Ежегодная научно-практическая конференция (РусКрипто'2024)
Секция: Перспективные исследования в области кибербезопасности

Прогнозирование уязвимостей в устройствах Интернета вещей

Дмитрий Левшун
Кандидат технических наук, Philosophy Doctor in Computer Science

Санкт-Петербургский государственный университет телекоммуникаций (СПбГУТ)
Санкт-Петербургский Федеральный исследовательский центр Российской академии наук (СПб ФИЦ РАН)

21 марта, 2024

Содержание

- Научная задача
- Основные понятия
- Цель исследования
- Подход к исследованию
- Экспериментальная оценка
- Дискуссия и заключение
- Контакты

Научная задача

- **Ученые** и **разработчики** по всему миру работают над обеспечением **информационной безопасности** сетевых систем.
- Данная задача является сложной и характеризуется **разнообразием угроз**, а также **широким спектром** требований безопасности. Более того, новые уязвимости обнаруживаются на **ежедневной основе**, в то время как известные уязвимости все еще **присутствуют** в работающих системах.
- При этом многие устройства **не представлены** в **открытых базах уязвимостей**, что не позволяет получить информацию об их **уязвимостях**, а также использовать данные уязвимости при построении **графов атак**.
- В данной работе была исследована производительность различных модификаций **BERT** при решении задачи прогнозирования **категорий уязвимостей** в **конфигурациях устройств**.
- В данном исследовании в качестве **конфигураций** рассматривается набор **CPE** (Common Platform Enumeration), где **CPE** это структурированный **шаблон описания** конфигураций устройств.

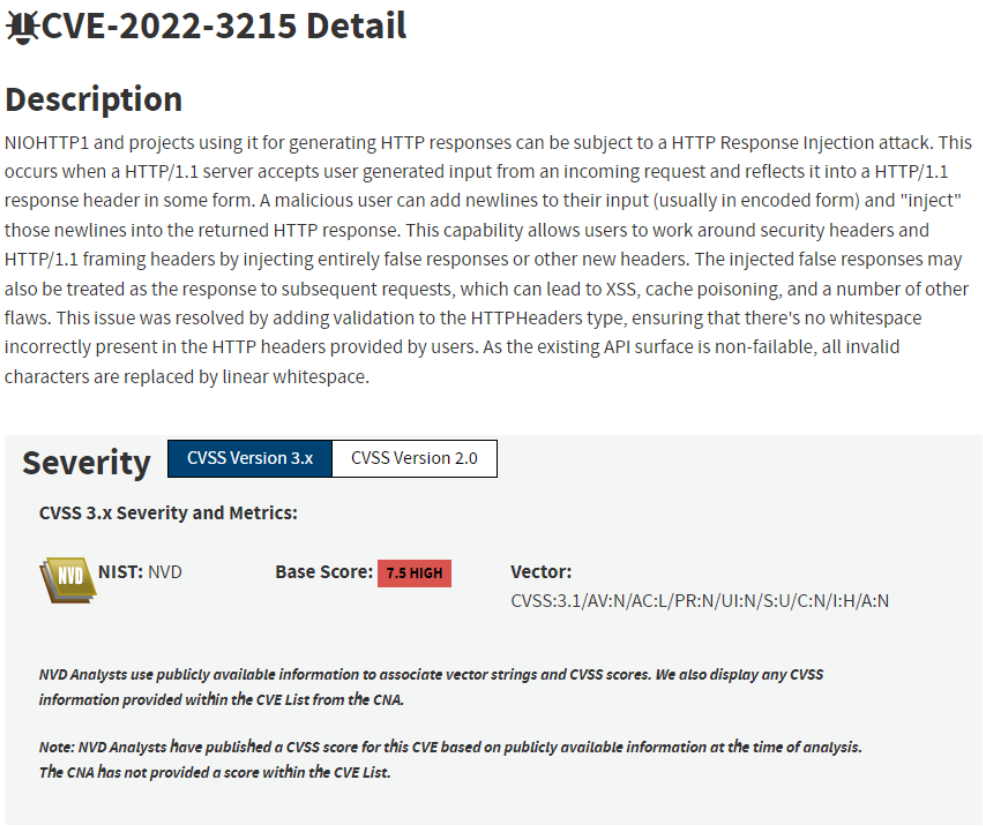
Основные понятия

- **CVE** (Common Vulnerabilities and Exposures) – **формат описания** уязвимостей.

