УДК 004.056.53
DOI 10.33764/2618-981X-2023-6-50-57

А. А. Емелина[1], Н. Е. Карпова[1], А. А. Саранский[1]*

Исследование действий пользователей в информационной среде

[1] Самарский государственный технический университет, г. Самара,
Российская Федерация
*e-mail: esib@samgtu.ru

Аннотация. Актуальной задачей является разработка и совершенствование методов обнаружения недопустимых событий в функционировании информационной системы, которые, как правило, является следствием действия пользователей в компьютерной системе. Сложности, с которыми сталкиваются разработчики моделей, связаны с тем, что все методы касаются описания поведения человека, которое плохо поддается формализации. В данной работе предложена математическая модель байесовской сети, а также разработан алгоритм выявлния неспецифичных действий пользователя в информационной системе. В результате анализа был сделан вывод, что сети Байеса являются наиболее гибким и адекватным математическим аппаратом, позволяющим даже в условиях некоторой неопределённости классифицировать поведение пользователя в информационной среде и определять необходимые меры по противодействию возможной реализации угрозы.

Ключевые слова: информационная безопасность, система мониторинга за действиями пользователя, сети Байеса, вредоносные действия пользователя, несанкционированный доступ, программно – аппаратные средства защиты информации

A. A. Emelina[1], N. E. Karpova[1], A. A. Saranskii[1]*

Research of User Actions in the Information Enviroment of the Enerprise

[1] Samara State Technical University, Samara, Russian Federation
*e-mail: esib@samgtu.ru

Abstract. An urgent task is the development and improvement of methods for detecting unacceptable events in the functioning of an information system, which, most commonly, are the result of user's actions in the computer system. The main difficulty faced by the developers of such systems is that it is necessary to choose a method that could describe the whole variety of human behavioral characteristics. The article presents a classification of threats emanating from a person, as well as analyzes existing methods for detecting and preventing these threats. This paper proposes a mathematical model of a Bayesian network and an algorithm of a system for determining the anomalous behavior of employees in the information environment of an enterprise. As a result of the analysis, it was concluded that Bayesian networks are the most flexible and adequate mathematical apparatus that allows, even under conditions of some uncertainty, to classify user's behavior in the informational environment and to determine the necessary measures to counter the possible implementation of the threat.

Keywords: information security, user activity monitoring system, Bayesian networks, malicious user actions, risk, unauthorized access, classification of information threats

Введение

С растущим объемом информации, которая обрабатывается в электронном виде и передается между различными информационными системами (ИС) в сети Интернет, организации и отдельные пользователи все чаще сталкиваются с необходимостью обеспечения её безопасности.

В настоящее время важной задачей является разработка и совершенствование методов обнаружения недопустимых событий в работе информационной системы. Обычно такие события происходят из-за неправомерных или ошибочных действий пользователей в компьютерной среде. Исследования показывают, что самой большой угрозой для любого предприятия являются его сотрудники. Поэтому необходимо выявлять как новые типы нарушений работы сети, так и вредоносные действия, продолжительные во времени. В настоящее время все большее значение приобретают программные и аппаратные средства защиты информации [1].

Мониторинг за действиями пользователя позволяет отслеживать не только текущее состояние, но и изменения как в динамической системе.

Мониторинг безопасности информационной среды требует анализировать все виды угроз, однако если естественные факторы достаточно просто формализируются в плане рисков и защита от них достаточно понятна, то к угрозам, имеющим причиной человеческий фактор необходимо особое внимание, так как невозможно предсказать действия человека даже при условии того, что он не намерен причинить вреда. Разработчики моделей сталкиваются с трудностями, поскольку все методы, связанные с описанием поведения человека, трудно поддаются формализации. Однако стоит отметить, что поведенческие системы более гибкие по сравнению с биометрическими и могут использоваться для анализа действий пользователя в информационной среде, включая выявление ранее неизвестных аномалий поведения. В общем, мониторинг действий пользователя может рассматриваться как непрерывное наблюдение за факторами, влияющими на функционирование информационной среды, а также как анализ результатов этого наблюдения.

Существует несколько видов классификации информационных угроз, например, угрозы делят по факторам возникновения.

Информационные угрозы по фактору возникновения:

а) природные угрозы;

б) человеческий фактор:

1) умышленные угрозы:

1.1 активные угрозы;

1.2 пассивные угрозы;

1.3 внутренние угрозы;

1.4 внешние угрозы;

2) непреднамеренные угрозы.

