УНИВЕРЗИТЕТ “Св. КИРИЛ И МЕТОДИЈ” - СКОПЈЕ

ФАКУЛТЕТ ЗА ЕЛЕКТРОТЕХНИКА И

ИНФОРМАЦИСКИ ТЕХНОЛОГИИ

- Докторска дисертација -

Тема:

Примена на федеративно машинско учење за зголемување на безбедност и приватност преку детекција на аномалии на мрежен сообраќај кај Интернет на нешта

Ментор:
Проф. д-р Марија Календар

Докторанд:
м-р Ана Чолакоска Цилакова,
индекс бр. 8/2018д
e-mail: acholak@feit.ukim.edu.mk

Скопје, 2023

УНИВЕРЗИТЕТ “Св. КИРИЛ И МЕТОДИЈ” - СКОПЈЕ

ФАКУЛТЕТ ЗА ЕЛЕКТРОТЕХНИКА И

ИНФОРМАЦИСКИ ТЕХНОЛОГИИ

м-р Ана Чолакоска Цилакова

ПРИМЕНА НА ФЕДЕРАТИВНО МАШИНСКО УЧЕЊЕ ЗА ЗГОЛЕМУВАЊЕ НА БЕЗБЕДНОСТ И ПРИВАТНОСТ ПРЕКУ ДЕТЕКЦИЈА НА АНОМАЛИИ НА МРЕЖЕН СООБРАЌАЈ КАЈ ИНТЕРНЕТ НА НЕШТА

-ДОКТОРСКА ДИСЕРТАЦИЈА-

Скопје, 2023

На моите родители, чија љубов и непоколеблива поддршка беа основата на моето патување. На мојот сакан сопруг, чие трпение и безгранично охрабрување ме доведоа до крајот. На мојата новородена ќерка, која внесува свет на инспирација и радост во мојот живот.Ова достигнување е доказ за љубовта и посветеноста на моето семејство.

Со сето мое срце,
Ана

Ментор: **Проф. д-р Марија Календар**
Универзитет „Св. Кирил и Методиј“– Скопје,
Факултет за електротехника и информациски технологии

Членови на комисијата: **вон. Проф. д-р Даниел Денковски**
Универзитет „Св. Кирил и Методиј“– Скопје,
Факултет за електротехника и информациски технологии

Проф. д-р Берт Арнрих (коментор)
Хасо Платнер Институт, Универзитет во Потсдам,
Факултет за дигитално инженерство,
Дигитално здравје - поврзано здравство

вон. Проф. д-р Христијан Ѓорески
Универзитет „Св. Кирил и Методиј“– Скопје,
Факултет за електротехника и информациски технологии

вон. Проф. д-р Валентин Раковиќ
Универзитет „Св. Кирил и Методиј“– Скопје,
Факултет за електротехника и информациски технологии

Датум на одбрана: _____
Датум на промоција: _____
Научна област: Електротехника и информациски технологии

Изјавувам дека докторскиот труд е оригинален труд што го имам изработено самостојно.

Своерачен потпис на докторандот

Изјавувам дека електронската верзија на докторскиот труд е идентична со отпечатениот докторски труд.

Потпис на авторот, с.р.

м-р Ана Чолакоска Цилакова

ПРИМЕНА НА ФЕДЕРАТИВНО МАШИНСКО УЧЕЊЕ ЗА ЗГОЛЕМУВАЊЕ НА БЕЗБЕДНОСТ И ПРИВАТНОСТ ПРЕКУ ДЕТЕКЦИЈА НА АНОМАЛИИ НА МРЕЖЕН СООБРАЌАЈ КАЈ ИНТЕРНЕТ НА НЕШТА

АПСТРАКТ:

