Софийски университет „Св. Климент Охридски“
Факултет по математика и информатика

Катедра „Софтуерни технологии“
Магистърска програма „Защита на информацията в компютърните системи и мрежи“

Проект

Тема: Сравнителен анализ на разработките за откриване на злонамерен софтуер при TLS трафик чрез методи за машинно обучение

Изготвил: Венета Кирева, ФН: 6MI3400573

Ръководител: д-р Венета Йосифова

Учебна дисциплина „Злонамерен софтуер (Malware)“

Зимен семестър, 2024/2025 г.

Съдържание

Въведение

Според доклад на Google, 96% от трафика, който минава през техните услуги, е криптиран (Google, 2024). Макар това да увеличава сигурността на кореспонденцията за потребителите, криптирането прави трафика труден за анализ, което затруднява откриването на зловреден софтуер. През 2021 година, 91.5% от злонамерения софтуер е бил предаден на потребителите чрез криптирана връзка (WatchGuard Threat Lab, 2021).

В недалечното минало, мрежовият трафик е бил анализиран чрез сравняване с шаблони (Pattern Matching NIDS) или чрез използване метаданните на пакетите. С оглед на невъзможността на традиционните системи за откриване на проникване да се справят с криптиран трафик, все повече компании залагат на методи, основани на машинно обучение, за да филтрират вредния трафик. Това се случва чрез анализ на характеристиките на пакетите и тяхното поведение, без да е необходимо декриптиране.

В настоящия проект са разгледани различни съвременни подходи за откриване на злонамерен софтуер при TLS трафик чрез методи на базата на решения с изкуствен интелект.

Подбор на променливи

При използване на методи за машинно обучение, основни части от предварителната подготовка са изборът и подготовката на данни. Добрият подбор на променливи, които ще се използват в модела, е ключов за успешното решаване на поставената задача.

В контекста на анализ на трафика, най-често това са данни, извлечени от устройствата в наблюдаваната мрежа.

McGrew & Anderson

(McGrew & Anderson, 2016) разглежда подобрен вид телеметрия и анализира желаните свойства на характеристиките в потока от данни спрямо системите за наблюдение на трафика. Правят се експерименти и се извличат заключения как най-ефективно да се използват тези данни класификатори, които да откриват недобронамерен трафик.

Авторите сравняват потенциални характеристики на мрежовия трафик спрямо следните 4 критерия:

- Компактност (Compact) – необходимият брой битове за представяне на данните да е значително по-малък от тях и да е ограничен от константа, независимо от броя на пакетите/битовете в потока;
- Съставимост (Composable) – трябва да е възможно да се комбинират два екземпляра от един тип данни;
- Информативност (Informative) – информацията, която представя характеристиката, да не може да се извлече чрез друга такава;

- Икономичност (Economical) –ресурсите, необходими да се пресметне и извлече, да са относително малки.

Въз основа на тях, в експериментите за обясняващи променливи са избрани:

- Дължина и време на пристигане на пакетите (Sequence of Packet Length and Times, SPLT) - брой на байтовете в пакета и брой милисекунди след получаване на предишния пакет. Събрани са SPLT данните за първите 50 пакета от всеки поток. Всяка стойност е дискретизирана чрез поставянето им в еднакви по големина контейнери, от които се изгражда матрица, представляваща верига на Марков.
- Разпределение на байтовете – представя вероятността на байта на определена позиция да приема дадена стойност. Тъй като съхранението на пълното разпределението би нарушило условието за компактност, вместо него се изчисляват и запазват ентропия на Шанън, средна стойност и стандартно отклонение на байтовете;
- Метаданни при TLS ръкостискане – списък с криптографски схеми (Cipher Suites), списък на обявените разширения и дължината на публичния ключ на клиента;
- Брой пакети и номер на използваните портове.

Stergiopoulos, Talavari, & Bitsikas

В (Stergiopoulos, Talavari, & Bitsikas, 2018) се разглежда система за откриване на зловреден трафик, основана на странични характеристики (Side Characteristics) на пакетите. Подходът може да бъде използван едновременно върху криптиран и некриптиран трафик.

