4. T. R. Gruber. Towards principles for the design of ontologies used for knowledge sharing. // International Journal of Human-Computer Studies. 1995. Vol. 43 (5/6). P. 907-928.

УДК 519.876.2

ГРНТИ 81.93.29

МОДЕЛЬ РАСПОЗНАВАНИЯ АНОМАЛЬНОГО ПОВЕДЕНИЯ ПРИ VPN ПОДКЛЮЧЕНИЯХ К ИНФРАСТРУКТУРЕ

Тынымбаев Болат Айткожинович
tynymbaevba@gmail.com

Докторант 2 курса, ЕНУ им. Л.Н. Гумилева, Нур-Султан, Казахстан

Научный руководитель – А.А. Адамов

Введение. В условиях введения чрезвычайного положения, в марте-апреле 2020 года многими организациями на территории Республики Казахстан было принято решение перевести своих работников на удалённый режим работы. При этом наиболее удобным и обеспечивающим безопасное соединение является VPN подключение, при условии выполнения отдельных правил и требований информационной безопасности. Однако, модель угроз организации увеличивается при данном виде подключении работников к критичным информационным активам. В данной работе будет описан ландшафт угроз, а также представлена модель для защиты инфраструктуры организации с применением машинного обучения при анализе VPN подключений работников, а также классификации пользователей-работников на отдельные виды группы

Модели угроз удалённого подключения. Угрозы, которые должна рассматривать организация, при предоставлении удалённого доступа своим работникам можно разделить согласно двум видам активов:

- рабочие станции работников, подключающихся удалённо;

- удалённое сетевое соединение.

Подобные модели угроз описаны в следующих статьях [1-2]. Угрозы, объединенные в группы по схожести типов атак, представлены в таблице 1.

Таблица 1. Модель угроз при удалённом подключении

Угроза	Описание
Угроза взлома, присутствия	Поскольку на личные рабочие станции работников

зловредного ПО на рабочей станции работника	не действуют групповые политики организаций, пользователи имеют возможность устанавливать самостоятельно различное программное обеспечение, с риском установки нелегитимного ПО, а также пренебрегать «гигиеной» кибербезопасности: открывая зловредные ссылки, фишинговые письма.
Атака «человек-по-середине», перехват удалённого соединения, взлом удалённого соединения	Существует риск перехвата соединения при использовании слабых алгоритмов шифрования при организации удалённого подключения. Также существует риск взлома модема, роутера работников при организации выхода в сеть Интернет.
Незапланированное предоставление удалённого доступа к критичным информационным активам	Предоставление доступа по открытому, незащищенному протоколу, к примеру, доступ к критичному серверу с ОС Windows Server по протоколу RDP. Существует риск некорректной сегментации сетей, при котором удалённый доступ будет получен к критичным информационным активам, к которым не должен предоставляться доступ. Может быть нарушена процедура предоставления доступов.
Взлом учётной записи работника, подключающегося удалённо	Угроза заключается в слабом пароле пользователя, ошибках в процедуре сброса пароля.

При этом в организациях следует пересмотреть регистр рисков, к примеру, для угрозы «Взлом учётной записи», её вероятность резко повышается, поскольку сейчас удалённый доступ предоставляется работникам, плохо обученным ИТ основам и с низким уровнем осведомлённости об основах информационной безопасности.

Анализ исследований в данном направлении. Тема безопасности VPN подключений не нова, к примеру, в работе [3] описана модель и подход для выявления аномальных подключений, представлен следующий алгоритм:

1. Определить количество кластеров (разделены группы пользователей с подключениями по будням и выходным).
2. Разделить тестовый набор данных на 10 различных наборов.
3. Выполнить расчёт максимального ожидания кластеров 10 раз с использованием 10 наборов.
4. Оценить вероятности показателей эффективности для каждой модели.

5. Если показатель увеличился после предыдущей итерации, увеличить количество кластеров и повторить. Если не увеличился, вернуть предыдущее значение итерации как оптимальное количество кластеров.

В исследовании [4] описан алгоритм кластеризации fpMAFIA и его применение для определения сетевых вторжений. Данный алгоритм схож с CLIQUE, использующий теорию графов.

