
- угрозы, связанные с передачей информации по каналам связи и работой в сети Internet.

К традиционным средствам обеспечения информационной безопасности (ИБ) корпоративных компьютерных сетей относят антивирусы, детекторы уязвимостей, межсетевые экраны и детекторы вторжений. Они решают отдельные задачи обеспечения ИБ корпоративной сети и, как правило, могут быть преодолены при командной работе квалифицированной группы нарушителей.

К часто используемым в компьютерных сетях интеллектуальным средствам относят базы знаний в составе экспертных систем, системы на основе байесовского метода, нечеткие логические системы, нейронные сети, эволюционные методы и гибридные интеллектуальные системы. Основными задачами, решаемыми интеллектуальными средствами обеспечения информационной безопасности компьютерной сети, являются классификация и кластеризация.

Нейронные сети используются для контроля трафика защищаемой локальной сети, поиска скрытых закономерностей в массивах первичных данных, своевременного выявления вторжений. Для предсказания значения целевого показателя используются наборы входных переменных, математических функций активации и весовых коэффициентов входных параметров. Выполняется итеративный обучающий цикл, нейронная сеть модифицирует весовые коэффициенты до тех пор, пока предсказываемый выходной параметр соответствует заданному значению. После обучения нейронная сеть становится моделью, используемой при прогнозировании.

Механизмы классификации применяются на первоначальном этапе, например, для систематизации способов защиты (нечеткие заключения) по вектору нечетких признаков угроз. Если достоверность классификации по известным угрозам меньше некоторого уровня, то при наличии признаков проведенной атаки классификация расширяется за счет введения новой градации в классификацию – решается задача кластеризации угроз. Ассоциации выявляют причинно-следственные связи и определяют вероятности (коэффициенты достоверности), позволяя делать соответствующие выводы.

Защита баз данных является одной из самых сложных задач систем защиты информации в корпоративных системах. Наиболее распространенными угрозами, характерными для баз данных, являются хищение, утрата, уничтожение, модификация данных и отказ от подлинности.

С интеллектуальными средствами обеспечения безопасности баз данных можно ознакомиться в [2]. Система информационной безопасности баз данных должна использовать средства и объекты применяемой системы управления базами данных (СУБД) и базы данных, набор правил и событий, характеризующих действия пользователей.

К средствам обеспечения безопасной работы систем обработки информации относятся механизмы предотвращения вторжений, авторизация, разграничение прав доступа, криптозащита (на носителях информации, в сетях, парольная защита), управление полномочиями пользователей. С целью контроля состояния системы используют базы сигнатур известных атак, а в качестве основных источников информации — системные журналы и файлы, анализируют содержимое сетевого трафика и файлов. Фиксация событий позволяет определить интересы каждого из пользователей, составить перечень регистрируемых событий.

В основном публикации о применении интеллектуальных систем для обеспечения безопасной работы систем обработки информации посвящены системам обнаружения атак, основанных на модели, предложенной Деннингом. Модель содержит набор профилей для легальных пользователей, сравнивает текущие действия с соответствующим профилем, обновляет профиль и сообщает о любых обнаруженных аномалиях.

При традиционном подходе к построению системы защиты с применением инструментария ИАД используются искусственные нейронные сети, деревья решений и алгоритмы классификации, методы нечеткой кластеризации, ассоциативные правила, алгоритмы ограниченного перебора, кластерный анализ.

Недостатками традиционного подхода являются следующие:

- базы знаний формируются экспертами, т.е. принцип включения в них ситуаций субъективен;
- базы знаний периодически обновляются, упорядочиваются, систематизируются, что является трудоемкой и дорогостоящей процедурой;
- существует задержка во времени между появлением атаки и средств защиты (запаздывающее противодействие);
- атаки видоизменяются, совершенствуются, «маскируются» под стандартные процедуры, что требует модификации средств защиты.

Перейдем к описанию особенностей построения адаптивной саморазвивающейся системы. В табл. 1 приведен перечень основных источников данных.