В изследването, характеристиките са избрани като сечение на множествата характеристики от предишни разработки, с цел максимално намаляване на обема на необходимите за разпознаването данни. В резултат са избрани следните 5:

- Големина на пакета (Packet Size, Ps) – добра за определяне на тип връзка и протокол;
- Големина на полезния товар (Payload Size, PAs) – може да подскаже за изтичане на информация;
- Съотношение на големината на полезния товар спрямо големината на пакета (Payload Ratio, PR);
- Съотношение на големината на пакета спрямо тази на предишния пакет (Patio to Previous Packet, Rpp);
- Разлика между времето на пристигане на текущия и предишния пакет (Time Difference, Td).

Shekhawat, Di Troia & Stamp

В (Shekhawat, Di Troia, & Stamp, 2019), авторите се стремят да извлекат максимален брой разнообразни характеристики от данните за връзката, SSL и сертификатите. В последствие, върху тях се прилага анализ за определяне на относителната важност (Relative Importance Analysis) и се вземат под внимание взаимодействията между тях, за да се определи финалното множество, с което ще се провеждат експериментите.

Lokoč, Kohout, Čech, Skopal, & Pevný

Подобно на другите изследвания, в (Lokoč, Kohout, Čech, Skopal, & Pevný, 2016) също за променливи са използвани:

- IP адрес на източник и дестинация;
- Брой изпратени и получени байтове;
- Продължителност на връзката;
- Време на осъществяване на връзката.

За определяне на това дали пакетите са зловредни, е използван SHA хешът на процеса-инициатор на връзката.

Gomez, Kotzias, Dell’Amico, Bilge, & Caballero

(Gomez, Kotzias, Dell’Amico, Bilge, & Caballero, 2023) представя откриване на злонамерен софтуер в криптиран с TLS 1.0, TLS 1.2 или TLS 1.3 трафик чрез използването на клъстеризация. За целта са определени 90 променливи, 67 от които не са използвани в предишни разработки. Променливите биват числови, които са нормализирани, и категорни, които са обработени чрез едноразрядно кодиране (One-Hot Encoding) и резултатът е умножен по TF-IDF на съответните стойности. Характеристиките могат да бъдат разделени в 4 категории:

- На клиента – най-нова поддържана версия на TLS, поддържани шифриращи алгоритми, методи за компресиране, списък поддържани протоколи на приложния слой и др.;
- На сървъра – аналогични на тези на клиента;

- На сертификатите – включва броя на сертификатите, изпратени от сървъра, и полетата на крайния (Leaf) сертификат. Ако са използвани клиентски сертификати, от тях се извличат аналогични данни;
- На полезния товар - размер в брой байтове, максимален последователен брой пакети в посока, съотношението на изпратените и получените байтове и др.

За разлика от другите разработки, времеви характеристики не се използват.

Meghdouri, Vázquez, & Zseby

(Meghdouri, Vázquez, & Zseby, 2020) разглежда начин за междуслойно (Cross-Layer) профилиране на криптиран трафик за откриване на аномалии. В частност, въвежда т.нар. *многоключов характеристичен вектор* (*Multi-key Feature Vector)*, чиято цел е да обедини трите основни начина за представяне на трафика (Application, Conversation и Endpoint) в една характеристика. Този подход се разглежда в съчетание с два протокола за криптиране – TLS и IPsec.

При използването на характеристичен вектор се събира на едно място информация, свързана с индивидуалните статистики за приложенията (CAIA и Consensus), статистики за разговорите (TA) и за единичните хостове (AGM). По този начин се създава по-дълбоко разбиране за анализирания трафик и могат да профилират заплахи въз основа на същността на приложението и разговора, на глобалното поведение на хоста или на всички тях едновременно. Това позволява откриване на зависимости, които не могат да бъдат засечени, ако се използват поотделно.

Криптирането предотвратява извличането на характеристики от определени слоеве - при TLS, всички характеристики над транспортния слой са недостъпни. Съответно, при използване на IPsec, характеристиките над мрежовия слой също. В следствие на това, при използване на IPsec, вместо пълния характеристичен вектор се създава редуциран такъв.

След определяне на данните, върху тях се прилагат нормализация и метода на главните елементи (Principal Component Analysis, PCA). Целта на PCA е намаляване на обема от данни за по-бързо изпълнение на кода без загуба на информация.

Откриване чрез машинно обучение

McGrew & Anderson

Методите се прилагат върху 1 101 599 злонамерени потока, събрани от ThreatGRID, и 1 107 450 доброкачествени, събрани от демилитаризираната зона (DMZ) на корпоративна мрежа. В таблица 1 са представени получените резултати – точност, 0.01% FDR и NPZ, получени при прилагането на логистична регресия с L1-пенализация и 10-разделно кръстосано валидиране.