Заметим, что большая часть угроз информационной безопасности связана с отсутствием у сотрудников необходимых компетенций, а также неисполнение служебных инструкций.

Также важным компонентом для разработки эффективной системы обнаружения вторжений в критическую инфраструктуру являются наборы данных, характеризующие различные виды атак (в т.ч. эксплуатация критических уязвимостей), а также анализ исходящего сетевого трафика.

Методы мониторинга:

– анализ эксплуатаций уязвимостей;
– анализ ресурсов, к которым обращается пользователь;
– анализ входящих соединений, с потенциально – опасных ресурсов.

В [4] приведен обзор некоторых наиболее важных аспектов безопасности Active Directory. Автор подчеркивает, что Active Directory является ключевым компонентом большинства сетей Windows и важно обеспечивать ее безопасность. Первая область, на которую обращает внимание автор, – это авторизация. Авторизация в Active Directory определяет, кто может получить доступ к данным и ресурсам в сети. Для обеспечения безопасности в этой области автор рекомендует использовать принцип наименьших прав, который позволяет ограничить доступ пользователей только к необходимой им информации.

Вторая область – это аудит безопасности. Логи событий в Active Directory могут использоваться для идентификации попыток несанкционированного доступа к данным и для определения уязвимостей в системе. Автор советует настроить систему аудита таким образом, чтобы она получала логи событий с наивысшим приоритетом.

Нейросетевой детектор атак, предложенный в работе [5], идентифицирует пользователя на основе количества запусков различных команд в течение дня. Авторы используют подход, основанный на машинном обучении с использованием многослойной нейронной сети. Для обучения был использован набор данных, содержащий информацию о пользователе и его поведении при работе с компьютером. В данной модели учитывается только количество команд и не учитывается их последовательность. Кроме того, количество подаваемых на вход нейронной сети команд ограничено (100 команд), хотя в реальных условиях оно может быть значительно выше.

Для описанных выше систем были разработаны математические модели, которые описывают действия пользователя. Примеры таких моделей включают модели принятия решения на основе теории вероятности с помощью экспертов, модели оперативного контроля с использованием математической статистики, модели, которые описывают схемы и потоки информационной системы на основе теории графов, модели, которые используют нечеткие множества, системы, которые используют нейронные сети, и методы раннего обнаружения внутренних нарушителей информационной безопасности, основанные на сетях Байеса [6 - 9].

Угрозами информационной безопасности (ИБ) могут являться кража информации, изменение конфигурации, шифрование данных и пр. Поэтому крайне важно максимально своевременно обнаруживать и предотвращать данные угрозы. Для решения данной задачи предлагаем использовать систему, построенную на основе сети Байеса.

Для составления математической модели обнаружения нарушителя ИБ требуются более глубокие исследования его поведения, поэтому выявление внутренних нарушителей ИБ с использованием байесовских сетей позволяет избавиться от недостатков, присущих описанным выше методам и средствам, а также создать инструмент для осуществления полноценного анализа поведенческих особенностей человека.

Методы и материалы

С помощью теоремы Байеса можно вычислить вероятность события при условии, что произошло другое событие, которое связано с ним статистически. Формула Байеса позволяет точнее учесть известную информацию и учесть новую информацию, полученную в результате наблюдений.

Байесовская сеть позволяет получать ответы на следующие типы вероятностных запросов [9]:

– поиск вероятности доказательства;

– определение априорных предельных вероятностей;

– определение апостериорных маргинальных вероятностей, в том числе:

а) предсказание или прямой вывод, - определение вероятности события по наблюдаемым причинам;

б) диагностика, или обратный вывод (похищение), - определение вероятности возникновения причины с наблюдаемыми последствиями.

Байесовские сети относятся к категории вероятностных графических моделей

Байесовская сеть представляет собой граф, которой ориентирован и ацикличен, где каждая вершина соответствует случайной величине, а связи между вершинами отображают условную независимость между этими переменными. В графе могут быть представлены различные типы переменных с помощью взвешенных параметров, скрытых переменных или гипотез. Модели неопределенных ориентированных графов основаны на изменении вероятностного происхождения событий, где для каждого случайного значения применяется таблица условной вероятности. Это позволяет моделировать вероятностную последовательность событий. Формула Байеса в общем виде (1):

$$P(A|B)=\frac{P(B|A)P(A)}{P(B)}, \quad (1)$$

где P(A) – априорная вероятность гипотезы; P(A|B) – вероятность гипотезы A при наступлении события B (апостериорная вероятность); P(B|A) – веро-

ятность наступления события B при истинности гипотезы A; P(B) – полная вероятность наступления события B.