Со оглед на брзата експанзија на мрежите од Интернет на нешта, од голема важност е да се овозможи безбедна интеракција помеѓу уредите без да се влошат нивните перформанси. Користењето на техники за машинско учење се покажува како добра опција за подобрување на детекцијата на аномалии во овие типови мрежи и целокупната безбедност, но нивната имплементација доведува до лоши перформанси и компромитирање на приватноста. Оваа дисертација се фокусира на развој на пристап кој ќе овозможи детектирање на аномалии на мрежен сообраќај во мрежи од Интернет на нешта, притоа зачувувајќи ја приватноста на корисниците и безбедноста на податоците. Предложени се и меѓусебно се споредени три пристапи, од кои специјален фокус е даден на федеративното машинско учење кое овозможува уредите во мрежата заеднички да тренираат и евалуираат модел на машинско учење при локално чување на личните податоци, со што се минимизира можноста за напад и се поттикнува дистрибуцијата на модели и учење во реално време. Резултатите покажуваат дека предложениот пристап има споредбени перформанси во споредба со методите на централизирано длабоко учење, со зачувување на приватноста на податоците и обезбедување на крајните корисници. Дополнително, докторската дисертација нуди темелна анализа на веќе постоечките пристапи за детекција на мрежни аномалии со помош на различни методи од машинско учење, од која се извлекуваат заклучоци за проширување и подобрување на самиот пристап.

КЛУЧНИ ЗБОРОВИ: федеративно машинско учење, интернет на нешта, безбедност, приватност, детекција на аномалии, мрежен сообраќај

Ana Cholakoska Cilakova, MSc.

A FEDERATED MACHINE LEARNING APPROACH FOR IMPROVING THE SECURITY AND PRIVACY OF IOT THROUGH NETWORK TRAFFIC ANOMALY DETECTION

ABSTRACT

Given the rapid expansion of Internet of Things networks, it is of great importance to enable secure interaction between devices without degrading their performance. Using machine learning techniques has proven to be a good option to improve anomaly detection in these types of networks and overall security, but their implementation leads to poor performance and compromising privacy. This dissertation focuses on the development of an approach that will enable the detection of network traffic anomalies in Internet of Things networks, while preserving user privacy and data security. Three approaches are proposed and compared, of which a special focus is given to federated machine learning, which allows devices in the network to jointly train and evaluate a machine learning model while storing personal data locally, thus minimising the possibility of attack and encouraging the distribution of models and real-time learning. The results show that the proposed approach has comparable performance compared to centralised deep learning methods, preserving data privacy and ensuring end users. In addition, the doctoral dissertation offers a thorough analysis of already existing approaches for detecting network anomalies using different methods of machine learning, from which conclusions are drawn for the extension and improvement of the approach itself.

KEYWORDS

federated machine learning, internet of things, security, privacy, anomaly detection, network traffic

Публикации

[ТРУД 1] **A.Cholakoska**, H. Gjoreski, V. Rakovic, D. Denkovski, M. Kalendar, B. Pfitzner, B. Arnrich, "Federated Learning for Network Intrusion Detection in Ambient Assisted Living Environments", *IEEE Internet Computing*, 2023. (accepted for publication) (**IF = 2.68**)

[ТРУД 2] B. Velichkovska, **A.Cholakoska**, V. Atanasovski, "Machine-learning Based Classification of IoT Traffic", *Radioengineering Journal*, 2023. (accepted for publication) (**IF=1.105**)

[ТРУД 3] **A.Cholakoska**, M. Karanfilovska, D. Efnusheva, "Survey of Security Issues, Requirements, Challenges and Attacks in Internet of Things", Informatics and Cybernetics in Intelligent Systems, CSOC 2021, Lecture Notes in Networks and Systems, vol. 228. **Springer**, Cham. 2021.

[ТРУД 4] **A.Cholakoska**, M. Shuslevska, Z. Todorov, D. Efnusheva, "Analysis of machine learning classification techniques for anomaly detection with NSL-KDD data set", Data Science and Intelligent Systems, CoMeSySo 2021, Lecture Notes in Networks and Systems, vol. 231. **Springer**, Cham. 2021.

[ТРУД 5] **A.Cholakoska**, B. Pfitzner, H. Gjoreski, V. Rakovic, B. Arnrich, M. Kalendar, "Differentially Private Federated Learning for Anomaly Detection in eHealth Networks", Proceedings of **ACM** International Joint Conference on Pervasive and Ubiquitous Computing, 2021.

[ТРУД 6] **A.Cholakoska**, B. Jakimovski, B. Pfitzner, H. Gjoreski, B. Arnrich, M. Kalendar, D. Efnusheva, "Network Anomaly Detection using Federated Learning for the Internet of Things", Pervasive Health and Smart Sensing conference, Ljubljana, Slovenia, 2022.