0.01% FDR (False Discovery Rate) се изразява в процент вярно определени зловредни потоци, ако на всеки 10 000 има точно 1 грешно определен като зловреден пакет. NZP (Non-Zero Parameters) е броят на параметрите със стойност, различна от 0. Резултатите са осреднени спрямо разделянето за кръстосано валидиране.

Данни	Точност	0.01% FDR	NZP	Брой параметри
Всички	**99.59**	**96.41**	116.2	463
SPLT	85.92	1.02	159.7	200
Разпределение на байтовете	90.86	19.68	139.3	256
Метаданни	95.68	0.05	7.0	7

Таблица 1. Резултатите от прилагането на логистична регресия с L1-пенализация и 10-разделно кръстосано валидиране.

Според постигнатите резултати, метаданните са най-значимите характеристики. Най-добра точност (99.59%) очаквано получаваме при използването на комбинацията от всички.

Stergiopoulos, Talavari, & Bitsikas

За провеждане на експерименти са използвани от наборите от данни FIRST 2015, Milicenso и CTU13-1, като са взети приблизително равно количество злонамерен и не злонамерен трафик (таблица 2). При необходимост, част от трафика е криптиран допълнително за някои видове злонамерен софтуер, с цел по-добра представителност.

Тестовете са проведени със 7 различни алгоритъма за машинно обучение:

- Логистична регресия (Logistic Regression);
- Линеен дискриминантен анализ (Linear Discriminant Analysis, LDA);
- K най-близки съседи (K-Nearest Neighbors, KNN);
- Дърво на решенията (Decision Tree Classifier, CART);
- Гаусов наивен Бейсов класификатор (Gaussian Naïve Bayes);
- Машината за поддържащи вектори за класифициране (Support Vector Classifier, SVC);
- Невронни мрежи (Neural Network, NN).

Брой не злонамерен, не криптиран	6 337 244
Брой не злонамерен, криптиран	1 631 274
Брой злонамерен, не криптиран	6 214 670
Брой злонамерен, криптиран	455 211
Общ брой	14 638 399

Таблица 2. Брой пакети, използвани в експериментите.

Резултатите от класификацията – прецизност (Precision), обхват (Recall), F1-оценка (F1-score) и точност (Accuracy) са представени в таблица 3.

Алгоритъм	Прецизност		Обхват		F1-оценка		Точност
	0	1	0	1	0	1	
LR	0.61	0.65	0.84	0.35	0.70	0.45	0.62
LDA	0.60	0.70	0.89	0.31	0.72	0.43	0.62
KNN	**0.95**	**0.91**	**0.92**	**0.94**	**0.94**	**0.93**	**0.93**
CART	**0.96**	**0.92**	**0.94**	**0.95**	**0.95**	**0.94**	**0.95**
NB	0.53	0.17	0.95	0.02	0.68	0.03	0.52
SVC	0.78	0.75	0.74	0.79	0.76	0.77	0.77
NN	0.86	0.84	0.86	0.83	0.86	0.83	0.85

Таблица 3. Резултати от класификацията – пълен набор от данни.

Най-добри резултати се получават при KNN и CART, което авторите обясняват със същността на двата алгоритъма – те вземат решения за класифициране въз основа на намиране на подобия между отделните екземпляри (т.нар. Instance-Based Learning). При модели, които се опитват да намерят и обобщят шаблони в поведението на данните (като LR, NN и т.н.) наблюдаваме значително по-лоши резултати. Това може да е в следствие на недостатъчно количество данни.

Проведен е последващ експеримент със значително по-малко пакети – по 20 000 пакета със злонамерен и незлонамерен трафик. Резултатите от него показват, че този подход е приложим дори при по-малко данни и по-кратко обучение, което го прави подходящ в голям набор от ситуации.

Алгоритъм	Прецизност		Обхват		F1-оценка		Точност
	0	1	0	1	0	1	
LR	0.55	0.52	0.61	0.45	0.58	0.48	0.53
LDA	0.55	0.53	0.64	0.44	0.59	0.48	0.54
KNN	**0.88**	**0.89**	**0.90**	**0.86**	**0.89**	**0.88**	**0.89**
CART	**0.87**	**0.90**	**0.92**	**0.85**	**0.89**	**0.88**	**0.89**
NB	0.90	0.50	0.90	0.99	0.16	0.67	0.54
SVC	0.82	0.96	0.97	0.78	0.89	0.86	0.87

Таблица 4. Резултати от класификацията - намален набор от данни.