Одним из вариантов управления описанными угрозами - является использование услуг аутентификации на третьей стороне, к примеру, на подобие схемы, описанной в работе [5]. Такие сервисы возможны у провайдеров MSSP, по организации VPN подключении с модулем аналитики поведения пользователей. К примеру, в работе [6] описано применение технологии CASB (cloud access security broker) с модулем аналитики поведения пользователей. Сведения об алгоритмах и сценариях выявления аномалий в системах и модулях аналитики поведения пользователей представлены в исследовании [7].

Модель распознавания аномального поведения. Для решения задачи распознавания аномального поведения пользователей с целью распознавания угроз в определенном объеме[1], описанных выше, предлагается следующая архитектура, представленная на рисунке 1.

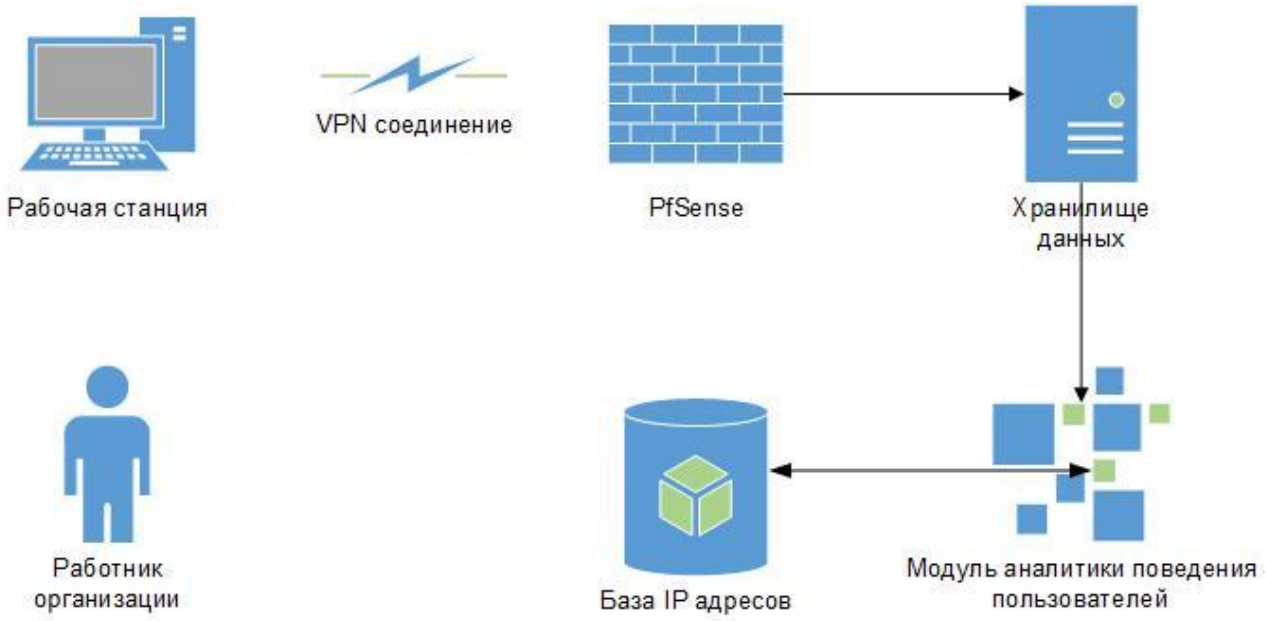

Рисунок 1. Архитектура системы распознавания аномального поведения при VPN подключениях.

Для исследования используем логи на протяжении 3 месяцев подключения 50 пользователей. В качестве сетевого оборудования используется программное обеспечение PfSense, являющимся VPN сервером, в качестве хранилища данных использовались несколько вариантов: Hadoop, а также программное обеспечение Microsoft Power BI Desktop. Стандартный лог о VPN подключении пользователя Pfsense, содержит информацию со следующими полями:

[1]Распознавание угроз не будет реализовано в полном объёме, поскольку для полноты исполнения процессов обеспечения информационной безопасности нужны дополнительные информационные модули, а также процессы, мероприятия, что выходит за рамки данного исследования.