[ТРУД 7] **A.Cholakoska**, V. Rakovic, H. Gjoreski, B. Pfitzner, B. Arnrich, M. Kalendar, "Machine Learning based Anomaly Detection in Ambient Assisted Living Environments", ETAI, 2021.

[ТРУД 8] M. Karanfilovska, T. Kochovska, Z. Todorov, **A.Cholakoska**, G. Jakimovski, D. Efnusheva, "Analysis and modelling of a ML-based NIDS for IoT networks", International Conference on Industry Sciences and Computer Science Innovation, Procedia Computer Science, vol. 204. **Elsevier**, 2022.

[ТРУД 9] I. Senchuk, **A. Cholakoska**, D. Efnusheva, "Analysis of Smart Home Security by Applying Machine Learning Algorithms", ETAI, 2021.

[ТРУД 10] M. Shushlevska, **A. Cholakoska**, D. Efnusheva, "Network Security Analysis by Applying Machine Learning Algorithms", ETAI, 2021.

[ТРУД 11] Z. Bogoevski, Z. Todorov, M. Gjosheva, D. Efnusheva, **A.Cholakoska**, ”A Monitoring System Design for Smart Agriculture”,Cybernetics Perspectives in Systems,CSOC 2022, Lecture Notes in Networks and Systems, vol 503. **Springer**, Cham. 2022.

[ТРУД 12] Z. Todorov, D. Efnusheva, **A. Cholakoska**, M. Kalendar, ”FPGA Implementation of IPV6 Header Processor”, 9th International Conference on Applied Innovations in IT (ICAIIT), Koethen, Germany, 2021.

[ТРУД 13] **A.Cholakoska**, D. Efnusheva, M. Kalendar, “Hardware Implementation of IP Packet Filtering in FPGA”, 7th International Conference on Applied Innovations in IT (ICAIIT), Koethen, Germany, 2019.

Содржина

Табела на слики

Табела на табели

Акроними

Кратенка	Значење	Англиски превод
6LoWPAN	IPv6 преку безжични лични мрежи со мала моќност	IPv6 over Low-power Wireless Personal Area Networks
AAL	Амбиентално асистивно живеење	Ambient Assisted Living
AD	Детекција на аномалии	Anomaly Detection
AES	Напреден стандард за шифрирање	Advanced Encryption Standard
ANN	Вештачки Невронски мрежи	Artificial Neural Network
AP	Точка на пристап	Access Point
API	Програмски интерфејс за апликации	Application Programming Interface
ARP	Протокол за резолуција на адреса	Address Resolution Protocol
AUC	Површина под ROC кривата	Area Under the ROC Curve
BLE	Bluetooth со мала енергија	Bluetooth Low Energy
CNN	Конволуциска невронска мрежа	Convolutional Neural Network
CoAP	Протокол за ограничени апликации	Constrained Application Protocol
CRM	Управување на односите со клиентите	Customer relationship management
CSMA-CA	Повеќекратен пристап на носител со избегнување на судири	Carrier Sense Multiple Access with Collision Avoidance
DDoS	Дистрибуирано одбивање на услуга	Distributed Denial of Service
DDS	Услуга за дистрибуција на податоци	Data Distribution Service
DL	Длабоко учење	Deep Learning
DNN	Длабока невронска мрежа	Deep Neural Network
DoS	Одбивање на услуга	Denial of Service
DP	Диференцијална приватност	Differential Privacy
DPFedAvg	Федеративен просек со диференцијална приватност	Differential Privacy Federated Averaging