Последният проведен тест е прилагане на KNN и CART само върху криптирани данни. При него се наблюдават най-добри стойности и по четирите измерени величини (таблица 5).

Алгоритъм	Прецизност		Обхват		F1-оценка		Точност
	0	1	0	1	0	1	
KNN	1.0	1.0	1.0	0.99	**1.0**	0.99	0.996
CART	1.0	1.0	1.0	1.0	**1.0**	**1.0**	**0.999**

Таблица 5. Резултати от експериментите при криптирани данни.

Shekhawat, Di Troia & Stamp

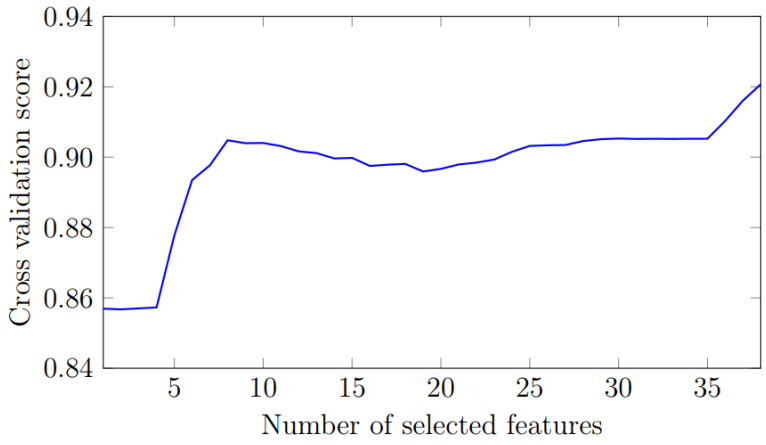
За проведените опити са използвани 3 алгоритъма за машинно обучение:

- Машина за поддържащи вектори за класифициране (SVC);
- Случайна гора (Random Forest, RF);
- Екстремно усилване на градиента (Extreme Gradient Boosting, XGBoost).

Експериментите са осъществени чрез 10-разделно кръстосано валидиране.

Тъй като SVC с линейно ядро присвоява на всяка характеристика тегло, съответстващо на важността ѝ, от опитите се установява, че най-голямо значение имат средната дължина на сертификата, периодичността и средната дължина на публичния ключ.

След като се определят теглата на характеристиките, авторите създават нови модели въз основа на направените заключения и рекурсивно премахване на характеристики (Recursive Feature Elimination, RFE). Резултатите от тях показват, че дори само с 6 от първоначалните 38 характеристики могат да се получат резултати с точност само с 2% по-малка от тази на изходния модел. С най-добрите 10, разликата в точността е близо до 1% (фигура 1).



Фигура 1. Осреднена точност по брой характеристики на модела, подредени спрямо RFE.

Характеристиките, подредени спрямо резултатите им от сортирането им с SVC с линейно ядно и RFE са представени в таблица 6.

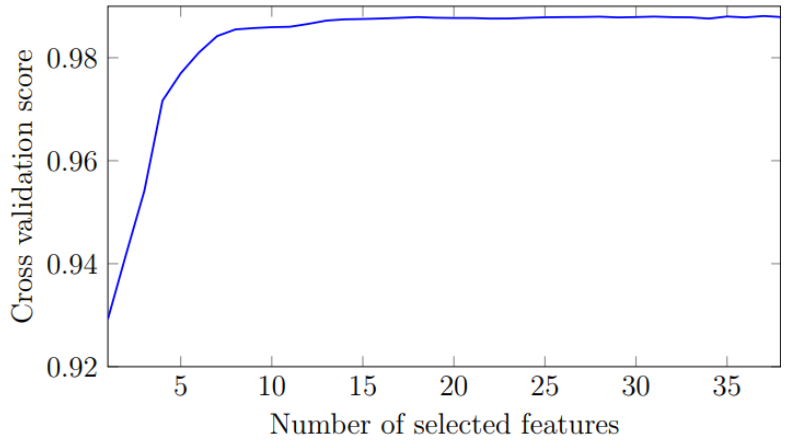
	Характеристики, подредени спрямо SVC с линейно ядро	Характеристики, подредени спрямо RFE
1	periodicity standard deviation	periodicity standard deviation
2	periodicity_average	periodicity_average
3	avg_of_certificate_length	avg_of_duration
4	avg_of_duration	is_SNIs_in_SAN_dns
5	standard_deviation_duration	avg_of_certificate_length
6	is_valid_certificate	avg_key_len
7	is_SNIs_in_SAN_dns	self_signed_ratio
8	avg_key_len	is_same_CN_and_SNI
9	standard_deviation_cert_length	percent_of_established_states
10	self_signed_ratio	differ_issuer_in_ssl_log
11	amount_diff_certificates	differ_subject_in_cert
12	x509_ssl_ratio	ratio_certificate_path_error
13	tls_version_ratio	ratio_of_same_issuer
14	ratio_of_same_subjects	size_of_resp_flows
15	percent_of_established_states	ratio_of_sizes

