Инженерный вестник Дона, №6 (2024)
ivdon.ru/ru/magazine/archive/n6y2024/9285

Определение злонамеренного поведения и формирование профиля злоумышленника в системах управления обучением

Н.А. Еритенко, А.А. Менщиков

Национальный исследовательский университет ИТМО, Санкт-Петербург

Аннотация: В процессе обеспечения информационной безопасности важным элементом является защита данных от злонамеренного воздействия. Один из этапов защиты - определение источника угрозы. Источником угрозы в ряде информационных систем может выступать сам злоумышленник, действуя через собственный профиль пользователя. Детектирование и определение злонамеренного профиля в данном случае – ключевой элемент построения защиты. В контексте профилирования пользователя хорошо исследованными системами являются социальные сети. Системы управления обучением, несмотря на свою схожесть с социальными сетями, остаются лишенными научного внимания. Ввиду схожести с социальными сетями и все большего внедрения систем управления обучением в информационные структуры, данный тип систем становится уязвимым звеном. В данной работе освещены аспекты определения злонамеренной активности и формирования профиля в системах управления обучением и приведен пример одного из методов.
Ключевые слова: информационная безопасность, информационная система, социальная сеть, система управления обучением, профилирование пользователя, злонамеренное поведение, классификация.

Введение. Определение злонамеренных пользователей в системах управления обучением

Определение и детектирование злоумышленника является одним из аспектов обеспечения информационной безопасности и составление профиля пользователя может играть ключевую роль в этом процессе. Для профилирования пользователя используются различные методы и техники, включая подходы, связанные с искусственным интеллектом и машинным обучением. Профилирование в общем случае может быть классифицировано двумя фундаментальными подходами: контент-ориентированными (явными) методами и коллаборативными (в т.ч. известными, как поведенческие или неявные) методы [1, 2]. Существует множество приложений для профилирования пользователя: предоставление услуг и сервиса [1], маркетинг, таргетинг и реклама [3], финансовый сектор и цифровая экономика [4], адаптивное обучение [5], и множество других.

В настоящее время хорошо изученным классом информационных систем с точки зрения обнаружения злонамеренного поведения являются социальные сети. Социальные сети считаются наиболее важным и уязвимым классом среди всех систем в рассматриваемом контексте. Одной из главных причин является то, что многие социальные сети были идентифицированы как платформы для скрытой террористической деятельности и использовались в качестве средства коммуникации для спящих ячеек (включая Twitter, Reddit и Facebook) [6]. Другая, не менее важная причина заключается в том, что злоумышленники часто изменяют количество просмотров, портят систему социальных сетей методами социальной инженерии и используют другие приемы, которые приводят к снижению доверия к платформе среди пользователей [7]. Однако, приложениям электронного обучения (E-learning), которые, как правило, аналогичны социальным сетям, уделяется критически мало внимания.

Согласно недавним исследованиям [1, 8], средства электронного обучения можно разделить на три группы: системы управления контентом (Content Management Systems – CMS), системы управления обучением (Learning Management System – LMS) и системы управления учебным контентом (Learning Content Management Systems – LCMS). CMS позволяют нам управлять учебными ресурсами и часто поддерживают все виды коллаборативного редактирования. LMS, в свою очередь, это программное обеспечение, созданное для управления учебными курсами и их проведения, которое может содержать некоторые дополнительные инструменты для оценки прогресса учащихся или визуализации успеваемости. LCMS представляют собой гибриды CMS и LMS с пропорциями функций, которые зависят от контекста системы. Недавние исследовательские работы показывают, что целый класс инструментов электронного обучения постоянно развивается и эволюционирует [9, 10]. Второе поколение систем

управления обучением (Learning Management System 2.0 – LMS 2.0), появившееся в 2004 году [10], интегрировало аспекты социальной сети и, как следствие, со временем внедрило некоторые их функции, попутно включая различные угрозы и уязвимости.