DT	Дрво на одлука	Decision Tree
ERP	Планирање на ресурси на претпријатија	Enterprise Resource planning
FedAdagad	Федеративен Adagrad	Federated Adagrad
FedAvg	Федеративен просек	Federated Averaging
FedMedian	Федеративна медијана	Federated Median
FedProx	Федеративен проксимал	Federated Proximal
FedYogi	Федеративен Yogi	Federated Yogi
FHSS	Техника на модулација на распространет спектар со фреквенциско скокање	Frequency Hopping Spread Spectrum
FL	Федеративно учење	Federated Learning
FN	Лажно негативен	False Negative
FP	Лажно позитивен	False Positive
FTFedAvg	Федеративен просек со толеранција на грешки	Fault Tolerant Federated Averaging
GFSK	Гаусова техника за модулација на фрекфенцијата	Gaussian Frequency Shift Keying
ICU	Оддел за интензивна нега	Intensive Care Unit
IoT	Интернет на нешта	Internet of Things
ITU	Меѓународна телекомуникациска унија	International Telecommunication Union
KNN	K најблиски соседи	K Nearest Neighbours
LR	Логистичка регресија	Logistic Regression
LSTM	Долга краткорочна меморија	Long Short Term Memory
MitM	Човек во средина	Man in the Middle
ML	Машинско учење	Machine Learning
MQTT	Телеметриски пренос во редица на пораки	Message Queuing Telemetry Transport
NB	Наивен Баесов (класификатор)	Naive Bayes

NB-IoT	Теснопојасен Интернет на нешта	Narrowband IoT
NIDS	Мрежен систем за детекција на упади	Network Intrusion Detection System
PCA	Анализа на главни компоненти	Principal Component Analysis
QoS	Квалитет на услуга	Quality of Service
R2L	Далечинско до локално	Remote to Local
RF	Случајна шума	Random Forest
RPL	Протокол за рутирање за мрежи со мала моќност и загуби	Routing Protocol for Low-power and Lossy Networks
SIEM	Систем за управување со безбедносни информации и настани	Security Information and Event Management
SQL	Структуриран јазик за пребарување	Structured Query Language
SSL	Ниво на безбедни приклучоци	Secure Sockets Layer
SVM	Векторски машини за поддршка	Support Vector Machine
TLS	Безбедност на транспортно ниво	Transport Layer Security
TN	Вистински негативен	True Negative
TP	Вистински позитивен	True Positive
U2R	Корисник до администратор	User to Root
USB	Универзална сериска магистрала	Universal Serial Bus
XGBoost	Екстремен градиент на засилување	Extreme Gradient Boosting
XSS	Криптирање меѓу страници	Cross Site Scripting

Глава 1 Вовед

Во последната деценија, сведоци сме на зачудувачки пораст на поврзаноста на уреди на мрежа. Претходно, концептот на поврзаност бил ограничен на интеракции помеѓу луѓето и машините. Денес, со напредокот на технологијата може да се овозможи поврзаност речиси помеѓу било какви уреди. Брзиот развој на Интернет на нештата (IoT) води до разновидност на апликации во општеството, овозможувајќи поедноставување и подобрување на квалитетот на животот на крајните корисници. Како поим, Интернет на нешта се однесува на мрежа од уреди поврзани на Интернет кои можат да комуницираат и разменуваат податоци меѓу себе. Овие уреди може да бидат во опсег од сензори и камери до апарати за домаќинство и автомобили.

Мрежите од Интернет на нешта денес имаат широка употреба - од подобрување на здравството, па се до зголемување на ефикасноста во индустријата, земјоделството и транспортот и урбаното планирање. Со помош на различни сензори за следење на човековото тело, можностите за дијагноза, рана превенција, лекување и администрација на лекови стануваат побрзи и полесни [1]. Исто така, мобилните уреди и паметни часовници со акцелерометри и пулсоксиметри денес играат клучна улога во далечинското следење на пациентите и нивната здравствена евалуација (детекција на пад, поставување дијагноза, прилагодување на терапија итн.) [2].

IoT уредите може да се користат за следење на опремата во производството и идентификување на потенцијалните проблеми со одржување пред тие да станат големи. Ова може да помогне да се намали времето на застој и да се зголеми целокупната продуктивност. Исто така, тие може да се користат за следење на залихите на производствените капацитети и автоматско преднарачување на набавките кога е потребно, подобрувајќи го управувањето со синџирот на снабдување [3,4].