Таблица 6. Характеристики, подредени спрямо важността им.

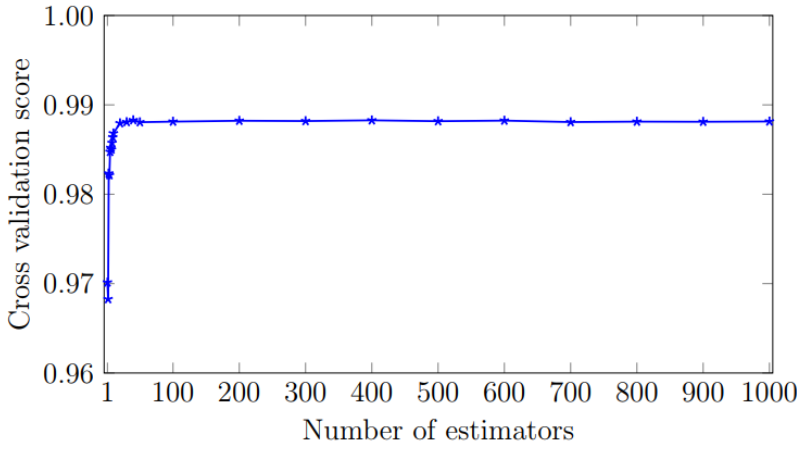
При случайната гора се наблюдава точност от над 96%, дори при моделите с малко предиктори. С увеличаване на броя на предикторите след 50, точността се увеличава незначително (фигура 2). Въпреки това, за оптимални резултати, са изследвани моделите с 500 предиктора.

Спрямо SVC, при RF се постигат по-добри резултати с по-малък брой използвани характеристики (фигура 3). Подредбата им по важност също се различава значително, като от четирите най-важни – size_of_orig_flows,

no_of_cert_path, standard_deviation_duration и ratio_of_sizes, само ratio_of_sizes е сред първите 15 при RFE, и то на последна позиция.

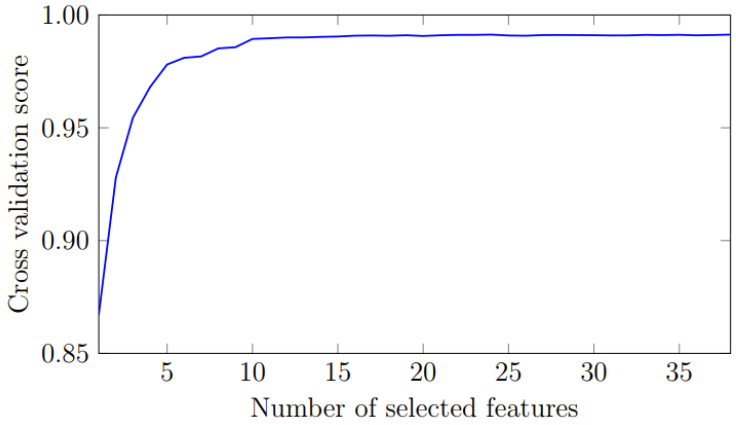


Фигура 2. Осреднена точност на RF спрямо броя на предикторите.



Фигура 3. Осреднена точност спрямо броя използвани характеристики.

При XGBoost се използвани 1000 предиктора (фигура 4). Получените тегла на характеристиките са различни, но подобни на тези от RF. Важно е да се отбележи, че, макар да са най-лесно интуитивно обясними, характеристиките, използвани от SVC, не са тези, които обясняват данните най-точно.



