В Диссертационный совет 24.2.385.09 при федеральном государственном бюджетном образовательном учреждении высшего образования «Санкт-Петербургский государственный университет промышленных технологий и дизайна»

ОТЗЫВ

официального оппонента к.т.н., доцента Красова Андрея Владимировича на диссертацию Баркова Вячеслава Валерьевича «Классификация противоправных и нежелательных мобильных приложений методами машинного обучения в потоковом режиме», представленную на соискание ученой степени кандидата технических наук по специальности 2.3.6. Методы и системы защиты информации, информационная безопасность.

1 Актуальность темы исследования

Задача обнаружения мобильных приложений, распространяющих противоправный, нежелательный или вредоносный контент, становится особенно актуальной в условиях стремительного развития мобильных технологий. Для решения подобных задач активно применяются методы интеллектуального анализа данных (Data Mining, DM) и машинного обучения (Machine Learning, ML), которые позволяют учитывать изменяющиеся характеристики сетевого трафика Интернет-ресурсов. Использование этих методов обеспечивает высокую эффективность в задачах классификации, анализа и фильтрации сетевого трафика мобильных приложений, распространяющих нежелательный контент.

Однако в существующих исследованиях, посвящённых классификации мобильных приложений на основе анализа сетевого трафика, часто не учитывается задача выявления неизвестного трафика. При построении моделей классификации предполагается наличие лишь известных классов. Обучение проводится на данных, относящихся к ограниченному числу классов, с последующей проверкой модели на других данных, относящихся к тем же самым классам. Недостаток информации о структуре фонового трафика существенно снижает точность классификации.

Целью исследования является повышение эффективности классификации мобильных приложений, осуществляющих распространение противоправного, нежелательного или вредоносного контента, на основе анализа сетевого трафика методами машинного обучения в потоковом режиме. Для достижения поставленной цели диссертантом предложено использовать искусственные нейронные сети (ИНС), в частности, автокодировщики (АК).

Объектом исследования в диссертационной работе является сетевой трафик, генерируемый мобильными приложениями.

Предметом исследования являются методы ML для классификация мобильных приложений на основе анализа сетевого трафика в условиях априорной неопределённости числа приложений.

Таким образом, диссертационное исследование Баркова Вячеслава Валерьевича, направленное на совершенствование алгоритмов классификации мобильных приложений, распространяющих противоправный, нежелательный или вредоносный контент, путём применения методов машинного обучения для анализа сетевого трафика в потоковом режиме в условиях априорной неопределённости, является актуальным.

2 Оценка содержания диссертации

Диссертация содержит введение, четыре главы, заключение, список использованных источников, содержащий 118 наименований, и пять приложений. Основной текст работы изложен на 107 страницах, проиллюстрирован 25 рисунками и 26 таблицами.

Во введении обоснованы актуальность и степень разработанности темы исследования, на основе которых сформулированы цель работы и решаемые задачи для ее достижения.

В первой главе проведён анализ основных подходов к классификации мобильных приложений, распространяющих противоправный, вредоносный и нежелательный контент, с применением методов машинного обучения в условиях априорной неопределенности и динамических изменений характеристик сетевого трафика. Для классификации использованы широко известные алгоритмы машинного обучения, такие как Logistic Regression (LR), K-Nearest Neighbors (KNN), Decision Tree (DT), Random Forest (RF), Gradient Boosting (GB), Naive Bayes (NB), C4.5, AdaBoost, Support Vector Machine (SVM), а также Искусственная Нейронная Сеть (ИНН). Рассмотрены пакетные и потоковые алгоритмы классификации и существующие методы обнаружения смены концепта.

Во второй главе разработан алгоритмов классификации мобильных приложений, распространяющих нежелательный или вредоносный контент, посредством анализа сетевого трафика в offline-режиме.

Задача классификации рассмотрена как в классической постановке обучения с учителем, так в условиях априорной неопределенности о числе и характеристиках классифицируемых приложений, что обусловлено наличием фонового трафика (ФТ).