Основная цель данной работы - описать и проанализировать разницу и схожесть процесса профилирования в LMS 2.0 и социальных сетях и предложить подход к анализу аномального поведения и профилирования. LMS 2.0 специфичны в контексте рассмотрения проблемы вредоносного поведения и, как следствие, специфичны для формирования профиля злоумышленника. В работе мы привели потенциальный гибридный метод анализа аномальной активности и профилирования на примере LMS 2.0 с использованием интеллектуального анализа данных.

Социальные сети и системы управления обучением

Как упоминалось ранее, существуют сходства и различия в информационных системах, которые относятся к классу социальных сетей и LMS 2.0. Оба класса выполняют различные функции по облегчению взаимодействия с пользователями и обмена контентом, но они существенно различаются по своим основным целям и организационным структурам. Процесс электронного обучения сопровождается определенной коммуникацией посредством взаимодействия и сотрудничества через информационную систему. Имеется большое количество инструментов для осуществления связи в онлайн-обучении. В работе [11] приведен анализ и сравнение ключевых характеристик и инструментов LMS-систем, в числе которых можно выделить следующие элементы, имеющие аналог в социальных сетях:

Таблица № 1

Сравнение инструментов в LMS 2.0 и социальных сетях

Инструмент LMS 2.0	Инструмент/аналог в социальных сетях
Группировка и управление обучающимися	Группы. Управление группами
Уведомление пользователей	Система подписок и уведомлений. Подписка на группы, чаты, новостные ленты
Видеоконференции	Видео- и аудио-звонки
Чаты и форумы, комментарии к заданиям	Диалоги, групповые чаты и обсуждения, стена, комментарии к медиаконтенту и новостям
Журнал успеваемости, результаты и прогресс обучения	Механизм достижений, прогресс в приложениях, заполненность секций контентом, числовые параметры и значения (число друзей, число фотографий)
Задания и тесты (контент преподавателя)	Медиаконтент: фото, видео, аудио (контент автора/владельца страницы)

Социальные сети и LMS 2.0 упрощают взаимодействие с пользователями и обмен контентом. Они позволяют пользователям создавать профили, входить в систему и получать доступ к контенту. Оба класса систем также позволяют пользователям обмениваться контентом, таким как записи, сообщения, файлы и мультимедиа. Кроме того, как социальные сети, так и LMS 2.0 управляют учетными записями, обеспечивая структурированную структуру для взаимодействия с пользователями.

Основной целью социальных сетей является социальное взаимодействие, развлечение и создание сообщества, в то время как LMS 2.0

предназначена для образовательных целей, таких, как преподавание, обучение и развитие навыков. Социальные сети обычно содержат пользовательский контент, такой, как посты, комментарии и мультимедиа, в то время, как LMS 2.0 ориентированы на структурированный образовательный контент, такой как курсы, уроки, задания и оценки. В социальных сетях пользователи часто играют схожие роли (друзья или подписчики), в то время как в LMS 2.0 у пользователей разные роли (преподаватели, студенты и администраторы), с разными уровнями доступа и ответственности. LMS 2.0 предназначены для упрощения механизмов оценки и обратной связи (викторины, экзамены и системы выставления оценок), которые обычно не используются в социальных сетях.

Подходы к анализу аномального поведения

Для обнаружения вредоносных аномалий существует несколько подходов и методов, в числе которых:

1. Статистические методы: Эти методы используют статистические модели для обнаружения аномалий на основе закономерностей в исторических данных. Однако эти методы могут быть чувствительны к изменениям в распределении данных и могут не обнаруживать новые атаки [12].

2. Методы машинного обучения: эти методы используют алгоритмы машинного обучения для изучения моделей нормального поведения и обнаружения отклонений, указывающих на вредоносную активность. Однако эти методы могут быть дорогостоящими с точки зрения вычислений и могут потребовать больших объемов помеченных обучающих данных [13].

3. Методы, основанные на правилах: в этих методах используются предопределенные правила для обнаружения аномалий на основе таких критериев, как шаблоны системных вызовов или сетевой трафик. Однако

способность этих методов обнаруживать сложные или новые атаки может быть ограничена [14].