Во современото земјоделие, IoT уредите се користат за следење на условите на почвата, временските шаблони и други фактори кои можат да влијаат на растот на културите. Со собирање и анализа на овие податоци, земјоделците можат да донесат подобри одлуки за распоредот на садење, наводнувањето и употребата на ѓубрива, што на крајот ќе доведе до повисоки приноси и намален отпад. Дополнително, сензорите може да се користат за следење на здравјето и благосостојбата на добитокот, овозможувајќи им на сточарите брзо да ги идентификуваат и третираат сите потенцијални проблеми [5-7].

IoT уредите се користат и за подобрување на безбедноста и ефикасноста на транспортот. Поврзаните возила можат да комуницираат едни со други за да избегнат несреќи и да го намалат застојот, додека сензорите може да се користат за следење на патничката инфраструктура како мостови и автопати за знаци на истрошеност [8]. Дополнително, логистичките компании користат IoT уреди за да ги следат пратките и да ги оптимизираат патиштата за испорака, подобрувајќи ја ефикасноста на синџирот на снабдување [9]. Во поглед на урбанистичкото планирање и јавната безбедност, паметните семафори и поврзаните возила можат да работат заедно за да го намалат сообраќајниот метеж и да го подобрат протокот на сообраќај. Дополнително, IoT уредите може да се користат за следење на квалитетот на воздухот и откривање на жариштата на загадување во градовите [10-12].

Овие уреди може да се користат за да се создаде поповрзана и поефикасна домашна средина [13-16]. На пример, паметните термостати можат да ги научат навиките за греење и ладење на корисниците и автоматски да ја приспособат температурата за да заштедат енергија. Слично на ова, паметните системи за осветлување може да се програмираат за автоматско вклучување/исклучување врз основа на распоредот на корисниците, намалувајќи ја потрошувачката на енергија и подобрувајќи ја удобноста.

IoT уредите може да се користат за подобрување на искуството на клиентите и зголемување на продажбата во малопродажниот сектор. На пример, трговците можат да користат IoT сензори за да ги следат движењата на клиентите во продавница, да ги анализираат моделите за купување и да прават персонализирани промоции и попусти врз основа на нивните навики за купување [17-20]. Секако, со секојдневниот пораст на бројот на уреди, се очекува навлегување и подобрување и на други сфери од човековото живеење.

Но, како што расте бројот на уреди, така расте и опасноста од можни напади. Паметна мрежа заснована на IoT уреди може да се состои од милиони јазли, што може да ја зголеми веројатноста од напад поради големата површина на располагање. Ваков фокусиран напад на пример, може да ја исклучи електричната мрежа, имајќи ги во предвид повеќето јавни комунални претпријатија, транспортните услуги и домашните уреди што работат на електрична енергија, при што последиците би биле катастрофални. Општо познато е дека штом еден уред е компромитиран, тогаш остатокот од мрежата станува ранлива. Што е уште полошо, верижниот ефект врз ваквите напади може да запре цели градови и да доведе до незамисливи финансиски и економски загуби [21-23].

Поврзаните уреди стануваат сè попопуларни и овој тренд исто така задава многу безбедносни предизвици и во домот. Зголеменото распоредување на уреди поврзани со Интернет во домот, ги изложува корисниците на ризик, бидејќи личните информации стануваат далечински достапни. Напаѓачот може, на пример, да го прислушкува безжичниот пренос на сензорите и да ги открие активностите на корисниците. Исто така, напаѓачот може од далечина да ја преземе контролата врз домашните уреди, користејќи ги како платформа за извршување напади кон трети страни.

Во областа на е-здравство, ризикот од нарушување на приватноста на податоците на луѓето/пациентите и нивната електронска здравствена евиденција исто така се зголемува. Вмрежените уреди може далечински и злонамерно да се исклучат, реконфигурираат или репрограмираат, што може да ги стави пациентите во ризик или да има катастрофални последици по нивното здравје [24,25]. На пример, постои злонамерен софтвер кој може да биде распореден на пејсмејкери или инсулински пумпи [26], што би можело брзо да резултира со исклучување на уредот и потенцијална смрт на пациентот.

За да се задоволат потребите на крајните корисници, IoT уредите мора да исполнуваат одредени безбедносни барања. Сепак, традиционалните мрежни барања како доверливост, интегритет и достапност не се доволни за овие специјализирани типови на уреди. За вакви системи да се сметаат за безбедни, тие мора да исполнуваат дополнителни безбедносни барања [27].