Предложена методика отбора рационального числа атрибутов на основе их информативности при заданной допустимой вероятности ложной классификации. Для улучшения показателей эффективности классификации мобильных приложений, распространяющих противоправный, вредоносный и нежелательный

контент, в условиях фонового трафика, предложено использовать искусственную нейронную сеть с архитектурой автокодировщик (АК) в качестве предварительного фильтра. Экспериментальные исследования показали, что применение АК в условиях фонового трафика значительно повышает качество классификации по сравнению с лучшим алгоритмом Random Forest (RF). Выигрыш составляет до 5% по метрике Accuracy, и около 2% по метрикам Precision, Recall и F1-score. Важным достоинством использования АК является отсутствие необходимости в разметке фоновых приложений, а также исключение необходимости внедрения дополнительных механизмов идентификации фонового трафика и повторного обучения при увеличении числа таких приложений.

В третьей главе рассмотрена разработка модели обнаружения смены концепта (МОСК) с использованием АК, обучающегося на статистических характеристиках нормального сетевого трафика. Процесс обнаружения смены концепта может быть реализован не только на этапах обучения и тестирования, но и в процессе предсказания, поскольку этот метод не требует наличия истинных меток. Он основывается либо на использовании меток модели классификации, либо на попытке восстановления данных с применением всех АК путём анализа ошибок восстановления. Для учета эффекта «старения» данных в условиях потоковой обработки предлагается МОСК, в которой изменение статистических характеристик атрибутов сетевого трафика мобильных приложений отслеживается с помощью двух скользящих окон во времени. Использование данной модели позволяет снизить вероятность ошибок классификации примерно на 5%. Экспериментально показано, что модели классификации, использующие МОСК с коэффициентами затухания, демонстрируют ухудшение производительности сразу после смены концепта, но в дальнейшем количество ошибок уменьшается на 5–7%, в отличие от моделей, не использующих МОСК.

В четвертой главе представлены результаты исследования потоковой классификации мобильных приложений, распространяющих противоправный, нежелательный или вредоносный контент, на основе анализа сетевого трафика. Рассмотрены два сценария распределения интенсивности трафика.

Первый сценарий предполагает равномерное и непрерывное поступление нежелательных или вредоносных приложений, второй — случайную интенсивность и длительность появления анализируемых приложений. Исследования показали, что алгоритм ARF эффективно решает задачи классификации как при равномерном, так и при случайном распределении сетевого трафика вредоносных или нежелательных мобильных приложений. Этот подход был реализован на базе модификации известного алгоритма ARF. Для обнаружения смены концепта в модификации алгоритма ARF (MARF) используется метод МОСК, основанный на автокодировщиках, который не требует наличия истинных меток. Применение MARF позволяет значительно ускорить процесс классификации в 2–3 раза по сравнению с базовым алгоритмом RF, а также с алгоритмами HAT, KNN и OB. Кроме того, этот подход позволяет эффективно обнаруживать смену

концепта в процессе работы модели, что делает MARF предпочтительным инструментом для решения задач классификации мобильных приложений в реальном времени на основе анализа сетевого трафика. Для автоматизации процесса сбора сетевого трафика и проведения исследований эффективности моделей классификации был разработан программный комплекс «Система анализа трафика» (ПК САТ).

В заключении выделены основные результаты, полученные в диссертации.

Изложение материала диссертационного исследования является структурированным и логичным, применяемая в работе терминология – корректна. Поставленная в работе цель достигнута.

Автореферат диссертации соответствует ее содержанию и достаточно полно отражает полученные научные и практические результаты.

3 Степень обоснованности и достоверность научных положений, выводов и рекомендаций

Обоснованность положений, выносимых диссертантом на защиту, подтверждается применением научного подхода при разработке алгоритмов, корректным их применением при решении поставленных задач.

Достоверность положений, выносимых диссертантом на защиту, подтверждается наличием имитационного моделирования с применением современного математического аппарата, а также внедрением полученных результатов в АО «Лаборатория Касперского» при разработке программного обеспечения для межсетевых экранов.