Фигура 4. Осреднена точност на XGBoost спрямо броя на предикторите.

Lokoč, Kohout, Čech, Skopal, & Pevný

За опитите в (Lokoč, Kohout, Čech, Skopal, & Pevný, 2016) са използвани 145 822 799 връзки до 475 695 сървъра. Резултатите са измерени в FP50 - вероятността за фалшиво откриване на зловреден трафик при 50% обхват. Тя е избрана в съответствие със стремежа да се намалят броя на фалшивите разпознавания. Приложено е 6-разделно кръстосано валидиране.

Тествани са 2 класификатора:

- Класификатор с условна максимизация на очакването (Expectation-Conditional Maximisation, ECM);
- K най-близки съседи (KNN) за K = 4.

Резултатите са представени в таблица 7. Изпробвани са различни варианти на KNN класификация, включително точни и приблизителни методи, с или без допълнителни техники (индексиране и TF-IDF).

Малките стойности на K дават по-добри резултати, вероятно тъй като проблемът е нелинеен и зловредните примери са малка част от всички (0.6%). Направена е и t-SNE визуализация, която потвърждава изводите за K -

зловредните данни са разпръснати, а не групирани, което обяснява по-лошото представяне при големи стойности.

Класификатор	FP-50	Време за	
		трениране	класификация
ECM	13.23	56 мин.	0.3 сек.
4-NN без индекс	2.015	0 сек.	63 мин.
4-NN	2.015	44 сек.	17 мин.
4-NN с IDF	2.247	44 сек.	23 мин.
приблиз. 4-NN (4 % БД)	2.017	44 сек.	3 мин.

Таблица 7. Резултати от проведените опити за класификация.

FP-50 стойността на 4-NN и приблизителния 4-NN (4-NN върху 4% от базата от данни) е почти еднаква, но приблизителният метод е значително по-бърз.

Gomez, Kotzias, Dell’Amico, Bilge, & Caballero

В изследването е използван йерархичен алгоритъм за клъстеризация, базиран на Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN), за оценка на 9 различни експеримента върху подмножества на данните:

- FD1: Включва всички характеристики;
- FD2, FD3, FD4, FD5: Всяка конфигурация изключва по една категория характеристики - на клиента, сървъра, полезния товар или сертификатите;
- FD6: Изключва всички освен тези на полезния товар;

- FD7: Изключва 67-те предложени нови характеристики, оставяйки само тези от предишни изследвания.
- FD8: Прилага се нормализация на числовите характеристики;
- FD9: Използва се търсене за оптимизация на хиперпараметрите на модела.

ID	Характеристики	Прецизност	Обхват	F1-оценка
FD1	Всички	0.993	0.872	0.928
FD2	Без клиент	0.992	0.877	0.931
FD3	Без сървър	0.747	0.905	0.819
FD4	Без сертификат	0.996	0.990	0.993
FD5	Без полезен товар	0.982	0.855	0.914
FD6	Полезен товар	0.994	0.990	0.992
FD7	Съществуващи	0.989	0.958	0.973
FD8	Без сертификат	0.995	0.982	0.988
FD9	Без сертификат	**0.996**	**0.990**	**0.993**

Таблица 8. Резултати от експериментите при подбора на различни характеристики.

Резултатите (таблица 8) показват, че характеристиките на сървъра и полезния товар предоставят най-много информация, докато използването на сертификатните е неефективно. Нормализацията допълнително подобрява резултата. Включването на новите предложени характеристики подобрява точността, увеличавайки F1-оценката от 0.973 (FD7) до 0.993 (FD9).

Meghdouri, Vázquez, & Zseby

За експерименти в (Meghdouri, Vázquez, & Zseby, 2020) са използвани 3 набора от данни – CICIDS2017, UNSW-NB15 и ISCX-Bot-2014.

Тествано е представянето на случайна гора. За максимизиране на F1-оценката, за стойностите на хиперпараметрите на модела е приложено търсене, основано на генетични алгоритми (Genetic Algorithms, GA). Избрано е 5-разделно кръстосано валидиране.

В таблици 9, 10 и 11 са представени резултатите. Използването на IPsec значително влияе на откриването на атаки като цяло, но не и на това на ботнети (Botnets). Възможно обяснение е, че потоците на ботнети са изолирани във всеки IPsec тунел и поведението, описано от характеристичния вектор, не е повлияно.

