ФОРМА 514. ОТЧЕТ О ВЫПОЛНЕНИИ НАУЧНОГО ПРОЕКТА

14.1. Номер проекта
14-37-50735

14.2 Руководитель проекта
Чечулин Андрей Алексеевич

14.2. Исполнитель проекта (Молодой ученый)
Носков Антон Николаевич

14.3. Название проекта
Исследование и разработка эвристических подходов к обнаружению атак на телекоммуникационные сети на базе методов интеллектуального анализа данных

14.4. Вид конкурса
мол_нр Конкурс научных проектов, выполняемых молодыми учеными под руководством кандидатов и докторов наук в научных организациях Российской Федерации

14.5. Год представления отчета
2014

14.6. Сведения о выполнении научного проекта
Основные цели проекта на 2014 год сводились к следующему:
1 Разработка метода классификации основных видов сетевых угроз, анализ существующих исследований в области. Анализ и оценка основных достоинств и недостатков существующих систем обнаружения вторжений (СОВ).
2 Исследование и определение существующих походов к выделению наиболее важных признаков нежелательного трафика или возможной атаки, методики минимизации пространства, основные принципы формирования обучающих выборок, выделение признаков из входящего трафика, формирование обучающей выборки для алгоритма SVM, оптимизация основных параметров SVM в заданных пределах, обучение алгоритма.
3 Построение модели представления входных данных в виде набора векторов, включающих в себя различные параметры характеризующие сетевые пакеты на сетевом и транспортном уровнях модели OSI;
4 Создание эвристического подхода к обнаружению атак на базе метода опорных векторов (SVM), позволяющего классифицировать входящий трафик на «подозрительный» и «чистый».

Степень достижения поставленных в проекте целей
Все задачи запланированные в проекте на 2014 год, выполнены полностью. Проведено исследование эвристических подходов к обнаружению атак на телекоммуникационные сети на базе методов интеллектуального анализа данных. Рассмотрены методы анализа и оценки систем обнаружения вторжения.

Полученные в 2014 года важнейшие результаты

1. Классификация основных видов сетевых угроз. Анализ и оценка основных достоинств и недостатков методов интеллектуального анализа данных для обнаружения атак, применяемых в системах обнаружения вторжений.
В настоящее время существует множество различных видов сетевых атак, которые используют как уязвимости операционной системы, так и иного

установленного программного обеспечения системного и прикладного характера. Для того чтобы своевременно обеспечить безопасность компьютера, важно знать какого рода сетевые атаки могут угрожать ему.
Известные сетевые угрозы можно условно разделить на три большие группы:
Сканирование портов – этот вид угроз сам по себе не является атакой, а обычно предшествует ей, поскольку является одним из основных способов получить сведения об удаленном компьютере. DoS-атаки или атаки, вызывающие отказ в обслуживании – это атаки, результатом которых является приведение атакуемой системы в нестабильное, либо полностью нерабочее состояние.
Существует два основных типа DoS атак:
- отправка компьютеру-жертве специально сформированных пакетов, не ожидаемых этим компьютером, что приводит к перезагрузке или остановке системы;
- отправка компьютеру-жертве большого количества пакетов в единицу времени, которые этот компьютер не в состоянии обработать, что приводит к исчерпанию ресурсов системы.
Яркими примерами данных групп атак являются следующие атаки:
- Атака Ping of death (группа 1);
- Атака Land (группа 1);
- Атака ICMP Flood (группа 2);
- Атака SYN Flood (группа 2).
Атаки, целью которых является «захват» системы. Данная группа является также самой большой по количеству включенных в нее атак. Их можно разделить на три подгруппы в зависимости от операционной системы: атаки на Microsoft Windows-системы, атаки на Unix-системы, а также общая группа для сетевых сервисов, использующихся в обеих операционных системах.
Наиболее распространенными видами атак, использующих сетевые сервисы операционной системы, являются:
- атаки на переполнение буфера;
- атаки, основанные на ошибках форматных строк.
Были проанализированы следующие методы интеллектуального анализа данных, подходящие для решения задачи обнаружения сетевых атак в эвристических системах обнаружения вторжений:
а) наивный байесовский классификатор (Naive Bayes Classifier)
б) метод K-ближайших соседей
в) нейронные сети
г) метод опорных векторов (SVM)
Выделены следующие параметры эффективности работы эвристических СОВ:
- CR (cost ratio, соотношение затрат) – количество корректно распознанных атакующих и легитимных пакетов, так как любые атаки, обычно, не являются нормальным трафиком сети – и классифицируютя как атакующий трафик;
- FP (False Positive, ложная тревога) – количество легитимных пакетов принятых за атакующие;
- FN (False Negative) – количество атакующих пакетов, принятых за нормальные;
- PPs (Packet per second, пакетооборот) – максимальное количество пакетов, которое система может обработать за 1 секунду.
В условиях реальных современных сетей на применение СОВ накладываются особые требования, связанные с высокими уровнями трафика (большими и сверхбольшими показателями пакетооборота в сети). При проверке большого количества пакетов в секунду, любое ложное срабатывание создает новое сообщение в журналах аномалий. Это приводит к тому, что если количество

