Proceedings of IYSW, (2020), vol. **9**, pp 249-262.

Journal homepage: http://journals.sdu.edu.kz/index.php/iysw

Прогнозирование уровня информационной безопасности в предприятии

Динара Хашимова

Университет Сулеймана Демиреля

Аннотация

В данной работе оценен потенциал машинного обучения для обнаружения вторжений на уровне приложений, с использованием параметров *DDoS-атаки, Атаки межсайтового скриптинга (XSS) и SQL-инъекций*. Результат, отражающий этот вопрос, будет представлен в виде набора классификаторов, способных идентифицировать вредоносные полезные данные, и какие функции дадут хорошие показатели результата для принятия решения, является ли ввод вредоносным или безвредным.

Ключевые слова: DDoS-атаки, атаки межсайтового скриптинга (XSS) и SQL-инъекций, информационная безопасность.

Прогнозирование уровня информационной безопасности в предприятии

В результате растущего использования компаниями веб-приложений, компании настаивают на защите своих конфиденциальных данных (таких как пароль, номера банковских карт и т.д.), Используя брандмауэры и демилитаризованную зону в качестве первой линии защиты, что недостаточно для обеспечения достойного уровня безопасности, и не гарантирует высокий уровень безопасности. По этой причине наступает вторая линия защиты с использованием систем обнаружения вторжений (IDS) и сканеров уязвимостей веб-приложений для обеспечения защиты конфиденциальных данных от попыток атаки и эксплуатации данных незаконным способом. Системы обнаружения вторжений можно разделить на две основные категории:

- Системы обнаружения вторжений на основе сигнатур;
- Системы обнаружения вторжений на основе аномалий.

Однако они страдают от многих недостатков, которые делают традиционные системы обнаружения вторжений неадекватными для обнаружения новых атак или применимыми в веб-приложениях. Таким образом, в последнее время многие исследователи применяют методы и алгоритмы интеллектуального анализа данных, машинного обучения и искусственного интеллекта, чтобы использовать их преимущества для улучшения производительности систем обнаружения вторжений, которая измеряется скоростью обнаружения и частотой ложных тревог [2]. Угроза может быть вызвана внутренними, внешними или как внешними, так и внутренними объектами. Огромные объемы данных, хранящихся в базе данных, делают его критической точкой защиты для любого предприятия - и ценным объектом для электронных сетей. Внешние угрозы или злоумышленники действуют извне компании и должны преодолеть вашу внешнюю

защиту, чтобы получить доступ к вашей базе данных. Внешние угрозы ограничены тем, какой доступ они могут получить за пределами сети передачи данных вашей компании. Они должны успешно обойти или отключить внешнюю защиту, прежде чем смогут войти в сеть и получить доступ к тем данным, которые доступны для непривилегированных учетных записей. Внутренние угрозы или диверсанты работают внутри компании и, таким образом, могут обойти внешнюю защиту. Будучи доверенными членами компании, они уже имеют гораздо больший доступ, чем любая внешняя угроза. Намерения, стоящие за угрозой для базы данных, являются еще одним ключевым вопросом. Внешние угрозы почти всегда являются злонамеренными: кража данных, сбои в работе служб - все возможные цели. Внутренние угрозы могут быть одинаково злобными и могут также включать в себя шантаж или другие незаконные действия. Однако внутренние угрозы не всегда являются вредоносными. Иногда угроза - это вовсе не личность, а слабые внутренние политики и меры безопасности, которые вызывают неожиданные или непреднамеренные нарушения базы данных или открывают уязвимости для внешних злоумышленников, если они нарушают внешние меры безопасности или обходят их [3].

Параметры атак

В [1] были раскрыты и рассмотрены распространенные виды атак в веб-приложениях, такие как DDoS-атаки, атаки SQL-инъекций, атаки межсайтового скриптинга (XSS), атаки нулевого дня (Zero-day Attacks), атаки бизнес-логики (Business Logic Attacks), атаки «человек посередине» (Man-in-the-middle attacks), вредоносные программы (Malware), Пороки (Defacements).