дата, время, ip адрес пользователя, ip адрес конечной точки подключения, сетевой порт, идентификатор пользователя, описание лога, тип события, наименование VPN туннеля. Однако не все поля используются в качестве набора данных. К примеру, поля название vpn тунеля, levl, тип события, log description не информативны для выявления аномалий, но могут понадобиться в будущем при расследовании инцидентов. Также используется база данных ip адресов, geolite2 (модуль для языка программирования Python) для последующего «обогащения» информации об ip адресах: страна, широта и долгота.

Методы машинного обучения использовались такие как статистические метрики, к примеру, медианное отклонение для вычисления расстояний локации пользователя и офиса; критерий Граббса, однако данный метод не совсем эффективный, так как применим только для данных с нормальным распределением. Метод может быть не эффективным, так как находит одну аномалию в каждой итерации [8]. Также используется модель локального уровня выброса, с помощью программного модуля sklearn.neighbors.LocalOutlierFactor языка Python.

Для указанного набора данных стоит отметить следующие особенности: практически у всех пользователей динамические ip адреса, тем самым невозможна строгая привязка пользователя к ip адресу, однако это применимо к mac адресу рабочих станций пользователей. Используя широту и долготу, получаемые из базы данных ip адресов, можно отслеживать удаленность пользователя от офиса. Данные о стране позволяют, в первую очередь, настраивать правило запрета подключения из других стран. Аномалии, которые определяем после обучения модели, при удалённом подключении пользователей, представлены в таблице 2.

Таблица 2. Список типов аномальных поведений

№	Тип аномального поведения	Описание
1	Нестандартное время подключения пользователя по VPN	В рамках обучения модели выявляются стандартные временные промежутки подключения пользователей, после чего определяется нестандартное поведение.
2	Подозрительный IP адрес подключения	Аналогично пополняется база ip адресов пользователей, при этом при использовании данных сторонней БД определяется геолокация пользователя и удалённость пользователя от офиса. К примеру, подключение с территории другого государства сразу отключается согласно настройкам сетевого устройства.
3	Нелегитимный MAC адрес рабочей станции пользователя	В отличие от ip адреса, данное поле должно быть неизменным, что позволяет настраивать процедуру предоставления доступа со строгим набором правил. При

		этом реестр mac адресов сформируется после обучения.
4	Измененная геолокация пользователя, доступ с нескольких локаций	Согласно имеющимся данным о широте и долготе подключения, можно сделать вывод о расстоянии между офисом и локацией пользователя для отслеживания расстояния.

Заключение. С учётом дальнейшего улучшения модели, рассматриваются случаи false positive данной модели. К примеру, неоднократные случаи с чем столкнулись, ip адреса отдельных пользователей, использующих услуги мобильного оператора, распознавались в базе данных, как ip адреса Эстонии, в связи с чем правило о запрете подключения с территории другого государства необходимо было отключать для выяснения обстоятельств.

В будущих работах планируется обучить модель на большем количестве пользователей для облачного сервис-провайдера с соответствующим разделением групп пользователей по уровню осведомлённости и навыкам соблюдения требований кибербезопасности в противовес угрозам из таблицы 1. Возможно разделение на группы в соответствии с работой [9], подобное разбиение позволяет создавать более точечные курсы повышения осведомлённости об угрозах кибербезопасности с последующей лучшей отдачей и гибче настраивать процедуру предоставления доступов. Также данную модель можно использовать при проведении аудитов информационной безопасности и расследовании инцидентов, немаловажным плюсом является, что используемые компоненты не требуют затрат.

Список использованных источников

1. Бизнес на расстоянии: как защитить инфраструктуру: 2020. URL: https://www.ptsecurity.com/ru-ru/research/knowledge-base/biznes-na-rasstoyanii-kak-zashchitit-infrastrukturu/

2. Пять уязвимостей, опасных для удалённой работы: 2020. URL:https://habr.com/ru/company/pt/blog/494472/

3. M. J. Chapple N. Chawla A. Striegel Authentication Anomaly Detection: A Case Study On A Virtual Private Network // Conference: Proceedings of the 3rd Annual ACM Workshop on Mining Network Data, MineNet 2007, San Diego, California, USA, June 12, 2007.