4 Новизна научных положений, выводов и рекомендаций

Научная новизна полученных автором результатов состоит в следующем:

1. Методика отбора рационального числа атрибутов на основе анализа их информативности при фиксировании допустимой вероятности ложной классификации. В отличии от известных, методика позволяет рассчитывать характеристики по выборки данных, и является инвариантной для различных типов сетевого трафика. Предложенная методика позволяет осуществить классификацию мобильных приложений, осуществляющих шифрование сетевого трафика с 90% вероятностью, при сокращении объема обучающей выборки, и снижает число ложной классификации в 2,5 раза. Отличительным признаком новизны методики является её инвариантность по отношению к разным типам сетевого трафика.

2. Модифицированный алгоритм классификации мобильных приложений в условиях неконтролируемого фонового трафика, отличается от известных алгоритмов каскадным включением нейронной сети с архитектурой автокодировщика, выполняющей предварительную фильтрацию, а также вторичной моделью классификации. Предложенный алгоритм позволяет повысить

достоверность классификации приложений на 7% и при этом не требует по сравнению с известными алгоритмами предварительной разметки фоновых приложений.

3. Статистическая модель обнаружения смены концепта при классификации мобильных приложений на основе анализа сетевого трафика, отличающаяся от известных включением автокодировщика в качестве базовой модели обнаружения смены концепта, в которой момент наступления смены концепта определяется посредством оценок ошибок восстановления анализируемых приложений и превышения пороговых значений, что повышает точность обнаружения смены концепта в реальном времени до 10% и на 5% снижает вероятность ошибки классификации по сравнению с моделями на основе коэффициентов затухания.

4. Новый алгоритм обнаружения смены концепта мобильных приложений в потоковом режиме с обработкой в скользящем окне в режиме накопления с «конечной памятью», как с равномерной, так и неравномерной интенсивностью поступления данных, отличающийся от известных учетом «старения» данных в окне обработки и негауссовским характером изменяющихся параметров классифицируемых приложений.

5. Модифицированный алгоритм Adaptive Random Forest (MARF) со встроенной моделью обнаружения смены концепта, позволяет обнаруживать смену концепта, быстрее чем алгоритмы Random Forest (RF), Hoeffding Adaptive Tree (HAT), K nearest neighbors (KNN), Oza Bagging (OB). Алгоритм отличается учетом ограничения памяти для скользящего окна, учитывающего старение данных, он позволяет осуществить классификацию в 2–3 раза быстрее, чем известные алгоритмы (RF, HAT, KNN, OB).

Научные результаты получены автором самостоятельно, что подтверждается наличием публикаций без соавторов, перечисленные результаты диссертационного исследований являются новыми и достоверными.

5 Теоретическая и практическая значимость результатов работы

Теоретическая значимость исследования заключается в разработке и оптимизации математических моделей и алгоритмов, которые с использованием методов машинного обучения обеспечивают возможность потоковой классификации мобильных приложений, распространяющих противоправный, вредоносный и нежелательный контент, и эффективно функционируют как в условиях априорной неопределенности относительно состава и количества приложений, так и в случае смены концепта.

Практическая значимость представленных в работе результатов заключается в разработке алгоритмов и создании программного комплекса для классификации мобильных приложений на основе анализа сетевого трафика в потоковом режиме, с учётом возможной смены концепта для предотвращения распространения противоправного, вредоносного и нежелательного контента.

Сформированная в рамках работы экспериментальная база данных сетевого трафика мобильных приложений имеет потенциал для применения в системах обнаружения вторжений, а также для блокировки приложений, осуществляющих распространение вредоносного контента, в том числе тех, которые используют шифрование при передаче.

6 Апробация работы и публикации по диссертации

Основные результаты диссертационных исследований обсуждены и одобрены на 7 конференциях и опубликованы в 18 научных печатных работах, в том числе: 5 – в научных журналах перечня ВАК; 1 – в научных рецензируемых изданиях по базе Scopus; 11 – в материалах конференций и других изданиях. Получено свидетельство о Государственной регистрации программы для ЭВМ. Среди опубликованных работ в научных журналах перечня ВАК две написаны лично диссертантом.

Результаты представляюсь на 4 международных и 1 всероссийской научно-технических конференциях.