Таблица 1

Источники анализируемых данных	
Источник данных	Анализируемая информация
log-файлы работающих подсистем	Время, тип, сущность операций, соответствие пароля, сбои при установке связи с удаленной машиной, диагностика аварийных остановов
Сетевые трафики	Загрузка сетевого оборудования, использование каналов связи, сетевая активность
Справочники и журналы регистрации пользователей и событий	ID-коды пользователей, корректность паролей, выполняемые действия
Перечни функциональных задач	Цепочки взаимосвязанных вызовов задач и процессов
Информация о правах доступа	Соблюдение регламента обращений к ресурсам
Сведения о работе почтовой системы	Статистика, объемы и адреса рассылок и почтовых поступлений, тематика сообщений
Текстовые файлы	Тематическая направленность

Окончание табл.1

Книги Excel	Безопасность, наличие/отсутствие макросов
Таблицы с атрибутами исполняемых файлов	Типы файлов, даты изменения, авторы изменений и их права, контроль «неизменности», адреса эталонных модулей, контрольные суммы

Рабочие гипотезы:

- активность пользователей, целевые обращения к ресурсам системы и происходящие в системе процессы можно зафиксировать и построить их адекватную модель;
- событие (последовательность событий), соответствующее обобщенной модели атаки, действительно является атакой, и применение алгоритмов опережающего или одновременного противодействия является обоснованным;
- система может отследить работу программного обеспечения системы и, обнаружив повреждение, восстановить защиту, выполнить автоматическую докачку потерянных или поврежденных файлов.

Механизм применения ИАД для адаптивной саморазвивающейся СЗИ можно разбить на ряд этапов.

Этап 1. Составление перечня основных источников данных и выбор информации, подлежащей анализу.

Этап 2. Постановка задачи. Выполняется анализ требований, определяются проблемы, которые будут решаться, метрики, по которым выполняется оценка модели, а также определяются задачи для проекта интеллектуального анализа данных. Здесь же исследуются уровни конфиденциальности данных, потребности и права пользователей в отношении доступных данных, методы идентификации и аутентификации, традиционно используемые на предприятии.

При этом риски информационной безопасности системы могут быть определены как функция трех переменных:

- вероятности существования угроз (потенциально возможных событий, преднамеренных или случайных, которые могут оказать нежелательное воздействие на корпоративную систему или её части либо на информационные активы и, как следствие, на бизнес-процессы компании);

- вероятности существования уязвимостей (недостатков) в системе, из-за которых возможно нежелательное воздействие на нее;

- потенциальных убытков, которыми являются потенциально возможные прямые и косвенные финансовые потери, полученные вследствие реализации угроз и наличия уязвимостей.

Этап 3. Сортировка и очистка данных. Данные упорядочиваются, удаляются недопустимые и ошибочные комбинации, определяются первоисточники данных, строятся согласования, подбираются, например, столбцы, данные из которых в дальнейшем могут быть использованы при анализе. При этом данные могут находиться в разных подразделениях компании и храниться в различных форматах, что потребует применения механизмов интеграции систем.

Алгоритмы сортировки оцениваются по скорости выполнения и эффективности использования памяти. Если алгоритм сортировки использует только абстрактную операцию сравнения ключей, то его вычислительная сложность составит $O(n \log n)$ операций сравнения. При параллельном вычислении n ситуаций можно отсортировать за $O(\log^2 n)$ операций, а худшим вариантом являются алгоритмы, вычислительная сложность которых соответствует $O(n^2)$ операций. Требуемый объем памяти при реализации алгоритмов, как правило, составляет $O(\log n)$ ячеек.

Методы сортировки, рекомендуемые для использования в системе: сортировка вставками (может сортировать список по мере его получения), блочная сортировка (относится к классу быстрых алгоритмов с линейным временем исполнения $O(N)$).

Этап 4. Классификация регистрируемых событий, например угроз по вектору признаков атак и механизмов защиты по вектору угроз, выделение кластеров. При этом используются:

- средства, упрощающие разделение данных на набор данных для обучения и проверочный набор данных;
- сортировка и очистка данных (упорядочение, удаление недопустимых и ошибочных комбинаций, согласование данных);
- формирование матриц адаптируемых экспертных оценок и на их основе создание исходных систем нечетких правил и классификаторов;
- классификация регистрируемых событий (например, угроз по вектору признаков атак и механизмов защиты по вектору угроз, выделение кластеров, упрощающие разделение данных на обучающий и проверочный наборы);
- предварительный статистический анализ данных, получение контрольных метрик и закономерностей;
- создание структуры интеллектуального анализа данных.

В дальнейшем эти данные могут использоваться несколькими подсистемами интеллектуального анализа, например уже упоминаемыми ранее подсистемой нерегламентированных действий пользователей или системой анализа вторжений, модели которых построены по одной структуре.