Со цел да се зачува безбедноста на крајните корисници и нивните податоци, се бараат нови и подобрени начини за детекција на аномалии (напади) во реално време за да може да се направат навремени одговори. Еден од најчестите пристапи кога станува збор за детекцијата на аномалии во мрежи на Интернет на нешта е користење на мрежни

системи за детекција на упад или системи за детекција на упад базирани на домаќин. И двата пристапа користат различни метрики за да ги одделат потписите (познати напади) и аномалиите (непознати напади) од нормален проток на мрежен сообраќај.

Пристапите засновани на потписи можат лесно да откриваат познати шаблони, но немаат способност за детекција на нови напади. Со учење од вообичаени мрежни податоци, пристапите базирани на детектирање на аномалиии имаат подобри резултати бидејќи се способни да детектираат и непознати напади покрај познатите. Сепак, ваквиот пристап носи поголеми пресметковни трошоци и намалување на точноста на детекција [28-30].

1.1 Предмет на истражување

Предмет на истражување на оваа докторска дисертација се можностите за развој, реализација и применливост на нов пристап за систем за детекција на аномалии во мрежен сообраќај кај Интернет на нешта. Во тој процес, најпрво е направено испитување на генералната архитектура на Интернет на нешта, протоколниот пласт, како и посочени се посебните барања кои мора да ги исполнат ваквите мрежи и уреди. Детално, анализирани се предизвиците и заканите со кои секојдневно се соочуваат мрежите на Интернет на нешта.

Посебен акцент во оваа докторска дисертација е даден на употребата на машинското учење како метод за детекција на аномалии во мрежен сообраќај кај Интернет на нешта. Во оваа фаза од истражувањето, направена е компаративна анализа на веќе постоечки системи за детекција на аномалии во мрежен сообраќај кај мрежи од Интернет на нешта, по што е направена анализа на постоечки податочни множества од мрежен сообраќај на Интернет на нешта. По изборот на соодветни податочни множества, предложен/дизајниран е модел кој користи методи од длабоко и машинско учење за детекција на аномалии во мрежен сообраќај на Интернет на нешта.

Бидејќи повеќето вакви системи се централизирани и постои можност од нарушување на приватноста и безбедноста, како и губење на податоците, во оваа докторска дисертација е истражена можноста за создавање на федеративен (дистрибуиран) пристап и модел за детекција на аномалии во мрежен сообраќај од Интернет на нешта, кој користи диференцијална приватност за да ја гарантира приватноста на самите податоци кои се испраќаат. Понатаму, направена е споредба на моделите кои користат централизирано и федеративно машинско учење, при што е даден соодветен заклучок (имајќи ги во предвид и загубите во точност при користење на диференцијалната приватност). Исто така, испитана е и применливоста на ваквиот пристап.

1.2 Цели на истражувањето

Основната цел на докторската дисертација е да се придонесе кон развојот на пристап кој ќе овозможи детектирање на аномалии на мрежен сообраќај во мрежи од Интернет на нешта, притоа зачувувајќи ја приватноста на корисниците и безбедноста на податоците. Истовремено, целта на ваквиот пристап е да се промовира меѓусебната соработка на нејзините ентитети (болници, паметни домови итн.). Во потесен контекст, тоа е постигнато преку развој, симулација и анализа на методи и техники од длабоко и федеративно машинско учење.

Во насока на постигнување на оваа цел, предложени се неколку подцели на

истражувањето:

- Преглед на релевантната литература во областа на мрежни системи за откривање на упади во Интернет на нешта со користење на методи на машинско учење;
- Компаративна анализа на постоечки системи за детекција на аномалии на мрежен сообраќај во мрежи на Интернет на нешта;
- Компаративна анализа на различни податочни множества од мрежен сообраќај во мрежи на Интернет на нешта;
- Конструирање, тренирање и евалуација на модел на длабоко машинско учење за детекција на аномалии во мрежен сообраќај кај Интернет на нешта;
- Конструирање, тренирање и евалуација на модел на федеративно машинско учење за детекција на аномалии во мрежен сообраќај кај Интернет на нешта;
- Предложување нов пристап за детекција на аномалии во мрежен сообраќај кај Интернет на нешта;
- Примена на предложените модели на тест мрежи, визуелен приказ и анализа на добиените резултати при споредба со останати постоечки системи за детекција на аномалии во мрежен сообраќај кај Интернет на нешта.