- Ключевой задачей **CVE Program** является идентификация, определение, и хранение **известных** уязвимостей.

- В данном каталоге для каждой уязвимости существует **только одна CVE запись**. При этом задача обнаружения и добавления уязвимостей возложена на организации по всему миру.

- Каждая **CVE запись** содержит следующую информацию: описание, ссылки, конфигурации, метрики **CVSS** (Common Vulnerability Scoring System) 2ой и 3ей версий), слабости **CWE** (Common Weakness Enumeration).



CVE-2022-3215 Detail

Description

NIOHTTP1 and projects using it for generating HTTP responses can be subject to a HTTP Response Injection attack. This occurs when a HTTP/1.1 server accepts user generated input from an incoming request and reflects it into a HTTP/1.1 response header in some form. A malicious user can add newlines to their input (usually in encoded form) and "inject" those newlines into the returned HTTP response. This capability allows users to work around security headers and HTTP/1.1 framing headers by injecting entirely false responses or other new headers. The injected false responses may also be treated as the response to subsequent requests, which can lead to XSS, cache poisoning, and a number of other flaws. This issue was resolved by adding validation to the HTTPHeaders type, ensuring that there's no whitespace incorrectly present in the HTTP headers provided by users. As the existing API surface is non-failable, all invalid characters are replaced by linear whitespace.

Severity CVSS Version 3.x | CVSS Version 2.0

CVSS 3.x Severity and Metrics:

NVD **NIST:** NVD **Base Score:** 7.5 HIGH **Vector:** CVSS:3.1/AV:N/AC:L/PR:N/UI:N/S:U/C:N/I:H/A:N

NVD Analysts use publicly available information to associate vector strings and CVSS scores. We also display any CVSS information provided within the CVE List from the CNA.

Note: NVD Analysts have published a CVSS score for this CVE based on publicly available information at the time of analysis. The CNA has not provided a score within the CVE List.

Источник: https://nvd.nist.gov/vuln/detail/CVE-2022-3215

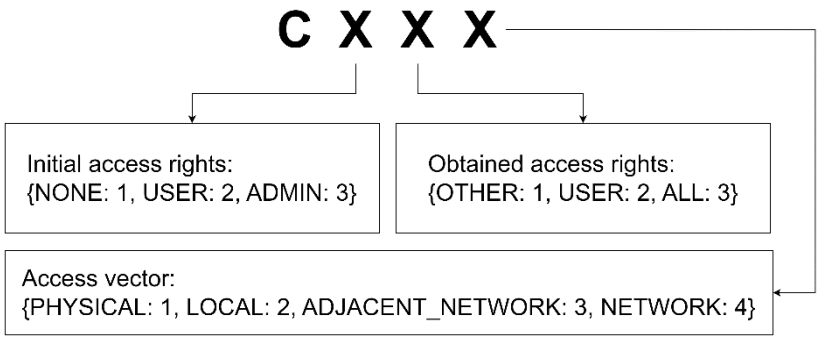
Основные понятия

- **CVSS** (Common Vulnerability Scoring System) – стандарт для описания метрик **CVEs**.
- LOCAL **access vector** из CVSS v2 был разделен на **PHYSICAL** и **LOCAL** access vectors в CVSS v3.
- В CVSS v2 не представлена метрика **privileges required**, а в CVSS v3 – не представлена метрика **obtained privileges**.
- На момент нашей последний проверки, в **NVD** было представлено **199996** CVEs, для **173952** из которых были известны v2 метрики и для **115651** - v3 метрики. Обе версии метрик были доступны только для **100581** CVEs.

		CVSS v2	CVSS v3
Access vector		LOCAL ADJACENT NETWORK NETWORK	PHYSICAL LOCAL ADJACENT NETWORK NETWORK
Privileges required			NONE LOW HIGH
Impact	Confidentiality	NONE PARTIAL COMPLETE	NONE LOW HIGH
	Integrity	NONE PARTIAL COMPLETE	NONE LOW HIGH
	Availability	NONE PARTIAL COMPLETE	NONE LOW HIGH
Obtained privileges	ALL	TRUE FALSE	
	USER	TRUE FALSE	
	OTHER	TRUE FALSE	

Основные понятия

- На основе значений таких метрик, как **access vector** (PHYSICAL, LOCAL, ADJACENT NETWORK, NETWORK), **privileges required** (NONE, USER, ADMIN), и **obtained privileges** (OTHER, USER, ALL) все CVEs были разделены на **24** категории.