Данни	Тип	Прецизност	Обхват	F1-оценка	Точност
TLS	трениране	1.000	0.999	**0.999**	1.000
	тестване	1.000	0.999	**0.999**	1.000
IPsec	трениране	0.997	0.792	**0.882**	0.999
	тестване	0.992	0.758	**0.859**	0.999

Таблица 9. Резултати при CICIDS2017.

Данни	Тип	Прецизност	Обхват	F1-оценка	Точност
TLS	трениране	0.939	0.934	**0.936**	0.994
	тестване	0.929	0.931	**0.930**	0.994
IPsec	трениране	0.965	0.992	**0.978**	0.994
	тестване	0.965	0.993	**0.979**	0.994

Таблица 10. Резултати при UNSW-NB15.

Данни	Тип	Прецизност	Обхват	F1-оценка	Точност
TLS	трениране	0.999	0.993	**0.996**	0.997
	тестване	0.994	0.989	**0.993**	0.994
IPsec	Трениране	0.999	0.993	**0.996**	0.997
	Тестване	0.998	0.989	**0.993**	0.994

Таблица 11. Резултати при ISCX-Bot-2014.

Друг интересен извод е, че при IPsec се наблюдава по-добра производителност в сравнение с TLS за UNSW-NB15. Авторите отдават това на дизайна на данните, които са събрани с фокус върху характеристики от приложния слой.

След създаване на модела са анализирани използваните променливи. Тези, които се установява, че са с най-голяма тежест, са:

- Обем на данните, изпратени от клиента;
- Стойност и броят на различните TTL стойности;
- Брой на различните IP адреси на дестинацията;
- Най-често използвани TCP флагове;
- Дължина на пакетите;
- Общ брой пакети.

Изводи

Машинното обучение ни дава мощен инструмент за решаване на задачата за откриване на злонамерен криптиран трафик.

Правилният подбор на данни е ключов за коректното класифициране. Противно на интуицията, дори характеристики, които привидно не носят особена конкретика, която да ги свързва със злонамерените пакети (напр. брой байтове и тяхното разпределение), са добър отличаващ белег.

Изборът на подходящ алгоритъм се оказва далеч по-ключов от данните за успеха на експериментите. За преобладаващото мнозинство от опити, при моделите от групата на т.нар. Instance-Based Learning алгоритми – от които с най-добро представяне се отличава случайната гора, се наблюдава много по-добро разграничаване на злонамерения от добронамерения трафик.

Цитирани източници

Gomez, G., Kotzias, P., Dell’Amico, M., Bilge, L., & Caballero, J. (2023). Unsupervised detection and clustering of malicious tls flows. *Security and Communication Networks*, 3676692.

Google. (2024). *Google Transparency Report: HTTPS encryption on the web.* Извлечено от https://transparencyreport.google.com/archive/https/overview

Lokoč, J., Kohout, J., Čech, P., Skopal, T., & Pevný, T. (2016). k-NN classification of malware in HTTPS traffic using the metric space approach. *Intelligence and Security Informatics: 11th Pacific Asia Workshop. PAISI 2016, Auckland, New Zealand, April 19, 2016, Proceedings 11*, 131-145.

McGrew, D., & Anderson, B. (2016). Enhanced Telemetry for Encrypted Threat Analytics. *2016 IEEE 24th International Conference on Network Protocols*.

Meghdouri, F., Vázquez, F. I., & Zseby, T. (2020). Cross-layer profiling of encrypted network data for anomaly detection. *2020 IEEE 7th International Conference on Data Science and Advanced Analytics* , 469-478.

Shekhawat, A. S., Di Troia, F., & Stamp, M. (2019). Feature analysis of encrypted malicious traffic. *Expert Systems with Applications 125*, 130-141.

Stergiopoulos, G., Talavari, A., & Bitsikas, E. (2018). Automatic Detection of Various Malicious Traffic Using Side Channel Features on TCP Packets. *Computer Security: 23rd European Symposium on Research in Computer Security, ESORICS 2018, Barcelona, Spain, September 3-7, 2018, Proceedings, Part I*, 346-362.

WatchGuard Threat Lab. (2021). *Internet Security Report - Q2 2021.* Извлечено от https://www.watchguard.com/wgrd-resource-center/security-report-q2-2021