Результаты предыдущих пунктов представляются в виде систем нечетких правил, которые реализуются в виде специализированных структур; осуществляется подбор классификаторов, формирование признака структуры (Ps). Анализ данных и формирование признаков происходят постоянно и независимо от дальнейшей работы алгоритма. Изменение параметра Ps свидетельствует об изменении условий внешней среды (появление, например, нового вида атаки и необходимости выбора иной модели процесса либо обучения прежней на новом наборе данных).

Правила представляются в виде «если» – «то» и используются для прогнозирования. На основе частоты встречаемости логических закономерностей делается вывод о возможном событии. Например, цепочка MD–COPYAZ–ARH–WWW–DEL ассоциируется с копированием информации из конфиденциального источника и передачей его по Internet-каналам.

Этап 5. Построение модели. На этом этапе модель представляет собой математическое выражение, контейнер, еще не наполненный данными. При этом, если данные изменяются, необходимо обновить структуру и модель интеллектуального анализа данных, изменить признак Ps.

Построение модели интеллектуального анализа данных является частью процесса, в который входят все задачи – от выбора и определения данных, создания модели до развертывания модели в рабочей среде. Для построения модели используют математические основы скрытых Марковских цепей, интеллектуальные мультиагентные технологии, аппарат нечетких множеств и семиотического моделирования и т.д. Однако главным требованием при этом является комплексный, системный подход, единый процесс построения адаптивной системы с учетом требований и методологии защиты информации.

Перед развертыванием модели в рабочей среде проверяется эффективность её работы. При этом создается несколько моделей с различной конфигурацией, анализируются все модели, определяется, какая из них обеспечивает лучшие результаты.

Этап 6. Передача опыта адаптивной СЗИ (наследование) по обеспечению информационной безопасности.

Этап 7. Обучение классификаторов на обучающей выборке – подмножестве входных векторов, формирование информационных полей четких классификаторов.

Этап 8. Адаптация системы к реальным условиям.

Этап 9. Коррекция матриц экспертных оценок и систем нечетких правил по результатам адаптации.

При этом решение о расширении классификаций атак и механизмов защиты производится в соответствии с системой оценок достоверности нейтрализации угроз в разрезе отдельных механизмов защиты. Обосновать целесообразность использования механизма защиты в составе многоуровневой СЗИ можно, например, по матрице достоверности использования механизмов защиты для нейтрализации угроз [3]:

$$x_i = \sqrt[n]{\prod_{j=1}^n m e_{ij}}, i = 1, \dots, m,$$

где $m e_{ij}$ – элементы матрицы достоверности «угрозы – защита».

Этап 10. Формулирование новых нечетких правил в случае расширения классификации, разработка спецификации на создание нового механизма защиты. Формирование комплекса оценок защищенности системы.

Этап 11. Анализ структуры классификаторов и выявление недостатков в системе защиты, оценка эффективности системы, включение в нее дополнительных механизмов защиты.

Этап 12. Контроль целостности данных и программных модулей, при необходимости – восстановление программной среды, изменение структуры системы информационной безопасности.

Порядок действий согласно методу проектирования адаптивных СЗИ может изменяться, но обязательными являются [3]:

- формирование многомерных матриц адаптируемых оценок и создание исходных систем нечетких правил и классификаторов (на нижних уровнях защиты — классификаторов «признаки атаки — угрозы», на верхних уровнях защиты – классификаторов «угрозы – механизмы защиты»);
- идентификация выявленной угрозы и при расширении поля угроз – кластеризация их с последующей адаптацией информационных полей путем обучения алгоритма ИАД;
- коррекция и расширение системы нечетких правил, вызванная изменением поля угроз;
- модификация систем нечетких правил и матриц экспертных оценок в результате обучения классификаторов уровней защиты;
- описание нового механизма защиты и формулировка спецификации на создание этого механизма;
- анализ защищенности ИТ-системы.

Одной из наиболее важных задач при обеспечении безопасности системы является сбалансированность уровня её защищенности и производительности. Применение средств защиты информации в корпоративной системе снижает ее производительность. Уровень защищенности и производительность находятся в обратно пропорциональной зависимости друг от друга. Поэтому построение интеллектуальных систем защиты невозможно без быстродействующих алгоритмов. Например, как было указано выше, одним из недостатков традиционного подхода является задержка во времени между появлением новой атаки и средств защиты от нее.

В теории СЗИ различают одновременное, опережающее и запаздывающее противодействия. Запаздывающее противодействие — когда реакция системы защиты начинается к моменту завершения угрозы или после нее. Одновременное противодействие – то, что начинает действовать с появлением угрозы. И, наконец, противодействие, носящее опережающий характер, — реакция системы защиты начинается до реализации угрозы.