- При этом название каждой **категории** CVE **закодировано** в формате **CXXX**, описание которого дано на рисунке выше.

	Description
C111	access PHYSICAL, required NONE, obtained NONE/OTHER
C112	access LOCAL, required NONE, obtained NONE/OTHER
C113	access ADJACENT_NETWORK, required NONE, obtained NONE/OTHER
C114	access NETWORK, required NONE, obtained NONE/OTHER
C121	access PHYSICAL, required NONE, obtained USER
C122	access LOCAL, required NONE, obtained USER
C123	access ADJACENT_NETWORK, required NONE, obtained USER
C124	access NETWORK, required NONE, obtained USER
C221	access PHYSICAL, required LOW, obtained NONE/OTHER/USER
C222	access LOCAL, required LOW, obtained NONE/OTHER/USER
C223	access ADJACENT_NETWORK, required LOW, obtained NONE/OTHER/USER
C224	access NETWORK, required LOW, obtained NONE/OTHER/USER
C131	access PHYSICAL, required NONE, obtained ALL
C132	access LOCAL, required NONE, obtained ALL
C133	access ADJACENT_NETWORK, required NONE, obtained ALL
C134	access NETWORK, required NONE, obtained ALL
C231	access PHYSICAL, required LOW, obtained ALL
C232	access LOCAL, required LOW, obtained ALL
C233	access ADJACENT_NETWORK, required LOW, obtained ALL
C234	access NETWORK, required LOW, obtained ALL
C331	access PHYSICAL, required HIGH
C332	access LOCAL, required HIGH
C333	access ADJACENT_NETWORK, required HIGH
C334	access NETWORK, required HIGH

Основные понятия

- **CPE** (Common Platform Enumeration) – структурированный шаблон для описания аппаратных и программных элементов, приложений, операционных систем и т.п.

cpe:<cpeversion>:<part>:<vendor>:<product>:<version>:<update>:<edition>:
<language>:<swedition>:<targetsw>:<targethw>:<other>

Search Results (Refine Search)

Sort results by: Publish Date Descending Sort

Search Parameters:

- Results Type: Overview
- Keyword (text search): cpe:2.3:a:samsung:account:10.8.0.4:*:*:*:*:*:*:*
- CPE Name Search: true

There are **14** matching records.
Displaying matches **1** through **14**.

Vuln ID	Summary	CVSS Severity
CVE-2022-39875	Improper component protection vulnerability in Samsung Account prior to version 13.5.0 allows attackers to unauthorized logout. **Published:** October 07, 2022; 11:15:23 AM -0400	*V3.1:* 4.4 MEDIUM *V2.0:*(not available)
CVE-2022-39874	Sensitive log information leakage vulnerability in Samsung Account prior to version 13.5.0 allows attackers to unauthorized logout. **Published:** October 07, 2022; 11:15:23 AM -0400	*V3.1:* 5.5 MEDIUM *V2.0:*(not available)

Основные понятия

- В **NVD** взаимосвязи между **CVEs** и **CPEs** заданы с помощью **логических выражений**.

Например, **CVE-1999-0016** связана со следующей **конфигурацией**:

AND(OR({"cpe23Uri": "cpe:2.3:o:cisco:ios:7000:*:*:*:*:*:*:*", "cpe_name": [], "vulnerable": true}), OR({"cpe23Uri": "cpe:2.3:a:gnu:inet:5.01:*:*:*:*:*:*:*", "cpe_name": [], "vulnerable": true}, {"cpe23Uri": "cpe:2.3:a:microsoft:winsock:2.0:*:*:*:*:*:*:*", "cpe_name": [], "vulnerable": true}))