Две основные известные уязвимости в веб-приложении являются SQL-инъекция и межсайтовый скриптинг (XSS), также являющая основной угрозой в интернете DDoS-

атаки, рассматриваем в нашей исследовательской работе и анализируем с использованием наборов данных этих уязвимостей. DDoS-атаки - типы внешних угроз, которые чаще всего случаются, направлены на подавление целевого веб-приложения веб-сайта/сервера ложным трафиком, снижая пропускную способность сети и делая ее недоступной для законных пользователей. Атаки межсайтового скриптинга (XSS), относящиеся также к внешним угрозам, направлены на пользователей уязвимых веб-приложений/веб-сайтов, чтобы получить доступ к браузерам и управлять ими. Атаки SQL-инъекций – типы внутренних угроз, в которых злоумышленник внедряет вредоносный код SQL в виде запросов или запросов в поля ввода пользователя в веб-приложениях, таких как формы отправки, контактные формы и т.д. Таким образом, они получают доступ к внутренней базе данных приложения, куда они проникают извлекать конфиденциальную информацию о клиентах или самой компании, получать несанкционированный административный доступ, изменять или удалять данные или даже получать полный контроль над веб-приложением.

Выявление параметров рассматриваемых атак

Параметры DDoS-атаки [4]:

1. Timestamp (метка времени) – время запроса. Информация о метке времени преобразуется в GMT для мобильности. Для расчета местного времени каждая метка времени должна быть скорректирована;
2. clientID (ID клиента) - для каждого клиента был задан уникальный целочисленный идентификатор, и из-за некоторых проблем конфиденциальности эти сопоставления не были выпущены;
3. objectID (идентификатор объекта) - уникальный целочисленный идентификатор для запрошенного URL;
4. size (размер) - количество байтов в ответе;
5. method (метод) - метод, содержащийся в запросе клиента;

6. status (статус) - это поле содержит две части информации; 2 биты высшего порядка содержат версию HTTP, указанную в запросе клиента, а остальные 6 битов указывают код состояния ответа;
7. type (тип) - тип запрашиваемого файла;
8. server (сервер) - дает информацию о том, какой сервер обработал запрос.

Параметры Атаки межсайтового скриптинга (XSS) и SQL-инъекций:

1. Length (длина) – первой пользовательской функцией, которая будет реализована, будет длина ввода;
2. Non-printable characters (непечатные символы) - непечатаемыми символами являются, например, символы табуляции, разрывы строк и нулевые символы, то есть символы, которые не представляют собой письменный символ;
3. Punctuation characters (знаки пунктуации) - эта функция описывает количество знаков препинания в полезной нагрузке. Сюда входят такие символы, как <и', которые обычно используются как в SQL-инъекциях, так и в межсайтовых скриптовых атаках;
4. Minimum byte (минимальный байт) - функция, описывающая минимальный байт во входных данных. Средний минимальный байт и стандартное отклонение;
5. Maximum byte (максимальный байт) - подобно минимальному байту, создает функции максимального байта;
6. Mean byte (средний байт) - функция, описывающая среднее значение байта значений входных символов;
7. Standard deviation byte (средний байт) - функция, описывающая среднее значение байта значений входных символов;
8. Distinct bytes (отдельные байты) - функция, описывающая количество различных байтов во входной строке;
9. SQL keywords - описание количества ключевых слов SQL внутри ввода;
10. Ключевые слова Javascript - функция, описывающая количество связанных с JavaScript слов для заданного ввода.

После создания нашего набора данных, его очистки и создания функций следующим шагом был выбор и обучение наших моделей. Выбранные классификаторы перечислены и раскрыты в следующем разделе.

Модели прогнозирования

1. *Логистическая регрессия (Logistic regression)*

Классификатор использует регрессию для соответствия границ между классами. Линия регрессии может быть произвольной функцией любого порядка, которая затем используется в качестве входных данных для сигмовидной функции. Функция сигмоида создает границу из линии регрессии, которая разделяет два класса.

2. *Машина опорных векторов (Support vector machine)*

Машина опорных векторов (SVM) является мощным классификатором, который признан хорошим выбором модели при подборе многомерных данных. Основная теория подгонки разделительной линии в пространстве признаков заключается в максимизации разрыва между ближайшими точками к линии каждого класса.