ложных срабатываний системы достаточно велико, то журналы системы очень быстро заполнятся ошибками распознавания и их анализ будет затруднен.

2.Проведены исследования существующих походов к выделению наиболее важных признаков нежелательного трафика или возможной атаки. Рассмотрены методики минимизации пространства, основные принципы формирования обучающих выборок, выделение признаков из входящего трафика, формирование обучающей выборки для алгоритма SVM, оптимизация основных параметров SVM в заданных пределах, обучение алгоритма.
В результате классификации вторжений, а также исследования современных атак на транспортном и/или межсетевом уровне сети, были выбраны следующие признаки сетевого трафика для дальнейшего анализа:
- Общие: протокол, характеристики фрагментации, TTL, ToS, количество отправленных/полученных байт, является ли широковещательным, IP-опции, корректность CRC и др.
- ICMP-пакетов: код и тип ICMP-сообщения.
- UDP-пакетов и TCP-сессий: сервис, land (равен ли порт клиента порту сервера).
- TCP-сессий: продолжительность, количество флагов в сессии, менялся ли размер окна, встречался ли нулевой sequence, количество пакетов, количество сессий у того же сервиса, отношения отправленных/полученных к общему количеству байт сессии, количество байт под опции TCP, тип и версия ОС инициатора сессии, MSS и др.
Для построения системы обнаружения атак был выбран подход, основанный на одноклассовом классификаторе. Задача классификации с обучением на одном классе (one-class classification, одноклассовой классификации) формулируется следующим образом: на основе обучающей выборки (записи трафика) алгоритм должен разделить пакеты на атакующие и легитимные.
Таким образом, в данном исследовании рассматривалась задача обучения по прецедентам, где X – пространство объектов, Y – множество. Требуется построить алгоритм, аппроксимирующий целевую зависимость на всём пространстве X.На практике для построения SVM решают именно эту задачу, так как гарантировать линейную разделимость выборки в общем случае не представляется возможным. Этот вариант алгоритма называют SVM с мягким зазором (soft-margin SVM), тогда как в линейно разделимом случае говорят об SVM с жёстким зазором (hard-margin SVM). В разработанной в рамках данного исследования методике используется обучение по тестовой выборке. В качестве тестовой базы была выбрана база KDD Cup'99. Каждая запись в данной базе представляет собой образ сетевого соединения. Соединение – последовательность TCP пакетов за некоторое конечное время, моменты начала и завершения которого четко определены, в течение которого данные передаются от IP-адреса источника на IP-адрес приемника, используя некоторый определенный протокол. Таким образом, обучение алгоритма производится на наборе данных объёмом около 1 000 000 tcp сессий, icmp и udp пакетов, характерных для некоторой сети.
Для решения проблемы линейной неразделимости, необходимо перейти от исходного пространства признаковых описаний объектов X к новому пространству H с помощью некоторого преобразования f: X > H. Если пространство H имеет достаточно высокую размерность, то можно надеяться, что в нём выборка окажется линейно разделимой (легко показать, что если выборка Xl не противоречива, то всегда найдётся пространство размерности не

более l, в котором она будет линейно разделима). Пространство H называют спрямляющим. Для такого подхода недавно был придуман термин «беспризнаковое распознавание» (featureless recognition), хотя многие давно известные метрические алгоритмы классификации (kNN, RBF и др.) также не требуют задания признаковых описаний. В разработанной в рамках данного исследования методике это позволит обеспечить более высокую скорость работы и возможность распараллеливания. Кроме перехода в пространство повышенной размерности, существуют и другие варианты решения проблемы нелинейности. Так, например, примером решения этой задачи является метод активных ограничений. Для обучения SVM применяются алгоритмы, учитывающие специфические особенности SVM. Специфика заключается в том, что число опорных векторов, как правило, невелико, и эти векторы находятся поблизости от границы классов. Именно эти особенности и позволяют ускорить поиск опорных объектов.

