DOI: 10.17117/na.2016.12.02.183

<http://ucom.ru/doc/na.2016.12.02.183.pdf>

Поступила (Received): 18.12.2016

Храмлюк С.А., Лапина Е.В.
Мониторинг работы пользователей систем
электронного документооборота

Khramlyuk S.A., Lapina E.V.
Monitoring of users of electronic document management systems

В работе проведен обзор методов анализа поведения пользователей в информационных системах. Предложен подход к разработке методики мониторинга работы пользователей СЭД, основанной на интеллектуальном анализе контентной информации и журналируемых событий систем электронного документооборота

Ключевые слова: анализ поведения пользователей, системы электронного документооборота

The paper reviewed the methods of analysis of user behavior in information systems. An approach to the development of a methodology for monitoring progress of EDM users, based on the intelligent analysis of the content of information and journaling of events of electronic document management systems

Key words: analysis of user behavior, an electronic document management system

Храмлюк Софья Анатольевна

Студент

Сибирский государственный аэрокосмический университет им. М.Ф. Решетнева
г. Красноярск, пр. им. газеты «Красноярский рабочий», 31

Khramlyuk Sofia Anatolievna

Student

Siberian state aerospace university named M.F. Reshetnev
Krasnoyarsk, Krasnoyarsky Rabochy ave., 31

Лапина Елена Владимировна

Старший преподаватель

Сибирский государственный аэрокосмический университет им. М.Ф. Решетнева
г. Красноярск, пр. им. газеты «Красноярский рабочий», 31

Lapina Elena Vladimirovna

Senior Lecturer

Siberian state aerospace university named M.F. Reshetnev
Krasnoyarsk, Krasnoyarsky Rabochy ave., 31

Несмотря на все методы и средства защиты информации в СЭД, одной из основных и статичных угроз являются действия легитимных пользователей СЭД. Полномочия таких пользователей довольно обширны, а средства защиты информации отслеживают только несанкционированный доступ. Таким образом, авторизованный пользователь или нарушитель/вредоносная программа под видом авторизованного пользователя остаются незамеченными для СЗИ от НСД. Важной задачей является контроль работы, как отдельных легитимных пользователей, так и групп пользователей.

Целью такого контроля является выявление действий легитимных пользователей, которые могут предшествовать нарушениям информационной

безопасности. Поэтому задача поиска и анализа действий пользователей, приводящих к нарушениям информационной безопасности, является актуальной. Как показывает практика, в большинстве случаев внутреннему вторжению предшествует активность пользователя, которая является легитимной, но является нехарактерной с точки зрения модели поведения пользователя или пользовательской группы. Особенностью обнаружения такой активности является тот факт, что всевозможные нелегитимные действия пользователей невозможно описать заранее и тем более формализовать. Поэтому решение данной задачи заключается в построении модели поведения пользователей и мониторинге за параметрами и характеристиками работы пользователей в системе. Целью мониторинга при этом является выявление соответствующих фактов или нарушений в работе.

Существующие подходы. На сегодняшний день для обнаружения интересующих фактов работы (активности) успешно применяется сигнатурный поиск, основанный на сравнении текущей активности с набором правил [1]. Правила задаются экспертом и описывают интересующие шаблоны активности или характеристики функционирования. Параметры текущей активности сравниваются с правилами, и в случае удовлетворения хотя бы одному из правил, система генерирует оповещение соответствующего типа. Основной проблемой использования сигнатурного подхода является невосприимчивость к еще неописанным экспертом шаблонам. Таким образом, метод не работает, пока не будет обновлена база сигнатур.

Обнаружение аномалий предполагает, что любые действия, отличающиеся от «обычной» активности, интерпретируются системой как аномалия. Для этого строится модель нормального поведения, а текущая активность сравнивается с моделью. Спецификой такой задачи является необходимость поиска неявных закономерностей в больших объемах данных. Поэтому для построения поведения пользователей и поиска аномалий требуются методы на основе алгоритмов машинного обучения и методов интеллектуального анализа данных, позволяющие выявлять скрытые закономерности и факты [2]. Минусом подхода, основанного на поиске аномалий, является то, что система не может самостоятельно без вмешательства аналитики отличить типы и корректность данных аномалий.