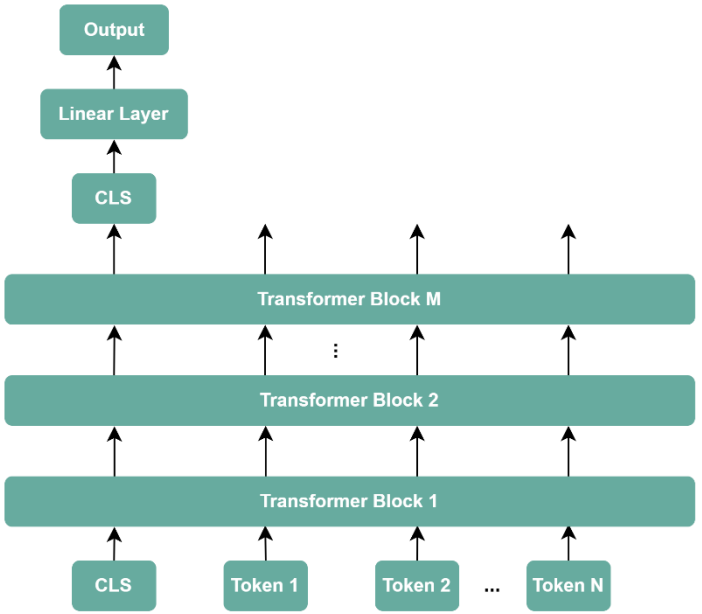
Таким образом, следующие **пары** из операционной **системы** и **приложений** необходимы на устройстве, чтобы оно было уязвимо для эксплуатации **CVE-1999-0016**:

cpe:2.3:o:cisco:ios:7000:*:*:*:*:*:*:* **AND**
cpe:2.3:a:gnu:inet:5.01:*:*:*:*:*:*:*

cpe:2.3:o:cisco:ios:7000:*:*:*:*:*:*:* **AND**
cpe:2.3:a:microsoft:winsock:2.0:*:*:*:*:*:*:*

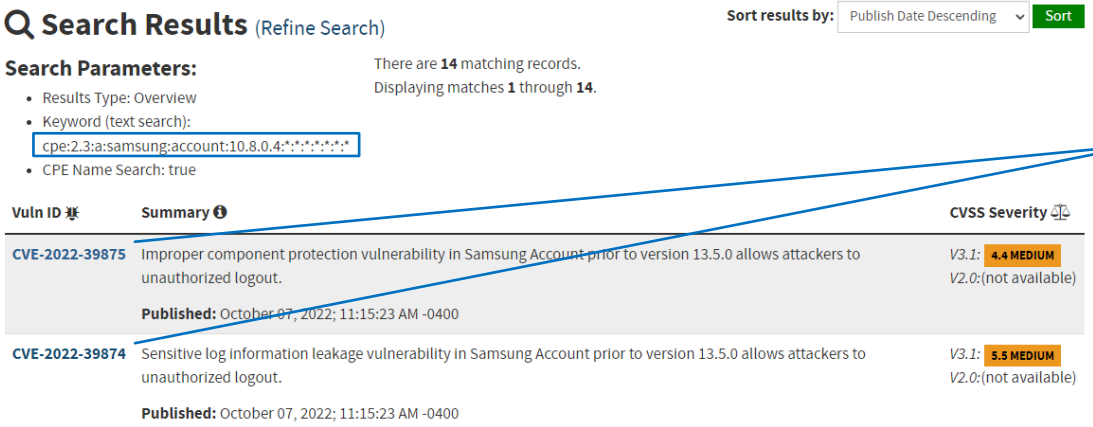
Основные понятия

- **BERT** представляет собой нейросетевую архитектуру на основе трансформеров, которая была разработана для решения задач обработки естественного языка (**NLP**). Данная архитектура использует стек слоев преобразователя-кодировщика и предобучение.
- **RoBERTa** (Robustly Optimized BERT Pretraining Approach) представляет собой модификацию BERT, где модель была предварительно обучена с использованием более длинных последовательностей.
- **XLM-RoBERTa** (Cross-lingual Language Model RoBERTa) представляет собой модификацию RoBERTa, которая была обучена на мультиязычном наборе данных.
- **DeBERTa** (Decoding-enhanced BERT with Disentangled Attention) представляется собой модификацию BERT, где используется механизм распутанного внимания.



Цель исследования

- Известно, что **конфигурации**, которые связаны со схожими **CVEs** могут иметь схожие **CPE URIs**. Предполагается, что **CPE URIs**, которые связаны со схожими **категориями CVE**, сильнее схожи друг с другом, чем **CPE URIs**, которые связаны с другими **категориями CVE**.