Преимущества байесовских сетей [11]:

– в модели устанавливаются связи между всеми переменными. Это позволяет легко обрабатывать те случаи, когда значения некоторых параметров отсутствуют;

– Байесовские сети позволяют проводить анализ возможных вариантов при прогнозном моделировании.

Все вышеизложенные преимущества делают применение сетей Байеса оправданным для моделирования поведения пользователя, так как данный математический аппарат позволяет учитывать инвариативность поведения пользователя, прогнозировать вероятность наступления события даже в условии некоторой неопределенности, а также находить связи между разными рассматриваемыми параметрами и находить зависимость между ними, что позволяет с более высокой точностью отнести поведение пользователя к категории «аномальное», «подозрительное» или «нормальное» поведение.

Для выявления нарушителей ИБ необходимо накопить информацию о поведенческих особенностях человека, после чего определить эталон поведения пользователя. При отклонении поведения пользователя от эталона пользователь переходит в группу потенциальных злоумышленников. В проектировании сложных систем, как правило, используется метод экспертных оценок. Сущность методов экспертных оценок заключается в том, что в основу принятого решения, прогноза, вывода закладывается мнение специалиста или коллектива специалистов, основанное на их знаниях и практическом профессиональном опыте. Человек с профильным образованием и опытом в области исследования считается экспертом.

Результаты

С математической точки зрения сеть Байеса – это модель для представления вероятностных зависимостей, а также отсутствия этих зависимостей.

Выходом каждой сети является мера принадлежности события к конкретному классу нарушений ИБ.

В нашем случае, выходом сети является рекомендация сотруднику службы безопасности или системе реагирования на инциденты о необходимой мере, которую следует применить к учетной записи или конечному узлу информационной системы, а именно:

1) заблокировать;

2) подозрение на инцидент – требуется проведение дополнительной экспертизы;

3) легитимные действия пользователя.

Сначала необходимо выявить определенные триггеры.

К компьютерным триггерам можно отнести аномальную почтовую активность, аномальную активность в мессенджерах, анализ журналов событий сред-

ствами Microsoft Windows, а именно события с индикатором 5136 – 5145. Данные события позволяют проанализировать доступ к объектам сетевого ресурса, выявить тип доступа, с которым пользователь обращался к объекту, а также определить, были ли модифицированы объекты. События с индикатором 4624 регистрируют все успешные входы в систему с указанием времени, учетной записи, типом входа, а также хоста-инициатора и целевого хоста и пр. Подробное описание данных событий описано в [14].

Также анализ событий средствами антивирусного программного обеспечения (АВПО), например, Kaspersky, а именно анализ действий со съемными устройствами.

Далее необходимо создать сеть Байеса. В данном случае можно предположить, что все триггеры являются условно независимыми друг от друга. Это позволяет считать, что кража информации может произойти при срабатывании любого триггера, а также при срабатывании нескольких триггеров одновременно. Список факторов, влияющих на наступление инцидент ИБ, может быть уменьшен или увеличен для каждой конкретной компании. Каждому триггеру также будет присвоено свое значение вероятности.

Данные вероятности можно получить различными методами:

– методами экспертных оценок (опрос n-го количества экспертов);

– на основе субъективной оценки определенного человека (например, руководителя отдела информационной безопасности) и пр.

После чего составляется сеть Байеса, а также заполняется таблица априорных вероятностей для каждого триггера.

Пример моделирования сети Байеса в программе GeNie представлен на рис. 1.