Идеальным вариантом, конечно, является опережающее противодействие, однако для него необходимо не только наличие методик, позволяющих своевременно обнаружить возникновение угрозы безопасности системе, но и применение алгоритмов, способных выполнить анализ ситуации и своевременно ликвидировать результаты вторжения.

Существует несколько подходов к решению данной задачи, использующих такие алгоритмы, как деревья принятия решений, нейронные сети, логистическая регрессия, метод опорных векторов, дискриминантный анализ, ассоциативные правила. Одним из эффективных алгоритмов в этой области является так называемый наивный (упрощенный) алгоритм Байеса. С точки зрения скорости обучения, стабильности на различных данных и простоты реализации, алгоритм Байеса превосходит практически все известные алгоритмы. Обучение алгоритма производится путем определения относительных частот значений всех атрибутов входных данных при фиксированных значениях атрибутов класса.

Классификация осуществляется путем применения правила Байеса для вычисления условной вероятности каждого класса для вектора входных атрибутов. Входной вектор приписывается классу, условная вероятность которого при данном значении входных атрибутов максимальна. Алгоритм строится на предположении, что входные атрибуты условно (для каждого значения класса) независимы друг от друга [4].

Эффективность функционирования СЗИ зависит от множества действующих взаимосвязанных между собой элементов и, как правило, оценивается совокупностью критериев, находящихся в сложных конфликтных взаимоотношениях. Отсутствие общего подхода к решению задач данного класса закономерно влечет за собой многообразие различных не- взаимосвязанных методов оценки качества.

Простейшей схемой построения защиты, устраняющей 20–30% угроз, является схема *Безопасность = Традиционные средства защиты*.

Более надежной схемой построения защиты, по данным «КомпьютерПресс», обеспечивающей 40–60% эффективности, является комплексный подход, наличие четко сформулированных и действующих политик безопасности, использование разветвленного перечня традиционных средств защиты, постоянный контроль ситуации и безотлагательное применение мер защиты. При этом схема защиты выглядит следующим образом:

$$\text{Безопасность} = \text{Политика безопасности} + \text{Традиционные средства защиты} + \text{Анализ риска} + \text{Реализация контрмер}.$$

И, наконец, модель адаптивного управления безопасностью тот же источник предлагает описывать формулой

$$\text{Безопасность} = \text{Анализ риска} + \text{Политика безопасности} + \text{Традиционные средства защиты} + \text{Реализация контрмер} + \text{Аудит} + \text{Мониторинг} + \text{Реагирование}.$$

Процесс определения эффективности систем защиты начинают с выбора и обоснования показателей (критериев) оценки эффективности системы защиты, а затем переходят к подбору или разработке методик расчета этих показателей. В [5] приведен перечень распространенных подходов к выбору критериев и оценке параметров, показатели эффективности систем защиты и методики их расчета.

Одним из способов оценки эффективности СЗИ является оптимизационный, или комбинаторный, подход. При этом решается задача оптимизации вида: максимизировать некую функцию при заданных ограничениях. В случае адаптивных систем указанную методику предлагается расширить параметром k , характеризующим этапы адаптивного процесса.

Введем следующие обозначения:

$U = \{u_j\}$ — множество угроз безопасности, $j=1, \dots, m$;

$A^k = \{a_i^k\}$ — множество механизмов безопасности, используемых на k -м этапе адаптивного процесса;

$X = \{x_i\}$ — множество требований безопасности, $i=1, \dots, n$; $k=0, \dots, R$;

$C^k = \{c_i^k\}$ — допустимые затраты на создание защиты (объем затрат на сопровождение системы с учетом реализации k -го этапа адаптивного процесса), причем c_i^1 — это затраты на начальную разработку, обучение и первоначальный запуск адаптивной системы;

$d^k(i, j)$ — эффективность нейтрализации i -м механизмом безопасности j -й угрозы на k -м этапе.