Search Results (Refine Search)

Sort results by: Publish Date Descending Sort

Search Parameters:

- Results Type: Overview
- Keyword (text search): cpe:2.3:a:samsung:account:10.8.0.4:*:*:*:*:*:*:*
- CPE Name Search: true

There are **14** matching records.
Displaying matches **1** through **14**.

Vuln ID	Summary	CVSS Severity
CVE-2022-39875	Improper component protection vulnerability in Samsung Account prior to version 13.5.0 allows attackers to unauthorized logout. **Published:** October 07, 2022; 11:15:23 AM -0400	*V3.1:* 4.4 MEDIUM *V2.0:*(not available)
CVE-2022-39874	Sensitive log information leakage vulnerability in Samsung Account prior to version 13.5.0 allows attackers to unauthorized logout. **Published:** October 07, 2022; 11:15:23 AM -0400	*V3.1:* 5.5 MEDIUM *V2.0:*(not available)

	Description
C111	access PHYSICAL, required NONE, obtained NONE/OTHER
C112	access LOCAL, required NONE, obtained NONE/OTHER
C113	access ADJACENT_NETWORK, required NONE, obtained NONE/OTHER
C114	access NETWORK, required NONE, obtained NONE/OTHER
C121	access PHYSICAL, required NONE, obtained USER
C122	access LOCAL, required NONE, obtained USER
C123	access ADJACENT_NETWORK, required NONE, obtained USER
C124	access NETWORK, required NONE, obtained USER
C221	access PHYSICAL, required LOW, obtained NONE/OTHER/USER
C222	access LOCAL, required LOW, obtained NONE/OTHER/USER
C223	access ADJACENT_NETWORK, required LOW, obtained NONE/OTHER/USER
C224	access NETWORK, required LOW, obtained NONE/OTHER/USER
C131	access PHYSICAL, required NONE, obtained ALL
C132	access LOCAL, required NONE, obtained ALL
C133	access ADJACENT_NETWORK, required NONE, obtained ALL
C134	access NETWORK, required NONE, obtained ALL
C231	access PHYSICAL, required LOW, obtained ALL
C232	access LOCAL, required LOW, obtained ALL
C233	access ADJACENT_NETWORK, required LOW, obtained ALL
C234	access NETWORK, required LOW, obtained ALL
C331	access PHYSICAL, required HIGH
C332	access LOCAL, required HIGH
C333	access ADJACENT_NETWORK, required HIGH
C334	access NETWORK, required HIGH

- Значит, можно прогнозировать **категории CVE** для устройств на основе списка их **CPE URIs**.

Подход к исследованию

- **Шаг 1.** Извлечение данных для базы уязвимостей.

Архивы с данными **CVEs** в **JSON** формате могут быть выгружены из **NVD** с их **веб страницы**. В рамках данных **файлов**, у каждой уязвимости есть **описание**, связанные с ней **ссылки**, метрики **CVSS**, уязвимые **конфигурации** и категории **слабостей**.

После того как все файлы **выгружены** и **распакованы**, необходимо подготовить **базу данных** для хранения **CVEs**. В рамках данного исследования, была использована СУБД **PostgreSQL**. После завершения разработки структуры базы данных, осуществляется **обработка** всех файлов и **добавление** извлеченных данных в базу уязвимостей.

Подход к исследованию

- **Шаг 2**. Подготовка наборов данных.

  В разработанной базе данных, информация об **уязвимостях** распределения между множеством **таблиц**. Следовательно, необходимо использовать специальные **SQL запросы** для извлечения **конфигураций**, связанных с **CVEs**.

  Также, при подготовке наборов данных, каждая строка **CPE URI** была предобработана:

  cpe:2.3:a:gnu:glibc:2.38:*:*:*:*:*:*:* → a gnu glibc 2.38

  И связана со значениями **access vector**, **privileges required** и **privileges obtained**.

Подход к исследованию

- **Шаг 3**. Оценка производительности моделей.

Производится оценка производительности моделей при прогнозировании значений таких метрик, как **access vector**, **privileges required** и **privileges obtained**. При этом модели оптимизировались под прогнозирование каждой метрики по отдельности.

Предобработка данных заключалась в обработке текста с помощью предварительно обученных токенизаторов и применяем заполнение до длины **192**. Для поиска лучших решений, был использован оптимизатор **AdamW**.