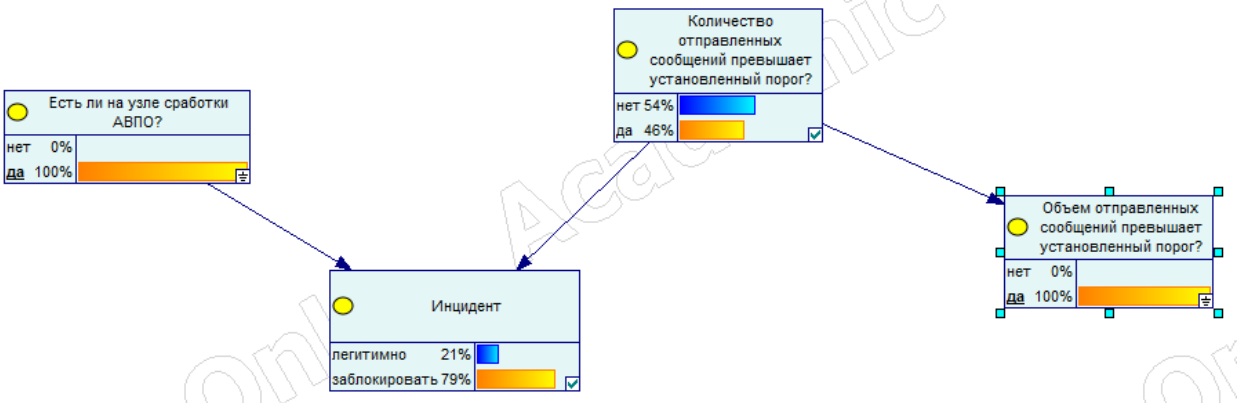

Рис. 1. Модель Сети Байеса для выявления нарушения информационной безопасности

На основании проведенного исследования был предложен алгоритм (рис.2) выявления неспецифичных действий пользователя в информационной системе,

основанный на сетях Байеса, затем на основе данного алгоритма разработан программный модуль для выявления подозрительной активности пользователя.