1.3 Хипотези

Во оваа докторска дисертација, главната хипотеза е дека предложениот дистрибуиран пристап за откривање аномалии во мрежи од Интернет на нешта ќе помогне во зачувувањето на приватноста на крајните корисници и ќе ја промовира соработката помеѓу субјектите (паметни домови, болници, одделенија за интензивна нега итн.) што ќе бидат вклучени во самиот систем, додека во исто време ќе ја обезбедува мрежата. Оваа хипотеза се заснова на неколку посебни хипотези:

Хипотеза 1: Предложениот модел ќе овозможи прецизна детекција на аномалии во мрежен сообраќај во мрежи од Интернет на нешта

Хипотеза 2: Предложениот модел ќе овозможи соработка меѓу ентитетите (паметни домови, болници, оддели за интензивна нега) кои што ќе бидат вклучени во самиот систем

Хипотеза 3: Предложениот модел ќе овозможи одржување на приватноста на корисниците на системот и нивните податоци

Хипотеза 4: Предложениот модел ќе даде прифатливи резултати во споредба со централизираните модели за детекција на аномалии во мрежи од Интернет на нешта

1.4 Научен придонес

Според целите на истражувањето главен научен придонес на оваа докторска дисертација е изградба на ефикасен систем за детекција на аномалии во мрежи од Интернет на нешта. Поспецифично, се очекува предложениот пристап да обезбеди безбедни и приватни податоци, како и соработка помеѓу ентитетите кои се вклучени. Во

рамките на ова истражување се очекуваат и следните придонеси:

- Теоретска анализа на техниките и методите кои се користат за детекција на аномалии во мрежен сообраќај кај мрежи од Интернет на нешта.

- Компаративна анализа на различни системи за детекција на аномалии во мрежен сообраќај кај мрежи од Интернет на нешта, нивни предности, недостатоци и применливост.

- Развој на пристап за дистрибуиран систем за детекција на аномалии во мрежен сообраќај преку:

  ◦ Развој и анализа на модел за бинарна класификација на аномалии во мрежен сообраќај и нормален сообраќај кај мрежи од Интернет на нешта
  ◦ Развој и анализа на модел за повеќекласно класифицирање на аномалии во мрежен сообраќај и нормален сообраќај кај мрежи од Интернет на нешта
  ◦ Развој и анализа на модел со длабоко учење за класифицирање на аномалии во мрежен сообраќај
  ◦ Развој и анализа на модел со примена на федеративно машинско учење за детекција на аномалии во мрежен сообраќај кај мрежи од Интернет на нешта

- Анализа на перформанси на новосоздадениот дистрибуиран систем и споредба со веќе постоечки системи.

- Анализа на применливост на новосоздадениот дистрибуиран систем.

1.5 Структура

Докторската дисертација е структурирана во **седум** глави.

Глава 1: Првата глава претставува вовед во проблематиката на безбедност во IoT мрежите. Таа ја изложува мотивацијата зад истражувањето. Врз основа на овој преглед е даден предметот на истражување на оваа докторска дисертација, хипотезите, како и нејзиниот научен придонес. Понатаму, поставени се целите во истражувањето и дадена е структура на самата дисертација.

Глава 2: Во глава 2 е даден детален преглед на архитектурата на IoT, слоевите кои ја сочинуваат и нивната улога во функционирањето на вакви типови системи. Понатаму, детално се протоколите специфично дефинирани за IoT и претставен е нивниот начин на работа.

Глава 3: Бидејќи главната тема на ова истражување е безбедноста во IoT мрежи, во оваа глава се анализирани предизвиците, како и барањата на кои ваквите типови на уреди и мрежи треба да одговорат за да се сметаат за безбедни за користење. Во посебен дел од оваа глава е дадена темелна анализа на нападите кои можат да се појават во IoT, издвоени по нивоата кои може да ги зафатат. На крај, направена е и компаративна анализа на сите обработени напади. Поради нивната разновидност и големиот број, тука се нагласува потребата од различни пристапи за заштита на овие мрежи.