4. K. Leung, C. Leckie Unsupervised Anomaly Detection in Network Intrusion Detection Using Clusters // Proc. Twenty-Eighth Australasian Computer Science Conference. Newcastle, Australia. P. 333-342

5. В.А. Десницкий, И.В. Котенко Модель защиты программного обеспечения на основе механизма «удалённого доверия» // ИЗВ. Вузов. Приборостроение. 2008, Т 51. N 11, С. 27-30.

6. Котенко И.В., Тынымбаев Б.А Архитектура перспективной системы UEBA для провайдеров облачных услуг // Актуальные проблемы инфо-телекоммуникаций в науке и образовании. VIII Международная научно-техническая и научно-методическая конференция: сб. науч. ст. в 4-х т. СПб.: СПбГУТ, 2019. Т. 1. С. 581-585.

7. Котенко И.В., Тынымбаев Б.А. Обзор решений класса UEBA // Актуальные проблемы инфо-телекоммуникаций в науке и образовании. VIII Международная научно-техническая и научно-методическая конференция: сб. науч. ст. в 4-х т. СПб.: СПбГУТ, 2019. Т. 1. С. 586-590.

8. C. Chio, D. Freeman Machine learning & Security. 2019. 365 p.

9. B. Tynymbayev, A. Adamov Mathematical model for Cloud provider's user behavioral analytics systems // II International Scientific Conference 11 - 14 December, 2019, Borovets, Bulgaria, Year III ISSUE 1(3)/2019. P. 19.

УДК 517.984

ГРАФТАҒЫ КЕЙБІР ДИНАМИКАЛЫҚ КЕРІ ЕСЕПТЕРДІ САНДЫҚ ШЕШУ

Умаров Мерей
umarov_mo@icloud.com
Казақстан, Нур-Султан қ.,
Л.Н. Гумилев атындағы Еуразия Ұлттық университеті,
Математикалық және компьютерлік модельдеу
кафедрасының магистранты
Ғылыми жетекшісі – К. Сулейменов

Бастапқы шекаралық есеп және Гурсат мәселесі.

Бірайнымалы және уақытқа тәуелді және $\Omega(x,t)=\{(x,t)\in[0,+\infty)\times[0,+\infty)\}$ жиынында берілген толқындық теңдеу үшін бастапқы шекаралық есепті қарастырамыз

$$\begin{cases} u_{tt}(x,t)-u_{xx}(x,t)+q(x)u(x,t)=0,\ x>0,t>0, \\ u(x,0)=u_t(x,0)=0,\ u(0,t)=f(t) \end{cases} \tag{1}$$

Мұндағы $q\in L^1_{loc}(\mathbb{R}_+)$ және $f\in L^2_{loc}(\mathbb{R}_+)$ f функциясы, шекаралық бақылау деп аталады. (1) есептің $u^f(x,t)\in L^2_{loc}\big(\Omega(x,t)\big)$ шешімі интегралдық ядро $\omega(x,s)$ арқылы жазылуы мүмкін, бұл Гурсат проблемасының бірегей шешімі болып табылады

$$\begin{cases} \omega_{tt}(x,t)-\omega_{xx}(x,t)+q(x)\omega(x,t)=0,\ 0<x<t, \\ \omega(0,t)=0,\ \omega(x,x)=-\frac{1}{2}\int_0^x q(s)ds\,. \end{cases} \tag{2}$$

$q\in L^1_{loc}(\mathbb{R}_+)$, яғни $[0,\infty)$ аралығында кез келген $[a,b]$ үшін $\int_a^b|q(x)|dx<\infty$ және $f\in L^2_{loc}(\mathbb{R}_+)$, яғни $[0,\infty)$ аралығындағы кез келген $[a,b]$ үшін $\int_a^b|f(t)|^2dt<\infty$, $u(x,t)\subset L^2_{loc}(\mathbb{R}^2_+)$ кез келген $[a,b]*[\alpha,\beta]=\Omega\Rightarrow\iint_\Omega|u(x,t)|^2dxdt<\infty$.

Сонымен, (1) жүйесінен (2) жүйеге өтуді дәлелдеп көрсетейік. Ол үшін

$$u_{tt}(x,t)-u_{xx}(x,t)+q(x)u(x,t)=0$$

теңдеуін

$$\omega(x,t)\in C^{2,2}(\mathbb{R}^2_+)$$

функциясына көбейтіп,