Специализированные алгоритмы настройки SVM успешно справляются с выборками из десятков тысяч объектов. Был рассмотрен алгоритм – последовательный метод активных ограничений (incremental active set method, INCAS).

Описанный процесс является частным случаем метода активных ограничений (active sets method), который применяется для решения произвольных задач математического программирования с ограничениями-неравенствами. В линейном программировании он эквивалентен симплекс-методу. Сходимость данного метода в общем случае не гарантируется, однако в задачах настройки SVM зацикливания случаются крайне редко. Хорошие эвристические показатели данного метода заключаются в том, что на каждом шаге выбирается объект i, для которого условие Куна-Таккера нарушается сильнее всего. Это способствует увеличению скорости сходимости.

Обнаружение сетевых атак связано с выделением большого числа признаков, по которым можно проводить классификацию. Так, например, в общедоступной базе KDD Cup'99, содержащей порядка 5 миллионов экземпляров атак, классифицированных по 22 типам, используется 41 признак. Все признаки информационно неравнозначны, причем уточнить их истинную значимость можно только после проведения дополнительных исследований. Задача оптимизации числа признаков является неотъемлемой частью процесса распознавания. В реальных условиях система должна автоматически выделять признаки и использовать их для решения задачи обнаружения атаки и определения ее типа. В настоящее время разрабатывается большое количество различных технологий обнаружения атак, основанных на различных методах интеллектуального анализа данных. К их недостаткам можно отнести уязвимость к новым атакам, низкую точность и скорость работы. В настоящей работе основной упор делается на сжатие пространства признаков для последующего использования в нейросетевой системе обнаружения атак.

Метод главных компонент (МГК) – один из основных способов уменьшить размерность данных, потеряв наименьшее количество информации. Вычисление главных компонент сводится к вычислению собственных векторов и собственных значений ковариационной матрицы исходных данных. Имея набор главных компонент, можно «свернуть» пространство любого вектора, полученного на выходе алгоритма преобразования, как на этапе обучения, так и при тестировании.

3.Построена модель представления входных данных в виде набора векторов,

включающих в себя различные параметры, характеризующие сетевые пакеты на сетевом и транспортном уровнях модели OSI.
Алгоритм СОВ включает в себя блок предобработки входящего трафика. В этом блоке из заголовков второго и третьего уровня модели OSI формируются входные вектора, включающие набор пакетных и временных признаков. Тем не менее, для корректной оценки работы алгоритма требуется объемная выборка экспертно проанализированных пакетов. Всего был выделен 41 признак, на основе которых были сформированы наборы векторов для обучения и тестирования эвристической системы обнаружения атак.
С помощью tcpdump были получены следующие признаки характеризующие сетевой трафик.
- duration: continuous
- protocol_type: symbolic
- service: symbolic
- flag: symbolic
- src_bytes: continuous
- dst_bytes: continuous
- land: symbolic
- wrong_fragment: continuous.
- urgent: continuous.
- hot: continuous.
- num_failed_logins: continuous.
- logged_in: symbolic.
- num_compromised: continuous.
- root_shell: continuous.
- su_attempted: continuous
- num_root: continuous.
- num_file_creations: continuous.
- num_shells: continuous.
- num_access_files: continuous.
- num_outbound_cmds: continuous.
- is_host_login: symbolic.
- is_guest_login: symbolic.
- count: continuous.
- srv_count: continuous.
- serror_rate: continuous.
- srv_serror_rate: continuous.
- rerror_rate: continuous.
- srv_rerror_rate: continuous.
- same_srv_rate: continuous.
- diff_srv_rate: continuous.
- srv_diff_host_rate: continuous.
- dst_host_count: continuous.
- dst_host_srv_count: continuous.
- dst_host_same_srv_rate: continuous.
- dst_host_diff_srv_rate: continuous.
- dst_host_same_src_port_rate: continuous.
- dst_host_srv_diff_host_rate: continuous.
- dst_host_serror_rate: continuous.
- dst_host_srv_serror_rate: continuous.
- dst_host_rerror_rate: continuous.

