ПРЕДСКАЗАТЕЛЬНАЯ МОДЕЛЬ МАШИННОГО ОБУЧЕНИЯ ДЛЯ РЕШЕНИЯ ЗАДАЧИ КЛАССИФИКАЦИИ УЯЗВИМОСТЕЙ КОМПЬЮТЕРНЫХ СИСТЕМ

Доронин А. К., Липницкий В. А.
Кафедра защиты информации, Белорусский государственный университет информатики и радиоэлектороники
Кафедра высшей математики, Военная академия Республики Беларусь
Минск, Республика Беларусь
E-mail: aliexei.doronin@gmail.com, valipnitski@yandex.by

В докладе рассматривается использование конволюционных нейронных сетей в сочетании с алгоритмом представления слов в многомерном векторном пространстве GloVe для задачи предсказания критичности уязвимости, основываясь лишь на её текстовом описании. В разделе I характеризуется стандарт оценок уязвимостей CVSS. В разделе II приводится анализ распределения оценок уязвимостей из базы данных NVD. В разделе III описывается алгоритм сопоставления слов векторам (GloVe). Раздел IV содержит информацию о практических особенностях построения предсказательной модели.

Введение

В мире с каждым днём появляются всё новые уязвимости компьютерных систем разной степени критичности. Некоторые из них практически не представляют опасности, а другие при эксплуатации способны нанести огромный ущерб не только инфраструктуре, но и человеку. Чем быстрее будет определена степень риска недавно появившейся уязвимости, тем скорее организации смогут предпринять меры для её нейтрализации. Американская национальная база данных уязвимостей NVD содержит около 100 000 записей о различных уязвимостях компьютерных систем, найденных почти за 20 лет развития компьютерных технологий [1]. В этой базе данных (БД) присутствуют следующие поля для описания каждой уязвимости: имя уязвимости, текстовое описание, версия уязвимого ПО, экспертные оценки о критичности. Записи из данной БД представляют собой обширное экспериментальное поле для построения различных предсказательных моделей, особенно в сочетании с применением современных алгоритмов машинного обучения и обработки естественного текста. В данной работе рассматривается попытка применить метод векторного представления слов GloVe [2] в комбинации с конволюционной нейронной сетью в целях предсказания оценки (классификации) уязвимостей по их текстовому описанию.

I. Американская национальная база данных уязвимостей NVD

Американская Национальная база данных уязвимостей NVD (National Vulnerabilities Database) основана на списке уязвимостей из проекта CVE, запущенном MITRE в 1999 году. CVE (Common Vulnerabilities and Exposures) - это общедоступный список записей, каждый из которых содержит идентификационный номер, описание и, по крайней мере, одну общедоступную ссылку для каждой уязвимости. Записи из списка CVE используются в многочисленных продуктах и услугах по кибербезопасности со всего мира, в том числе и NVD. Это означает, что база данных уязвимостей NVD полностью синхронизирована с CVE, однако содержит некоторую дополнительную информацию и предоставляет расширенные функции поиска. Все записи в базе данных NVD имеют 14 полей, из них основными являются следующие:

1. CVE Id: уникальный ID (номер) уязвимости (например, CVE-2011-1585)
2. Date published: дата первой публикации о данной уязвимости
3. Date modified: последняя дата изменения записи
4. Summary: текстовое описание уязвимости

Оставшиеся поля относятся к параметрам оценки уязвимости, выставляемой в соответствии со стандартом CVSS [3]:

1. CVSS Base: базовая оценка уязвимости (десятичное значение от 0 до 10, например, 9.5)
2. CVSS Impact: оценка элементов воздействия (десятичное значение от 0 до 10, например, 9.5)
3. CVSS Exploit: оценка возможности эксплуатации (десятичное значение от 0 до 10, например, 9.5)
4. CVSS Access vector (AV): вектор доступа (локальный/сетевой/локально-сетевой)
5. CVSS Access complexity (AC): сложность доступа (низкая/средняя/высокая)
6. CVSS Authentication (Au): уровень требуемой аутентификации (нулевая/однократная/многократная)
7. CVSS Confidentiality impact (C): воздействие на конфиденциальность (полное/частичное/нулевое)

8. CVSS Integrity impact (I): воздействие на целостность (полное/частичное/нулевое)
9. CVSS Availability impact (A): воздействие на доступность (полное/частичное/нулевое)
10. CVSS Vector: базовый вектор уязвимости содержит в себе значения других полей (например, AV:N/AC:M/Au:N/C:N/I:P/A:C)

II. Анализ базы данных NVD

Для проведения дальнейших исследований критерий оценок CVSS Base был выбран в качестве «главной оценки» уязвимости (рисунок 1).