Подход к исследованию

- **Шаг 4**. Оптимизация гиперпараметров моделей.

  На предыдущем шаге был получен ряд моделей для дальнейшего улучшения. Задача данного шага заключается в выборе **лучших гиперпараметров** для каждой модели по каждой задачи, не допуская переобучения (**overfitting**).

  Оптимизация гиперпараметров была осуществлена с помощью фреймворка **Optuna**. Были использованы **HyperbandPruner** и **TPESampler**, при этом количество попыток было ограничено **150**. Каждая модель была обучена в течении **4 эпох**.

  Итогом данного шага являются **3 оптимизированные модели**.

Подход к исследованию

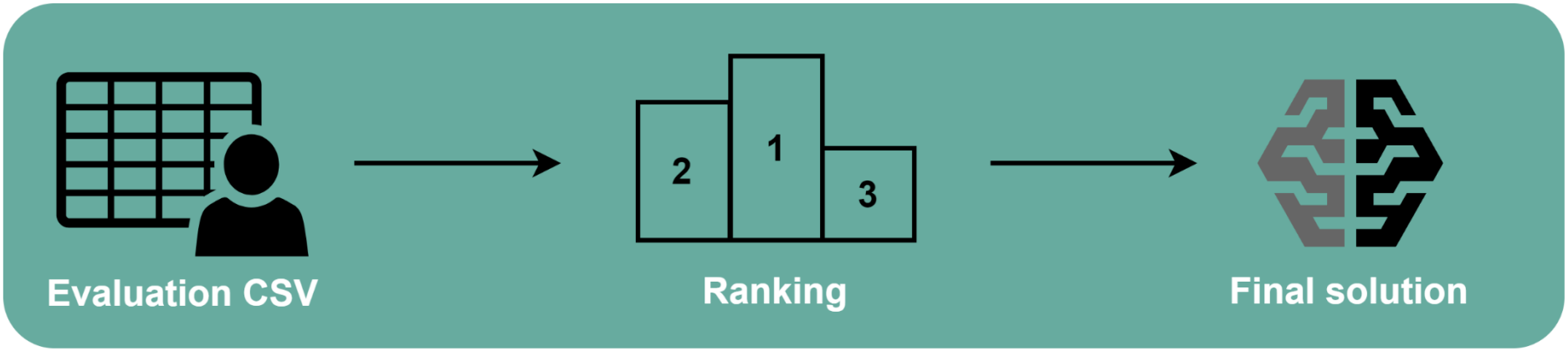
- **Шаг 5**. Оценка прогнозирования категорий уязвимостей.

На данном шаге, мы **объединяем прогнозы моделей**, в список **категорий уязвимостей** в соответствии с **пороговыми значениями**, которые определяют **минимальную вероятность** принятия прогноза:

[0.92: NONE, 0.82: USER, 0.01: ADMIN]

если пороговое значение равно **0.80**, тогда подходят **NONE** и **USER**.

Затем оценивается **качество прогнозирования** категорий уязвимостей, при котором результаты прогнозирования разделяются на **true**, **useful** и **false**.



Экспериментальная оценка

- Набор данных для **access vector**:

cpe,av_nw,av_an,av_lc,av_ph
a redhat jboss_enterprise_web_platform 5.2.0,1,0,1,0
a redhat jboss_enterprise_web_server 2.0.1,1,0,0,0

- Набор данных для **privileges required**:

cpe,pr_none,pr_user,pr_admin
a redhat jboss_enterprise_web_platform 5.2.0,1,0,0
a redhat jboss_enterprise_web_server 2.0.1,1,0,0

- Набор данных для **privileges obtained**:

cpe,po_other,po_user,po_all
a redhat jboss_enterprise_web_platform 5.2.0,1,0,0
a redhat jboss_enterprise_web_server 2.0.1,1,0,0

		Values		Examples	Multilabel
		True	False		
Access vector	*av_nw*	217 126	29 199	246 325	33 554
	av_an	13 330	232 995		
	av_lc	54 364	191 961		
	av_ph	2 684	243 641		
Privileges required	*pr_none*	77 817	25 263	103 080	26 408
	pr_user	43 924	59 156		
	pr_admin	15 164	87 916		
Privileges obtained	*po_other*	229 155	9 248	238 403	16 512
	po_user	9 837	228 566		
	po_all	19 542	218 861		

- **85%** данных были использованы для **обучения** и **15%** для **тестирования**. Все цифры, представленные в таблице, приведены после **удаления дубликатов** данных. **Несбалансированность данных**, была сохранена, т.к. она отражает их **природу**.