Для построения математической модели вводят переменную $p(i, j)$, равную 1, если j -я угроза устраняется с помощью i -го механизма, и нулю — в противном случае. Также вводится переменная q , такая, что

$$q(i, j) = \begin{cases} 1 & \text{— если } i\text{-й механизм безопасности} \\ & \text{используется для устранения } j\text{-й угрозы,} \\ 0 & \text{— в противном случае.} \end{cases}$$

Если информационные угрозы между собой не связаны, то требуется найти максимальный эффект от нейтрализации множества информационных угроз U с помощью задекларированных в системе средств защиты A при ограничениях на общий объем затрат C .

$$\sum_{k=1}^r \sum_{j=1}^m \sum_{i=1}^n d^k(i, j) p(i, j) \Rightarrow \max \quad \text{при ограничениях} \quad \sum_{i=1}^n c(i) * \text{sign} \sum_{uj \in U} p(i, j) \leq C, \quad p(i, j) \in (0, 1), \quad j=1, \dots, m; \\ i=1, \dots, n).$$

Видоизменив постановку задачи и введя понятие функции принадлежности $\mu^A(x_i)$ в соответствии с [6], получим вариант оценки эффективности информационной системы защиты в случае нечётких показателей.

Пусть W — счётное множество показателей $W = \{w_i\}$, $i = 1, \dots, n$, n — количество показателей. Принадлежность к определённому уровню безопасности определяем на заданном промежутке, например $[0, T]$. Тогда множество значений V , определяющих выполнение требований безопасности, определяется как

$$V = \sum_{i=1}^n \frac{\mu^A(x_i)}{x_i},$$

где $\frac{\mu^A(x_i)}{x_i}$ — пара «функция принадлежности/элемент».

Разные состояния безопасности системы выделяются в виде подмножеств нечёткого множества, а вероятность взлома оцениваемой системы может соответствовать кардинальному числу (мощности) нечёткого множества. Интерпретация вышесказанного приведена в табл. 2.

Соответствие требований безопасности состояниям системы безопасности

	Требование безопасности	Состояние безопасности системы
1	1	Абсолютно незащищённая
2	2	Недостаточно защищённая
3	3	Защищённая
4	4	Достаточно защищённая
5	5	Абсолютно защищённая

При таком подходе для оценки эффективности системы защиты необходимы данные о требованиях защищённости и данные о полноте выполнения этих требований. Подобный подход позволяет добиться постоянного мониторинга состояния информационной безопасности системы, выполнить прогноз возможности атак, провести изменение требований к переменным безопасности.

Выводы. Особенностью систем защиты информации в сложных современных информационных структурах, ориентированных на крупные компании, является комбинация как минимум трех проблем: защита информации в компьютерных сетях; обеспечение безопасности баз данных; гарантия работоспособности узлов ввода, обработки и хранения информации. Организовать бесперебойную работу сложной структуры, защитить ее от внешних и внутренних угроз традиционными методами сложно. Одним из перспективных направлений является применение интеллектуальных алгоритмов и, в частности, интеллектуального анализа данных.

Целью работы является описание методики построения интеллектуальной адаптивной системы защиты информации, основанной на применении принципов интеллектуального анализа данных.

Приведены основные источники данных, подлежащих анализу, выдвинуты рабочие гипотезы, сформулированы и охарактеризованы этапы построения интеллектуальной системы. Важным моментом является использование быстродействующих алгоритмов.

Принципы интеллектуального анализа могут использоваться как при защите блоков самой системы, так и при обеспечении целостности баз данных, анализе сетевого трафика распределенной сети с целью более раннего выявления сетевых атак, в процессе контроля приема и передачи данных и т.д. И именно комплексный подход, использование единой методики анализа и применение ее к данным, полученным из различных источников, составляют новизну разработки и делают методику удобной в реальном применении.

Библиографические ссылки

1. **Шаньгин В.Ф.** Защита информации в распределенных корпоративных сетях и системах. / В.Ф. Шаньгин, А.В. Соколов. – М., – 2002. – 134 с.
2. Базы данных: интеллектуальная обработка информации / В.В. Корнеев, А.Ф. Гареев, С.В. Васютин, В.В. Райх. – М.: Нолидж, 2000. – 352 с.
3. **Нестерук Ф.Г.** Основы организации адаптивных систем защиты информации: учеб. пособие / Ф.Г. Нестерук, Г.Ф. Нестерук, Л.Г. Осовецкий. – СПб.: СПбГУ ИТМО 2008. – 112 с.
4. **Гончаров М.** Модифицированный древовидный алгоритм Байеса для решения задач классификации / М. Гончаров // Spellabs. – 2007.
5. **Маслова Н.А.** Методы оценки эффективности систем защиты информационных систем / Н.А. Маслова // Штучний інтелект. – Донецьк: ІПШІ. – 2008, – № 4 – С.253–264.
6. **Щербаков А.Ю.** Компьютерная безопасность. Теория и практика / А. Ю. Щербаков. – М.: Нолидж, 2001. – 352 с.

Надійшла до редколегії 04.12.2012 р.