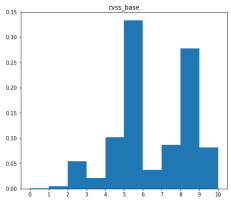

Рис. 1 – Гистограмма распределения оценок CVSS Base

На гистограмме видно, что примерно половина всех оценок сгруппирована вокруг отметки в 4.5 балла, а остальная половина - вокруг 8. Напомним, что данные оценки выставляются экспертами. Перед нами поставлена задача смоделировать систему, которая смогла бы автоматически предсказывать, к какой группе относится новая уязвимость: к группе с низкими оценками (меньше или равно 5 баллам), или же к группе с высокими оценками (более 5 баллов). В качестве входных данных для данной модели наиболее интересным выглядит текстовое описание уязвимости. Таким образом, имеем задачу классификации с двумя классами: класс 0 - уязвимости с низкими оценками, класс 1 - с высокими оценками.

III. Векторное представление слов

Очевидно, что для построения предсказательной модели по текстовому описанию необходимо каким-либо образом осуществить переход от слов к некоторому числовому их эквиваленту. Одним из решений данной задачи является алгоритм GloVe, разработанный в 2013 году [2]. Нами использовался наиболее простой словарь GloVe, содержащий в себе пары "слово – вектор": размерность каждого вектора - 50, общий объём словаря - 400.000. Применение данного словаря позволяет сделать переход к 50-мерному векторному пространству, в котором каждая точка - это вектор, поставленный в соответствие некоторому слову. Ключевая особенность данного алгоритма в том, что на этапе построения вектора слова учитывается контекст окружающих его слов. Это позволяет говорить о том, что при близком расположении точек в векторном пространстве значения соответствующих слов с высокой вероятностью будут так же близки по смыслу.

IV. Построение и обучение нейронной сети

В качестве предиктивного алгоритма было принято решение использовать конволюционную нейронную сеть, так как данный тип нейросетей обычно показывает хорошие результаты в задачах обработки естественного текста. В результате была построена многослойная нейронная сеть с 2 выходами, классифицирующая уязвимости по их текстовому описанию с точностью 85% на отложенной тестовой выборке. На вход нейросети подаётся текстовое описание уязвимости, которое преобразуется в вектор в соответствии со словарём GloVe. Далее преобразованные данные проходят три конволюционных слоя (conv1d), после каждого из них - следует обобщающий (MaxPooling) слой. С целью уменьшения переобучения затем идут три полносвязных (Dense) слоя в сочетании с исключающими (Dropout) слоями. Нейросеть обучалась на 20% от всей выборки, из которых 20% были использованы в качестве валидирующего набора данных. Более подробно с исходным кодом и комментариями можно ознакомиться по ссылке: `https://github.com/teacherlex/cve_vulns_classifier`

Заключение

В докладе дано построение модели, предсказывающей критичность той или иной уязвимости, основываясь лишь на её текстовом описании. В качестве звена, преобразующего текст в векторное пространство, используется словарь проекта GloVe, а в качестве основного обучающегося алгоритма - многослойная конволюционная нейронная сеть. Наилучшая из построенных нами моделей достигла показателя 91% по метрике площади под AUC-ROC кривой. Отметим, что построенную модель можно использовать для оценки любого текста, а не только являющегося непосредственно описанием уязвимости. Например, используя данную модель, можно построить фильтр, отсеивающий лишь наиболее «критичные» тексты. Кроме того, перспективным может оказаться переход к задаче регрессии: в этом случае можно отсортировывать тексты в порядке критичности содержащейся в них информации с точки зрения уязвимостей компьютерных систем.

Список литературы

1. National Vulnerability Database [Электронный ресурс] / The Natural Language Processing Group at Stanford University. -– Режим доступа: https://cve.mitre.org. – Дата доступа: 25.06.2018.
2. GloVe: Global Vectors for Word Representation [Электронный ресурс] / The Natural Language Processing Group at Stanford University. -– Режим доступа: https://nlp.stanford.edu/projects/glove/. – Дата доступа: 07.07.2018.
3. Общая система оценки уязвимостей CVSS [Электронный ресурс] / BIS Expert. -– Режим доступа: http://bis-expert.ru/blog/5345/43124. – Дата доступа: 18.07.2018.