Особенностью мониторинга пользователей СЭД является необходимость сбора узкоспециализированной информации о работе пользователей в системе и совокупного анализа параметров работы. Для анализа требуется объединение информации, собранной из учетной записи пользователя. Указанная задача может решаться различными методами, однако с ростом пользователей СЭД становится необходима консолидация и централизованное хранение всех собранных параметров и характеристик функционирования.

Одним из основных вопросов построения систем мониторинга является выбор исходных данных. Можно выделить следующие основные источники в СЭД:

1) контентная информация – содержательные данные, с которыми работают пользователи;

2) журналируемая информация – данные из системного журнала СЭД.

На основе контроля перемещения контентной информации с большой вероятностью могут обнаруживаться утечки корпоративных данных. Существует несколько видов контентного анализа, такие как цифровые отпечатки, файловые метки, анализатор шаблонов, лингвистический анализ. Каждая из этих технологий имеет свои преимущества и ограничения [3].

Однако работа с контентной информацией имеет сложности:

- контентная информация является объемной, поэтому задачи ее консолидации, хранения и анализа – ресурсоемкие;
- контентная информация не формализована, требуется разработка методов извлечения требуемых данных из контента различных форматов;
- анализ контентной информации потенциально снижает уровень защищенности СЭД из-за предоставления экспертам, разрабатывающим правила анализа, ключевых слов или иных признаков конфиденциальных документов и доступа СЗИ ко всей информации СЭД.

Использование событий журналируемой информации позволяет избежать обозначенных проблем. Журналируемая информация легко формализуется, содержит некоторые параметры и характеристики работы пользователей. Такая информация не компрометирует безопасность СЭД, так как не содержит конфиденциальной информации. Однако журналируемая информация СЭД не содержит достаточных данных для построения модели поведения пользователя, а позволяет отслеживать лишь общие параметры работы в системе, такие как фиксация попыток удаления документа, время работы в системе и т.д.

Предлагаемое решение. В работе предлагается создание методики, основанной на интеллектуальном анализе контентной информации и журналируемых событий СЭД. Методика должна решать следующие задачи:

- 1) сбор и хранение статистики работы пользователей и приложений в СЭД;
- 2) обработка контентной информации и журналируемых событий СЭД;
- 3) построение моделей поведения пользователей;
- 4) анализ исходных данных;
- 5) выявление аномалий поведений и угроз информационной безопасности.

Процесс обработки информации заключается в формировании из большого множества исходных собираемых событий нескольких сотен значимых событий. Можно выделить следующие общие этапы формирования значимых событий [4] (Рисунок 1).

Для решения каждого конкретного типа задач требуется специализированный набор данных [5]. В связи с этим, актуальным является разработка унифицированных моделей и методов, позволяющих производить сбор, консолидацию и анализ журналированных событий и анализ контентной информации, с учетом особенностей и ограничений использования систем электронного документооборота и особенностей работы в этих системах.

Рис. 1. Этапы обработки событий

На основе собранной активности пользователей может быть построена математическая модель поведения пользователя. Модель строится на основе набора записей, отражающих легитимное поведение пользователей. Алгоритм выявления аномалий «обучается» на построенной модели. Анализ предлагается проводить методами Data Mining.

Следующими этапами работы по разработке методики системы мониторинга работы пользователей СЭД являются:

- 1) определение перечня, свойств и атрибутов исходных данных, необходимых для анализа;
- 2) разработка алгоритма сбора, хранения и консолидации данных;
- 3) разработка (или выбор из существующих) алгоритмов построения и обучения моделей поведения;
- 4) разработка (или выбор из существующих) алгоритмов интеллектуального анализа данных.
- 5) разработка архитектуры системы мониторинга.

Список используемых источников:

1. Amoroso, Edward, G., *Intrusion Detection* // 1st ed., Intrusion.Net Books, Sparta, New Jersey, USA, 1999.
2. Jiawei Han, MichelineKamber // *Data Mining. Concepts and Techniques*. USA, Academicpress. 2001. С. 5-16.
URL: https://www.infowatch.ru/sites/default/files/docs/infowatch_content_filtering_database_whitepaper_russian.pdf
3. *Безопасность корпоративных сетей: мониторинг, анализ, управление*.
URL: http://storagenews.ru/30/sec_corp_network-1.pdf
4. Трошин С. В. *Консолидация и предобработка информации из журналов регистрации ВС для систем обнаружения вторжений* // Программные системы и инструменты. №6. М.: МАКС Пресс, 2005. С. 68-80.