3. *Наивный байесовский классификатор (Multinomial Naive Bayes)*

Байесовский классификатор использует байесовское правило, чтобы назначить вероятность того, что точка данных находится в определенном классе. Затем можно определить порог того, насколько вероятным должен быть результат, чтобы классифицировать точку входных данных как положительный класс.

4. *Случайный лес (Random forest)*

Чтобы описать классификатор случайных лесов, сначала необходимо определить дерево решений. По сути, дерево решений - это серия вопросов «да / нет», задаваемых об образце данных. Узлы представляют вопросы, а края представляют ответы. Случайный

лес говорит сам за себя; он состоит из нескольких случайно сгенерированных деревьев решений. Модой всех деревьев решений будет вывод классификации из случайного леса.

5. *AdaBoost*

Как и случайные леса, классификатор AdaBoost объединяет несколько классификаторов. AdaBoost лучше всего использовать при объединении слабых классификаторов, таких как деревья решений только высоты 1. Алгоритм начинается с обучения модели на основе обучающих данных, а затем создания второй модели с целью исправления ошибок из первой модели. Эта аддитивная процедура выполняется до тех пор, пока не будет создано максимальное количество моделей или пока не будет достигнут идеальный прогноз.

6. *Искусственная нейронная сеть* (*Artificial neural network)*

В последние годы популярность искусственных нейронных сетей (ANN) возросла из-за увеличения объема доступных данных и вычислительной мощности, необходимой для обучения этих сетей. Аргумент, почему этот метод более мощный, чем любой другой классификатор машинного обучения, заключается в том, что он может представлять любую функцию, например иметь структуру произвольной сложности, чтобы соответствовать даже самым сложным данным.

7. *Bag-of-words*

Поскольку классификаторам машинного обучения всегда нужны числа в качестве входных данных, необработанные текстовые данные не подходят для непосредственного использования в модели. Популярный метод решения проблемы с вводом текстовых данных называется мешком слов, который преобразует текстовую строку в вектор количества слов. В результате преобразования пакета слов каждое уникальное слово во

всем наборе данных представляется как его собственная особенность. Вектор объектов точки данных просто равен нулю для каждого объекта, кроме объектов, представляющих слова, которые существуют в строке. Значения этих функций равны количеству появлений слов в строке.

8. *Метод главных компонент* (*Principle component analysis)*

Принцип компонентного анализа (PCA) попадает в категорию алгоритмов обучения без учителя. Это метод уменьшения размерности, который проецирует набор данных из исходного пространства объектов в новое сокращенное пространство объектов. Новые функции представляют собой линейные комбинации оригинальных функций. Целью процесса преобразования является максимальное сохранение дисперсии.

9. Decision Trees

Древовидные модели, в которых целевая переменная может принимать дискретный набор значений, называются деревьями классификации; в этих древовидных структурах листья представляют метки классов, а ветви представляют соединения функций, которые ведут к этим меткам классов. Деревья решений, в которых целевая переменная может принимать непрерывные значения (обычно действительные числа), называются деревьями регрессии.

Анализ данных с использованием моделей прогнозирования

Для получения хороших данных мы получили четыре различных категории: легальный ввод, XSS-атаки, SQL-инъекции и DDoS атак.

Очистка данных

После фазы обработки данных следующим этапом является очистка набор данных. Проблемой после получения данных из нескольких источников было количество дубликатов. Был создан скрипт Python, удаляющий все дубликаты, входные данные, которые были слишком короткими (и, следовательно, не вызывали реальной атаки), пустые точки данных и нулевые данные

Выявление данных DDoS атак

Рисунок 1

```
np.set_printoptions(precision=3)
pd.set_option('display.float_format', lambda x: '%.3f' % x)
warnings.filterwarnings('ignore')
np.random.seed(8)
%matplotlib inline


def timeit(method):
    def timed(*args, **kw):
        ts = time.time()
        result = method(*args, **kw)
        te = time.time()
        if 'log_time' in kw:
            name = kw.get('log_name', method.__name__.upper())
            kw['log_time'][name] = int((te - ts) * 1000)
        else:
            print('%r  %2.2f ms' % \
                  (method.__name__, (te - ts) * 1000))
        return result
    return timed


data = pd.read_csv('/DDoS_2019_update_dataset.csv')
data_ = data[(data[' Label']!='BENIGN')&(data[' Label']!='WebDDoS')]
len(data_[' Label'].value_counts())
```