Предлагаемый нами подход сочетает в себе систему, основанную на правилах, с алгоритмами машинного обучения для повышения точности обнаружения вредоносных аномалий. Система, основанная на правилах, предоставляет детальное представление о потенциальных аномалиях, в то время как алгоритм машинного обучения предоставляет более детальное представление о конкретных атаках. Обобщенная схема процесса классификации профиля приведена на рисунке 1.

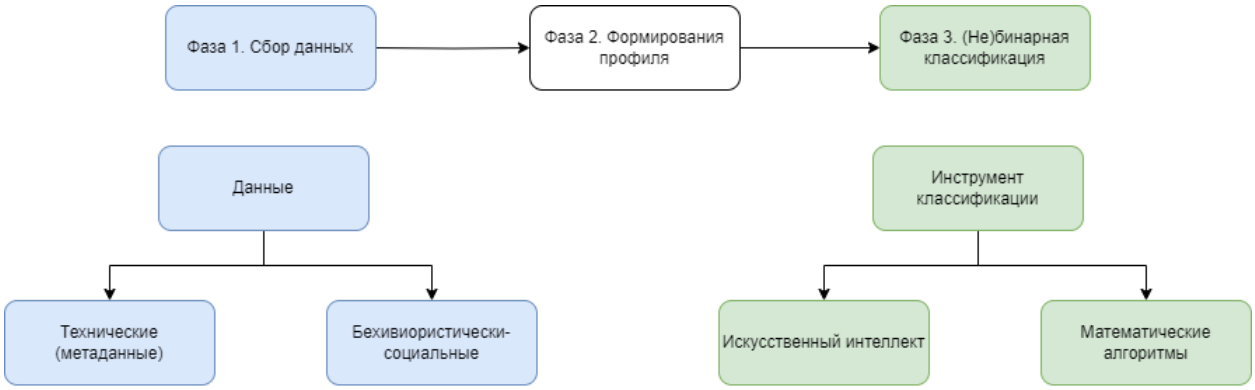

Рис. 1. – Процесс классификации

Анализ аномального поведения на основе правил и методах машинного обучения

На основе собранных на первом этапе данных можно провести анализ поведения и сформировать профиль каждого пользователя.

Простейший пример создания фильтра аномального поведения на основе правил выглядит следующим образом: по каждому пользователю ведется счетчик действий, которые нарушают правила. Изначально значение счетчика – нуль, но при отклонении от нормального поведения значение счетчика увеличивается. Когда счетчик достигает порогового значения, профиль пользователя помечается с помощью «красного флага» –

специальной отметки о том, что его поведение или действия являются подозрительными.

Каждый из пунктов правил требует настройки «нормы» и чувствительности к выбросам в виде некоторого предельного значения, превышение которого приводит к увеличению счетчика «красного флага».

Перечислим возможные для пользователя LMS 2.0 правила/параметры:

1. Числовой идентификатор устройства в компьютерной сети (IP адрес), устройство, браузер, прочие параметры браузера. При изменении адреса интернет-протокола, устройства, с которого был совершен вход в систему больше порогового значения, означает увеличение счетчика «красного флага».
2. Время, длительность и частота сеансов входа в систему, активность в различных разделах, подразделов курсов. Производится сбор сведений о среднем времени сеанса пользователя, в соответствии с иерархической структурой LMS 2.0.
3. Ввод информации и нажатия. Отслеживаются изменения клавиатурного почерка, паттернов кликов, скроллинга страниц.
4. Ввод правильного ответа с первого раза.
5. Ввод чужих ответов.
6. Время, затраченное на решение задания относительно нормы.
7. Переходы по ссылкам на сопроводительные материалы.
8. Поведение в комментариях.
9. Вводимая в поле ответа информация.
10. Информация из единой системы профилей.

При обнаружении нарушении правил проводится анализ потенциальных угроз, которые могут быть осуществлены с помощью этого конкретного набора. Например, сочетания нарушений по пунктам 1, 3, 6 могут означать использование аккаунта другим человеком; сочетания

нарушений по пунктам 4-6 указывают на вмешательство в образовательный процесс и обход системы оценивания.