Экспериментальная оценка

- Изначально были использованы **базовые версии** модификаций BERT; их параметры представлены в таблице.
- Каждая модель была протестирована **10 раз** на каждом наборе данных. Усредненные результаты (**accuracy**) с их отклонением представлены в таблице.
- Также было оценено **время**, необходимое для обработки данных: BERT – **78 мс**, RoBERTa – **74 мс**, XLM-RoBERTa – **74 мс**, и DeBERTa-v3 – **102 мс**.
- Диапазоны значений гиперпараметров, которые были проверены при оптимизации **BERT** для каждой метрики, представлены в таблице.

	Transformer layers	Hidden dimension	Parameters, in millions
BERT	12	768	110
RoBERTa	12	768	125
XLM-RoBERTa	12	768	125
DeBERTaV3	12	768	184

	Privileges required	Privileges obtained	Access vector
BERT	0.7549 ± 0.0029	0.9462 ± 0.0017	0.8990 ± 0.0016
RoBERTa	0.7487 ± 0.0010	0.9432 ± 0.0025	0.8895 ± 0.0038
XLM-RoBERTa	0.6883 ± 0.0530	0.9319 ± 0.0060	0.8733 ± 0.0019
DeBERTa-v3	0.7501 ± 0.0037	0.9432 ± 0.0013	0.8579 ± 0.0620

	Range of values
Learning rate	from 9e-5 to 1e-5 with 1e-5 step, plus 9e-4 and 8e-4
Warm up epochs	from 0.00 to 1.50 with 0.10 step
Weight decay	from 0.00 to 0.05 with 0.01 step

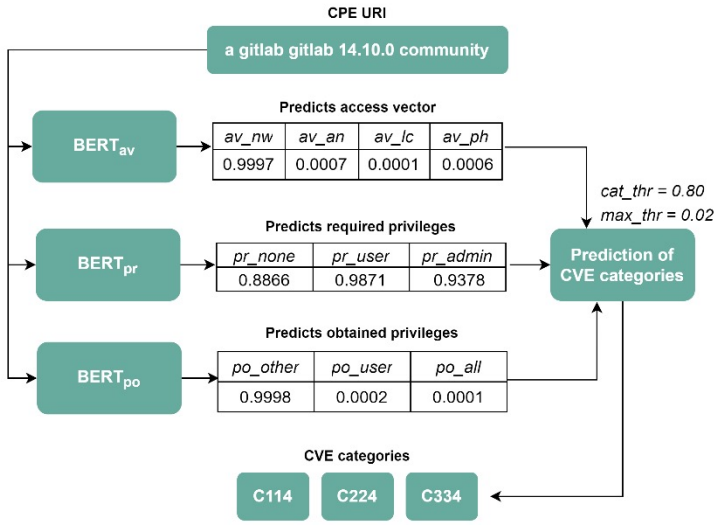
Экспериментальная оценка

		Accuracy	Precision	Recall	F-measure	Support
Access vector	*av_nw*	0.9068	0.97	0.98	0.98	32597
	av_an		0.84	0.89	0.86	2013
	av_lc		0.86	0.84	0.85	8197
	av_ph		0.77	0.61	0.68	386
Privileges required	*po_none*	0.7724	0.91	0.95	0.93	11685
	po_user		0.85	0.82	0.83	6585
	po_admin		0.78	0.65	0.71	2342
Privileges obtained	*po_other*	0.9472	0.99	0.99	0.99	34412
	po_user		0.76	0.74	0.75	1477
	po_all		0.83	0.77	0.80	2920

	Hyperparameter		
	Learning rate	**Warm-up steps**	**Weight decay**
Access vector	7e-05	0.50	0.01
Privileges required	6e-05	0.30	0.04
Privileges obtained	7e-05	0.00	0.05

Экспериментальная оценка

- Прогнозирование основано на **двух порогах**:
  - *cat_thr* определяет **минимум** вероятности для прогнозирования **метрики CVSS**;
  - *max_thr* определяет **допустимый диапазон** значений относительно **максимума** вероятности, если **нет значений** выше *cat_thr*.
- Были проанализированы различные **значения** *cat_thr* и *max_thr* для выбора их **комбинации**:
  - *cat_thr* в диапазоне [0.65; 0.99] с шагом 0.01;
  - *max_thr* в диапазоне [0.00; 0.10] с шагом 0.01.
- Для определения **полезных** прогнозов, рассмотрим пример с ответом **C114 C224 C334**:
  - любые комбинации C114, C224 и C334 рассматриваются в качестве полезных;
  - C114 C224 C334 **C131** не является полезным.