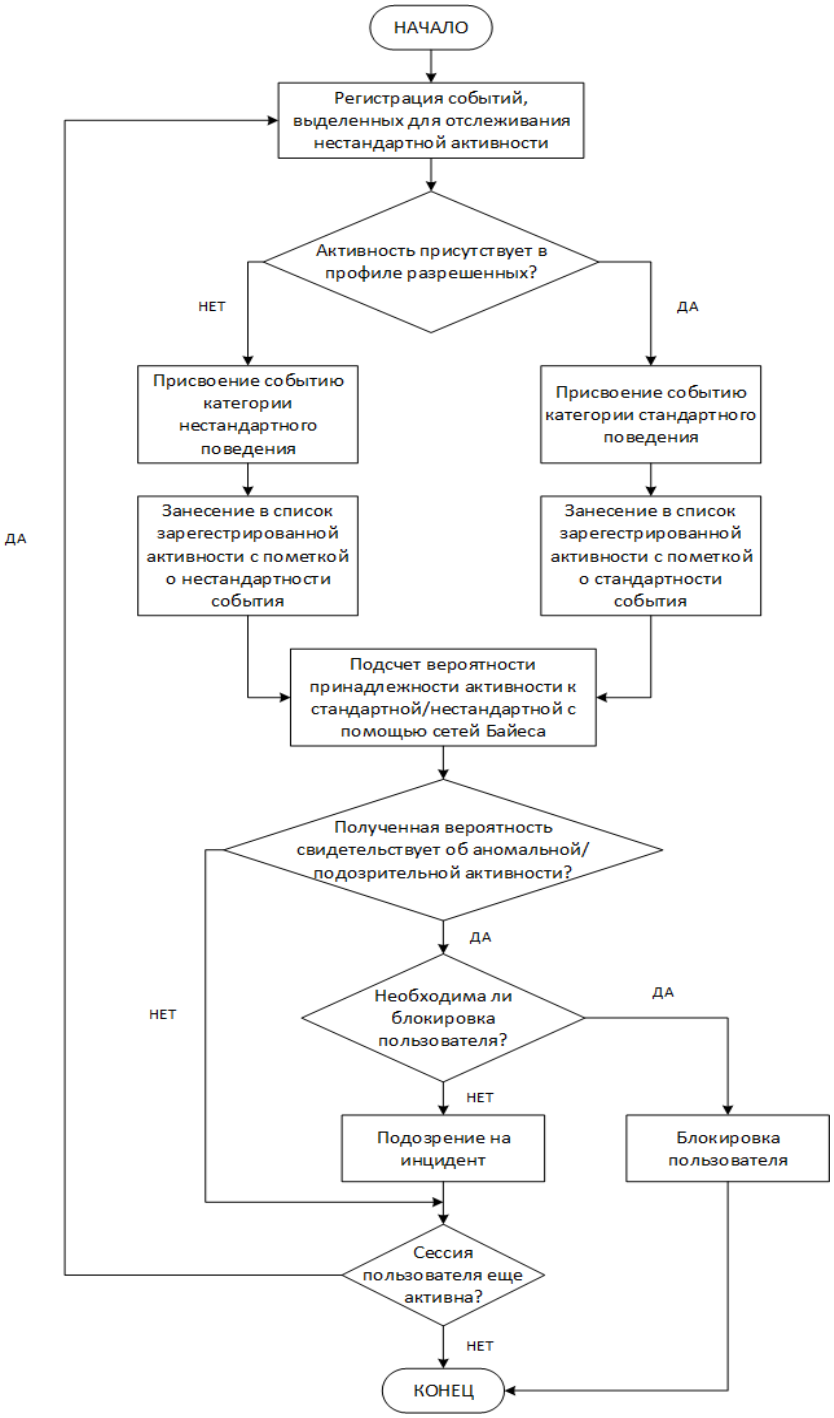

Рис. 2 Алгоритм анализа действий пользователя в информационной системе

Заключение

Сети Байеса являются наиболее гибким и адекватным математическим аппаратом, позволяющим даже в условиях некоторой неопределённости классифицировать поведение пользователя в информационной среде и определять необходимые меры по противодействию возможной реализации угрозы.

Таким образом в статье был произведен анализ угроз, исходящих от действий пользователя, проанализированы существующие системы мониторинга, а также рассмотрены математические аппараты, применяющиеся для задачи мониторинга действий пользователя. Авторами был выбран наиболее гибкий математический аппарат для определения риска кражи информации на предприятии, составлена байесовская сеть и предложена структурная схема системы, построенной на основе сети Байеса.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Tara Seals Fear of Insider Threats Hits an All-Time High, URL: https://www.infosecurity-magazine.com/news/fear-of-insider-threats-hits-an/ (дата обращения 14.03.2023).
2. Karpova N., Panfilova I. Ensuring the Safety of Information Processes in Sociotechnical Systems Based on an Analysis of the Behavioral Characteristics of a Person as a Subject of Such a System // XXI International Conference Complex Systems: Control and Modeling Problems (CSCMP), Samara, 2019. P. 751–753.
3. Sriram Raghavan S. V. Raghavan SDN Security: Developing an organic escalation framework for operational automation on security incidents //SpringerLink. URL: https://link.springer.com/article/10.1007/s40012-020-00266-8. (дата обращения 18.04.2023).
4. Doug White. Three Key Areas in Active Directory Security/ Security Weekly. URL: https://securityweekly.com/2018/09/06/three-key-areas-in-active-directory-security/ (дата обращения 02.03.2023).
5. Obaidat M.S., Macchairolo D.T. A multilayer neural network system for computer access security. EEE Trans. On Syst., Man. And Cybern. Vol. 24, No 5. Pp. 806-813, (1994).
6. Domanetska I., Khaddad A., Krasovska H., Yeremenko B. Corporate System Users Identification by the Keyboard Handwriting based on Neural Networks. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 2019, vol. 9, iss. 1, pp. 4156–4161.
7. Elike H., Xavier B., Andrew H., Pierre-Louis D., Ephraim I., Christos T., Robert A. Threat analysis of IoT networks using artificial neural network intrusion detection system. Proceedings of the 3th International Symposium on Networks, Computers and Communications (Hammamet, Tunisia, 11-13 May 2016). New York, IEEE, 2016. DOI: 10.1109/ISNCC.2016.7746067.
8. Zvyagin L.S. Iterative and non-iterative methods of Monte Carlo as actual computing methods Bayesian analysis. Proceedings of 20th IEEE International Conference on Soft Computing and Measurements (St. Petersburg, Russia, 24-26 May 2017). New York, IEEE, 2016. DOI: 10.1109/SCM.2017.7970482.
9. Zhou Z.-H. Rule Extraction: Using neural networks or for neural networks? Journal of Computer Science and Technology, 2004, vol. 19, iss. 2, pp. 249–253.
10. Adnan Darwiche. Modeling and Reasoning with Bayesian Networks. – Cambridge University Press, 2009. – 526 p.
11. Judea Pearl. Causality: Models, Reasoning, and Inference. – 2-nd Edition. – Cambridge University Press, 2009. – 464 p.
12. D. MacKay. Information Theory, Inference, and Learning Algorithms. - Cambridge University Press, 2003 – 640 p.
13. Звягин Л.С. Метод байесовских сетей и ключевые аспекты байесовского моделирования / Л.С. Звягин // XXII Международная конференция по мягким вычислениям и измерениям (SCM-2019). Сборник докладов. Санкт-Петербург. 23-25 мая 2019 г. - СПб.: СПбГЭТУ «ЛЭТИ». - С. 30-34.
14. Security auditing. URL: https://docs.microsoft.com/ru-ru/windows/security/threat-protection/auditing/security-auditing-overview (дата обращения 04.04.2023).