Система обнаружения злонамеренного поведения, основанная на системе правил, требует тонкой, правильной настройки границы (чувствительности) для каждого из пунктов правил.

Приведенные выше потенциальные правила могут быть проанализированы средствами машинного обучения. К примеру, пункт правил 3 может быть дополнен классификацией по нескольким параметрам. Пункт 8 может быть усилен введением классификатора, основанного на методе опорных векторов (Support Vector Machine – SVM) [15].

Литература

1. Cufoglu A. User profiling - a short review. International Journal of Computer Applications. 2014. Volume 108. № 3. Pp. 1–9.

2. Kanoje S. Girase S. Mukhopadhyay D. User Profiling Trends, Techniques and Applications. International Journal of Advance Foundation and Research in Computer. 2014. Volume 1. Issue 1. URL: doi.org/10.48550/arXiv.1503.07474/.

3. Trusov M., Ma L., Jamal Z. Crumbs of the Cookie: User Profiling in Customer-Base Analysis and Behavioral Targeting. Marketing Science. 2016. Volume 35. No 3. URL: doi.org/10.1287/mksc.2015.0956/.

4. Ilina T. Piatina E. Approaches to Digital Profiling in the Financial Market. Journal of Corporate Finance Research. 2020. Vol. 14. No 4. Pp. 47–60.

5. Froschi C. User Modeling and User Profiling in Adaptive E-learning Systems. Graz, Austria. Faculty of Computer Science. Institute for Information Systems and Computer Media. 2005. 175 p.

6. Al-Qurishi M., Hossain M.S., Airubaian M., Rahman S.M.M., Alamri A. Leveraging Analysis of User Behavior to Identify Malicious Activities in Large-

Scale Social Networks. IEEE Transactions on Industrial Informatics. 2018. Vol. 14. No. 2. Pp. 799–813.

7. Ikwu R., Giommoni L., Javed A., Burnap P., Williams M. Digital fingerprinting for identifying malicious collusive groups on Twitter. Journal of Cybersecurity. 2023. Volume 9. Issue 1. URL: doi.org/10.1093/cybsec/tyad014/.

8. Nath J. Ghosh S., Agarwal S., Nath A. E-learning methodologies and its trends in modern information technology. Journal of Global Research in Computer Science. 2012. Volume 3. No. 4. Pp. 48–52.

9. Maes J-M. Chamilo 2.0: A Second Generation Open Source E-learning and Collaboration Platform. International Journal of Advanced Corporate Learning. 2010. Volume 3. No. 3. Pp. 26–31.

10. Sahin M., Yurdugul H. Learners' Needs in Online Learning Environments and Third Generation Learning Management Systems (LMS 3.0). 2022. Technology, Knowledge and Learning. Volume 27. Pp. 33–48.

11. Исаева Е.С. Современные LMS платформы дистанционного обучения: анализ и сравнение // Педагогика. Вопросы теории и практики. 2021. Том 6. Выпуск 6. С. 1045-1050.

12. Шелухин О.И., Филинова А.С., Васина А.В. Обнаружение аномальных вторжений в компьютерные сети статистическими методами // T-Comm – Телекоммуникации и Транспорт. 2015. Том 9. Выпуск 10. С. 42-49.

13. Labayen V., Magana E., Morato D., Izal M. Online classification of user activities using machine learning on network traffic. Computer Networks. 2020. Volume 181. URL: doi.org/10.1016/j.comnet.2020.107557/.

14. Шкодырев В.П., Ягафаров К.И., Баштовенко В.А., Ильина Е.Э. Обзор методов обнаружения аномалий в потоках данных. The Second Conference on Software Engineering and Information Management. Санкт-Петербург. CEUR Workshop Proceedings. 2017. С. 50-55.

15. Клековкина М.В., Котельников Е.В. Метод автоматической классификации текстов по тональности, основанный на словаре эмоциональной лексики. Материалы XIV Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции». Переяславль-Залесский. CEUR Workshop Proceedings. 2012. С. 118-123.

References

1. Cufoglu A. User profiling - a short review. International Journal of Computer Applications. 2014. Volume 108. No 3. Pp. 1–9.