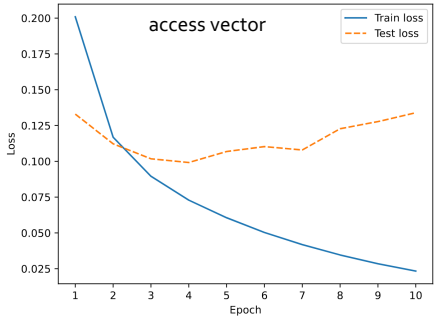


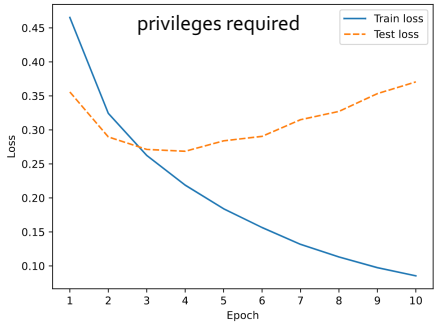
cat_thr	max_thr	Predictions: true	Predictions: partially	Predictions: false	Accuracy	Useful
0.81	0.01	70127	10564	14349	0.7379	0.8490
0.81	0.02	70109	10529	14402	0.7377	0.8485
0.80	0.01	70179	10347	14514	0.7384	0.8473
0.80	**0.02**	**70159**	**10315**	**14566**	**0.7382**	**0.8467**
0.80	0.03	70138	10290	14612	0.7380	0.8463
0.80	0.04	70112	10264	14664	0.7377	0.8457
0.79	0.01	70197	10143	14700	0.7386	0.8453
0.79	0.02	70175	10113	14752	0.7384	0.8448
0.79	0.03	70156	10088	14796	0.7382	0.8443
0.79	0.04	70129	10065	14846	0.7379	0.8438

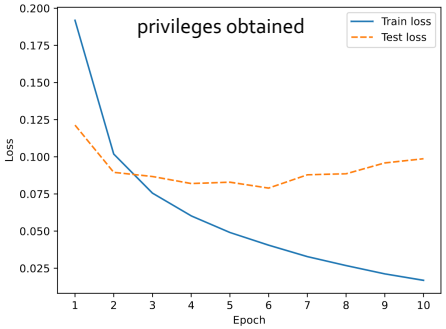
Дискуссия и заключение

- Использованные данные характеризуются сильной **несбалансированностью данных** с **сильным перекосом** в сторону отдельных классов. Были опробованы различные подходы, однако они не принесли результатов.
- Анализ графиков функций потерь показал, что каждая из использованных моделей достаточно быстро начинает переобучаться (**overfitting**). Именно поэтому обучение модификаций BERT ограничивалось **4 эпохами**.
- Во время экспериментов, использовались только те **CPE URI**, которые связаны с CVE **1 к 1**. Отметим, что такие связи характерны для **90%** конфигураций из открытых баз.

	Previous	Current
Approach	Direct prediction of CVE categories	Prediction of CVSS metrics and their combination into CVE categories
Classification	Multi-class	Multi-label
Models	Random Forest	BERT
Accuracy	0.6450	**0.7382**







Контакты

Лаборатория проблем компьютерной безопасности Санкт-Петербургского Федерального исследовательского центра Российской академии наук:

- Адрес: 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, Россия
- Телефон: +7(812)328-71-81
- Факс: +7(812)328-44-50
- URL: http://comsec.spb.ru

Автор:

- Дмитрий Левшун, levshun@comsec.spb.ru, http://comsec.spb.ru/levshun

Благодарность:

- Исследование выполнено за счет гранта Российского научного фонда № 22-71-00107, https://rscf.ru/en/project/ 22-71-00107/.