- dst_host_srv_rerror_rate: continuous

В качестве тестовой базы была выбрана база KDD Cup'99. Каждая запись в данной базе представляет собой образ сетевого соединения. Соединение – последовательность TCP пакетов за некоторое конечное время, моменты начала и завершения которого четко определены, в течение которого данные передаются от IP-адреса источника на IP-адрес приемника, используя некоторый определенный протокол. Отдельная запись KDD Cup'99 включает 41 информационный признак и промаркирована как «атака» или «не атака». Например, первый параметр определяет длительность соединения, второй – указывает используемый протокол, третий – целевую службу и т.д.

Атаки, представленные в тестовом наборе: back dos; buffer_overflow u2r; ftp_write r2l; guess_passwd r2l; imap r2l; ipsweep probe; land dos; loadmodule u2r; multihop r2l; neptune dos; nmap probe; perl u2r; phf r2l; pod dos; portsweep probe; rootkit u2r; satan probe; smurf dos; spy r2l; teardrop dos; warezclient r2l; warezmaster r2l.

При этом атаки делятся на четыре основные категории: DoS, U2R, R2L и Probe:

- DoS – отказ в обслуживании, характеризуется генерацией большого объема трафика, что приводит к перегрузке и блокированию сервера;
- U2R предполагает получение зарегистрированным пользователем привилегий локального суперпользователя (администратора);
- R2L характеризуется получением доступа незарегистрированного пользователя к компьютеру со стороны удаленной машины;
- Probe заключается в сканировании портов с целью получения конфиденциальной информации.

4. Разработан эвристический подход к обнаружению атак на базе метода опорных векторов (SVM), позволяющий классифицировать входящий трафик на «атакующий» и «легитимный».

Для реализации и тестирования алгоритма было необходимо:

- обеспечить алгоритм преобразования входящего трафика в векторы признаков вторжения;
- реализовать алгоритм снижения размерности пространства признаков методом главных компонент;
- определить пределы допустимых значений для управляющего параметра C и выбрать сам параметр;
- найти параметры SVM при помощи алгоритма INCAS и обучить SVM на обучающей выборке;
- протестировать работу алгоритма и проанализировать результаты.

Принимая во внимание критерий быстродействия, необходимый для СОВ, был разработан алгоритм, в котором размерность входных векторов снижается с помощью МГК, а настройка параметров SVM выполняется при помощи INCAS.

Выбор обусловлен следующими критериями:

По сравнению с независимым анализом компонент (ICA) или нелинейными методами сокращения размерности пространства, МГК обеспечивает более высокую скорость работы и возможность распараллеливания (например, с применением GPGPU)

Задача построения СОВ соответствует условиям эффективной работы последовательного метода активных ограничений, в частности:

а) сравнительно малому числу опорных векторов;

б) последовательному поступлению входных векторов в блок анализатора.

Для формирования исходных данных был проведён анализ информации, полученной при обработке реальных IP-трафиков, информация о которых

представлена в общедоступной базе образцов сетевого трафика KDD Cup'99. База содержит характеристики легитимных соединений и атак (отказ в обслуживании, сканирование портов, получение зарегистрированным пользователем привилегий локального суперпользователя, получение доступа незарегистрированного пользователя к компьютеру со стороны удаленной машины).

Таким образом, обучение алгоритма производилось на наборе данных большого объёма - около 1 000 000 tcp сессий, icmp и udp пакетов, характерных для некоторой сети.

Для снижение размерности пространства признаков был использован МГК. Данный метод представляет собой итерационную процедуру, в которой новые компоненты добавляются последовательно, одна за другой. Важно знать, когда остановить этот процесс, т.е. как определить правильное число главных компонент. Если это число слишком мало, то описание данных будет не полным. С другой стороны, избыточное число главных компонент приводит к переоценке, т.е. к ситуации, когда моделируется шум, а не содержательная информация. Результаты исследования показали, что сокращение размерности признаков при помощи МГК не оказывает значительного влияния на уровень верного распознавания легитимных и атакующих пакетов. В итоговой реализации размерность с 41 признака была сокращена до 27 (на 34%).