2. Kanoje S. Girase S. Mukhopadhyay D. User Profiling Trends, Techniques and Applications. International Journal of Advance Foundation and Research in Computer. 2014. Volume 1. Issue 1. URL: doi.org/10.48550/arXiv.1503.07474/.

3. Trusov M., Ma L., Jamal Z. Crumbs of the Cookie: User Profiling in Customer-Base Analysis and Behavioral Targeting. Marketing Science. 2016. Volume 35. No 3. URL: doi.org/10.1287/mksc.2015.0956/.

4. Ilina T. Piatina E. Approaches to Digital Profiling in the Financial Market. Journal of Corporate Finance Research. 2020. Vol. 14. No 4. Pp. 47–60.

5. Froschi C. User Modeling and User Profiling in Adaptive E-learning Systems. Graz, Austria. Faculty of Computer Science. Institute for Information Systems and Computer Media. 2005. 175 p.

6. Al-Qurishi M., Hossain M.S., Airubaian M., Rahman S.M.M., Alamri A. Leveraging Analysis of User Behavior to Identify Malicious Activities in Large-Scale Social Networks. IEEE Transactions on Industrial Informatics. 2018. Vol. 14. No. 2. Pp. 799–813.

7. Ikwu R., Giommoni L., Javed A., Burnap P., Williams M. Digital fingerprinting for identifying malicious collusive groups on Twitter. Journal of Cybersecurity. 2023. Volume 9. Issue 1. URL: doi.org/10.1093/cybsec/tyad014/.

8. Nath J. Ghosh S., Agarwal S., Nath A. E-learning methodologies and its trends in modern information technology. Journal of Global Research in Computer Science. 2012. Volume 3. No. 4. Pp. 48–52.

9. Maes J-M. Chamilo 2.0: A Second Generation Open Source E-learning and Collaboration Platform. International Journal of Advanced Corporate Learning. 2010. Volume 3. No. 3. Pp. 26–31.

10. Sahin M., Yurdugul H. Learners' Needs in Online Learning Environments and Third Generation Learning Management Systems (LMS 3.0). 2022. Technology, Knowledge and Learning. Volume 27. Pp. 33–48.

11. Isaeva E.S. Sovremenny`e LMS platformy` distancionnogo obucheniya: analiz i sravnenie. Pedagogika. Voprosy` teorii i praktiki [Pedagogy. Questions of theory and practice]. 2021. Tom 6. Vy`pusk 6. Pp. 1045-1050.

12. Sheluxin O.I., Filinova A.S., Vasina A.V. Obnaruzhenie anomal`ny`x vtorzhenij v komp`yuterny`e seti statisticheskimi metodami. T-Comm – Telekommunikacii i Transport [T-Comm – Telecommunications and Transport]. 2015. Tom 9. Vy`pusk 10. Pp. 42-49.

13. Labayen V., Magana E., Morato D., Izal M. Online classification of user activities using machine learning on network traffic. Computer Networks. 2020. Volume 181. URL: doi.org/10.1016/j.comnet.2020.107557/.

14. Shkody`rev V.P., Yagafarov K.I., Bashtovenko V.A., Il`ina E.E`. Obzor metodov obnaruzheniya anomalij v potokax danny`x. The Second Conference on Software Engineering and Information Management. Sankt-Peterburg. CEUR Workshop Proceedings. 2017. Pp. 50-55.

15. Klekovkina M.V., Kotel`nikov E.V. Metod avtomaticheskoj klassifikacii tekstov po tonal`nosti, osnovanny`j na slovare e`mocional`noj leksiki. Materialy` XIV Vserossijskoj nauchnoj konferencii «E`lektronny`e biblioteki: perspektivny`e metody` i texnologii, e`lektronny`e kollekcii» [Proceedings of the 14th All-Russian Scientific Conference «Digital libraries: Advanced Methods and

Technologies, Digital Collections»]. Pereyaslavl`-Zalesskij. CEUR Workshop Proceedings. 2012. Pp. 118-123.

Дата поступления: 6.04.2024

Дата публикации: 30.05.2024