7 Соответствие работы паспорту научной специальности

Диссертация соответствует двум пунктам паспорта специальности 2.3.6 «Методы и системы защиты информации, информационная безопасность (технические науки)»:

п. 13. «Методы и модели выявления и противодействия распространению ложной и вредоносной информации» (1-3 научные результаты);

п. 15. «Принципы и решения (технические, математические, организационные и др.) по созданию новых и совершенствованию существующих средств защиты информации и обеспечения информационной безопасности» (4-5 научные результаты).

8 Замечания по диссертации

1. Из работы не понятно, чем отличаются 4 и 5 результаты, представляемые на защиту. Фактически предложенный алгоритм обнаружения смены концепта и классификации мобильных приложений и предложенный алгоритм MARF, обеспечивающий обнаружение смены концепта очень похожи и основываются на применении скользящего окна с учетом старения данных, отличия, заявленные в виде учета предложенной автором статистической модели обнаружения смены концепта, что может являться частным случаем оценки старения данных.

2. Публикации автора по теме диссертации трудно поставить в соответствие конкретным научным результатам, они не равномерно распределены по научным результатам и по годам.

3. Во второй главе методика отбора рационального числа атрибутов на основе анализа их информативности при фиксировании допустимой вероятности ложной классификации, четка не описана, приводится только констатация 5 шагов отбора, выводы по главе не отражают отличие данной методики от известных.

4. Модифицированный алгоритм классификации мобильных приложений в условиях неконтролируемого фонового трафика фактически представлен в тексте диссертации не алгоритмом, заявленным как научный результат, а как модель бинарной классификации, выводы по эффективности сформулированы только в советующем параграфе, а не в выводах по главе.

5. Четвертая глава посвящена разработке программного комплекса системы анализа трафика, реализующего представленные выше научные результаты. Данное программное обеспечение является безусловно важным практическим, но не научным результатом, раскрывающим, например, новые принципы построения подобных программ.

6. Имеют место стилистические ошибки в тексте.

Вместе с тем отмеченные недостатки не носят принципиального характера и не снижают ценности представленной диссертационной работы, которая, несомненно, имеет большую практическую ценность. Важным достоинством работы является то, что научные результаты, предложенные автором, доведены до готовой реализации в виде программного обеспечения, используемого в одном из лидеров Российского рынка решений в данной области – компании АО «Лаборатория Касперского».

Диссертация выполнена в рамках работы по гранту аспирантам, соискателям и молодым ученым на исследования, направленные на обеспечение информационной безопасности для задач цифровой экономики при государственной поддержке ведущих научных школ Российской Федерации в области информационной безопасности (Грант ИБ).

9 Заключение о соответствии диссертации критериям, установленным Положением о присуждении учёных степеней

Диссертационная работа Баркова Вячеслава Валерьевича «Классификация противоправных и нежелательных мобильных приложений методами машинного обучения в потоковом режиме» по актуальности, научной новизне, теоретической и практической значимости соответствует требованиям п. 9-14 «Положения о присуждении ученых степеней» ВАК Минобрнауки России, предъявляемым к диссертациям на соискание ученой степени кандидата технических наук, так как является научно-квалификационной работой, в которой изложены научно обоснованные технические и программные решения в области классификации мобильных приложений, осуществляющих распространение противоправного, нежелательного и вредоносного контента, имеющие существенное значение для

развития систем обеспечения информационной безопасности страны, использующих методы интеллектуального анализа данных.

Тема и содержание диссертации «Классификация противоправных и нежелательных мобильных приложений методами машинного обучения в потоковом режиме» полностью соответствует выбранной специальности 2.3.6. «Методы и системы защиты информации, информационная безопасность (технические науки)».

Барков Вячеслав Валерьевич заслуживает присуждения ученой степени кандидата технических наук по специальности 2.3.6. Методы и системы защиты информации, информационная безопасность (технические науки).

Официальный оппонент:
кандидат технических наук (специальность 05.13.01. Системный анализ, управление и обработка информации (технические науки)), доцент,
Заведующий кафедрой «Защищённые системы связи»,
Федеральное государственное бюджетное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича»
193232, г. Санкт-Петербург,
пр. Большевиков, д.22, корп.1, литера А, Ж
Тел.: +7 (812) 305-12-55, доб. 1255
E-mail: krasov.av@sut.ru

12 ноября 2024 г. Красов Андрей Владимирович