Выявление XSS-атак и SQL-инъекции

Рисунок 2

First 5 lines of SQL

	payload	is_malicious	injection_type
0	'\n	1	SQL
1	a' or 1=1-- \n	1	SQL
2	"a"" or 1=1--"\n	1	SQL
3	or a = a\n	1	SQL
4	a' or 'a' = 'a\n	1	SQL

Рисунок 3

First 5 lines of XSS

	payload	is_malicious	injection_type
0	data:text/html;alert(1)/*,<svg%20onload=eval(...	1	XSS
1	">*/--></title></style></textarea></script%0A...	1	XSS
2	" onclick=alert(1)//<button ‘ onclick=alert(1)...	1	XSS
3	';alert(String.fromCharCode(88,83,83))//';aler...	1	XSS
4	">><marquee><img src=x onerror=confirm(1)></ma...	1	XSS

Для преобразования входных данных в числовые объекты использовались два типа методов. Техника с пакетом слов и другая методика использовались в качестве пользовательских функций, таких как длин**а** и байтовое распределение данных.

Набор слов с пространством

N-граммовый подход для преобразования наших выборок данных полезной нагрузки в «слова» размером N.

Пример того, как строка «<script>» будет преобразована при использовании 1-грамма, 2-грамма и 3-грамма.

Пользовательское пространство функций

Перед тренировкой классификаторов было выбрано характеристики с использованием выбора критерий Хи-квадрат $X_c^2 = \sum \frac{(O_i - E_i)^2}{E_i}$.

Критерий Хи-квадрат используется в статистике для проверки независимости двух событий. Учитывая данные двух переменных, мы можем получить наблюдаемый счет *O* и ожидаемый счет *E*. Хи-квадрат показывает, как ожидаемый счет *E* и наблюдаемый счет *O* отклоняются друг от друга.

После создания нашего набора данных, его очистки и создания функций следующим шагом был выбор и обучение наших моделей. Выбранные классификаторы перечислены выше, и они прошли обучение по всем семи пространствам признаков отдельно. Большинство моделей показывают выдающийся результат с общим средним показателем F1 0,9949. Классификатор случайных лесов с наибольшей эффективностью среди всех пространств признаков был классификатором случайного леса со средним баллом F1 0,9913.

Результаты.

В зависимости от инфраструктуры организации определяются их требования к обеспечению безопасности. Проведенный анализ в данной диссертационной работе дает возможность формирования новых правил, политики безопасности в организации или корректировать их, что представляет собой модель обеспечения безопасности.

При наличии:

- БД необходимо обеспечить защиту от SQL-инъекций;
- Финансовых транзакций необходимо обеспечить защиту сессии, канал связи, методов доступа.

Заключение

Основной целью исследования было изучение возможности использования машинного обучения для обнаружения вредоносных угроз с помощью брандмауэра веб-приложений для проведения аудита с высокой точностью для прогнозирования уровня безопасности предприятия. В будущей работе мы можем выдвигать гипотезы, проверяя другие атаки в WAF.

Список использованной литературы

Khashimova, D.A., Atymtayeva, L., Baimuratov, O. Review of threat to information security in the enterprise. *Suleyman Demirel University Bulletin: Natural and Technical Sciences,* 1 (51), 2020.

ElBachir El Moussaid, N., Toumanar, A. F. M. (2014). Web Application Attacks Detection: A Survey and Classification. *International Journal of Computer Applications*, pp. 1-6.

Difference Between Internal & External Threats to an IT Database. URL: https://itstillworks.com/difference-between-internal-external-threats-database-26979.html

Sharmila, D., Umaran, S. Predicting Application Layer DDoS Attacks Using Machine Learning Algorithms. E*ngineering and Technology International Journal of Computer and Systems Engineering*

.