Для оценки параметров использовался метод скользящего контроля. Скользящий контроль, или кросс-проверка, или кросс-валидация (cross-validation, CV) – процедура эмпирической оценки обобщающей способности алгоритмов, обучаемых по прецедентам. Фиксируется некоторое множество разбиений исходной выборки на две подвыборки: обучающую и контрольную. Для каждого разбиения выполняется настройка алгоритма по обучающей подвыборке, затем оценивается его средняя ошибка на объектах контрольной подвыборки. Оценкой скользящего контроля называется средняя по всем разбиениям величина ошибки на контрольных подвыборках.

Если выборка независима, то средняя ошибка скользящего контроля даёт несмещённую оценку вероятности ошибки. Это выгодно отличает её от средней ошибки на обучающей выборке, которая может оказаться смещённой (оптимистически заниженной) оценкой вероятности ошибки, что связано с явлением переобучения. Скользящий контроль является стандартной методикой тестирования и сравнения алгоритмов классификации, регрессии и прогнозирования.

От глобального параметра C в данном случае зависит, будет ли акцент на максимизации зазора между классами или на минимизации функции ошибки. Если C слишком большой, то в алгоритме будет слишком большой штраф для неразделимых точек; если C недостаточен, то может возрасти число ошибок на этапе обучения. При C, стремящемся к бесконечности, получим SVM с жёстким зазором. Можно определить граничные значения параметра C между 0.01 и 35 000. Задача метода оптимизации – нахождение экстремума функции f, характеризующей процент корректно классифицированных алгоритмом пакетов. При задании C в определенных пределах и использовании метода оптимизации остальных параметров SVM колебания в выборе C не влияют на конечную эффективность работы алгоритма. Для реализации данного алгоритма было выбрано значение: C = 225.

Один из ключевых недостатков алгоритма SVM – высокая чувствительность к экспертной ошибке в обучающем наборе. Этот недостаток может оказаться критичным при работе с реальными данными. Из эксперимента можно сделать

вывод, что при уровне экспертной ошибки около 0.5% алгоритм функционирует в нормальном режиме; эти значения достижимы при обучении на реальном трафике.
Для исследования эффективности и скорости работы алгоритма SVM с различными параметрами используем метод скользящего контроля и контрольные выборки по 5000 тренировочных и 5000 тестовых пакетов.

Авторы проекта опубликовали полученные результаты в научном журнале, относящемся к перечню ВАК, (Носков А.Н., Чечулин А.А., Тарасова Д.А. Исследование эвристических подходов к обнаружению атак на телекоммуникационные сети на базе методов интеллектуального анализа данных // Труды СПИИРАН. Вып.6 (37). СПб.: Наука, 2014.) сборниках и трудах конференций, а также апробировали результаты на ряде различных российских и международных конференций, в частности на XIV Санкт-Петербургской Международной Конференции «Региональная информатика-2014» («РИ-2014»), г. Санкт-Петербург, 29-31 октября 2014 г. (Левшун Д.С., Чечулин А.А. Построение классификационной схемы существующих методов корреляции событий безопасности // XIV Санкт-Петербургская Международная Конференция “Региональная информатика-2014” (“РИ-2014”). 29-31 октября 2014 г. Материалы конференции. СПб., 2014. С. 148-149); Научно-технической конференции «Инновации Северо-Запада», г. Санкт-Петербург, 15-16 декабря 2014 г. (Чечулин А.А., Котенко И.В., Дойникова Е.В. Методика анализа истории событий безопасности, прогнозирования действий нарушителя и их последствий // Научно-техническая конференция «Инновации Северо-Запада». Материалы конференции. 15-16 декабря 2014 г. СПб.: Изд-во СПбГЭТУ «ЛЭТИ». 2004. С.69-72).

14.7.1. Полное название организации, где выполнялся научный проект
Федеральное государственное бюджетное учреждение науки Санкт-Петербургский институт информатики и автоматизации Российской академии наук

14.7.2. Сокращенное название организации, где выполнялся научный проект
СПИИРАН

14.8. Количество подготовленных публикаций по результатам выполнения научного проекта
3

*Подпись Руководителя проекта*_____