Глава 4: Во глава 4 е направена темелна анализа на типовите на системи за детекција на аномалии во мрежен сообраќај. Понатаму, објаснети се процесите на машинско и длабоко учење, како и нивното значење и употреба во системите за детекција на аномалии во мрежен сообраќај во ваквите системи. Во овој дел исто така е направена посебна компаративна анализа на секој тип учење преку моменталната актуелна литература. На крај на оваа глава, направена е компаративна анализа и на податочните множества кои се користат при дизајнирање на системи за детекција на аномалии во мрежен сообраќај.

Глава 5: Во овој дел од докторската дисертација се дадени основите на федеративното машинско учење и прикажан е процесот на учење. Понатаму, направена е посебна компаративна анализа на системи за детекција на аномалии во мрежен сообраќај на IoT со користење на федеративно машинско учење.

Глава 6: Врз основа на претходните анализи во Глава 2,3,4 и 5, во оваа глава прикажан е развојот на пристап за детекција на аномалии во мрежен сообраќај во IoT, кој претставува еден од главните придонеси на оваа докторска дисертација. Најпрвин, прикажан е изборот на податочните множества и нивната поделба. Понатаму, дадени се резултатите од експериментите со машинско, длабоко и федеративно машинско учење извршени врз овие множества, како и компаративна дискусија помеѓу пристапите.

Глава 7: Седмата глава дава заклучок на сработеното, можности за проширување на постоечкото истражување, како и насоки за идни истражувања.

Глава 2 Интернет на нешта

Концептот на Интернет на нешта како таков постои веќе неколку децении, но дури во 2000-тите години IoT уредите почнуваат да стануваат се поприсутни. Притоа, како што се зголемувал бројот на IoT уреди, станало јасно дека традиционалниот протоколен пласт TCP/IP, кој е базиран на клиент/сервер архитектура и дизајниран за општо вмрежување, нема да биде доволен за да ги задоволи уникатните барања на IoT уредите.

Во 2010 година, Меѓународната телекомуникациска унија (ITU) го формира извештајот „Y.2060 Overview of the Internet of Things“ [31], кој обезбедува сеопфатна анализа на техничките барања за IoT мрежите. Извештајот ја идентификува потребата од различен протоколен пласт за поддршка на уникатните карактеристики на уреди и апликации на Интернет на нешта.

Оттогаш, развиени се повеќе протоколи специфични за Интернет на нешта, вклучувајќи ги протоколот за ограничена апликација (CoAP), телеметриски пренос на редица од пораки (MQTT) и услуга за дистрибуција на податоци (DDS). Овие протоколи се широко прифатени во индустријата и сега се сметаат за стандардни компоненти на протоколниот пласт за Интернет на нешта.

2.1 Архитектура

Архитектурниот модел на Интернет на нешта (Слика 2.1.) дефинира како овие уреди, мрежи и апликации треба да бидат структурирани за да се изгради сигурен и ефикасен систем. Во референтната литература [32-36] овој модел обично се состои од четири нивоа: ниво на уреди, мрежа, сервисни платформи и ниво на апликации. Секое ниво извршува одредена функција и работи заедно со другите за да создаде беспрекорен и скалабилен IoT систем.

2.1.1 Ниво на уреди

Нивото на уреди (ниво на перцепција) е фундаментално за архитектурниот модел на Интернет на нешта. Ова ниво се состои од сите физички уреди или сензори кои собираат податоци од околината. Во овие уреди може да се вклучат камери, сензори за температура, сензори за притисок, сензори за движење и многу други. Исто така, ова ниво може да вклучува и актуатори - уреди кои ги контролираат физичките уреди, како што се: прекинувачи, вентили, мотори итн.

Како главни карактеристики на ова ниво се можноста за перцепција - детектирање на физички промени во околината и зачувување на податоците (точно и ефикасно) во локалната меморија на уредот за понатамошно процесирање. IoT уредите имаат ограничена процесирачка моќ, па затоа процесирањето на податоците се врши локално со цел да се намали количината на податоци кои што ќе треба да бидат испратени до погорните нивоа. Во делот за процесирање се вклучуваат техники како: филтрирање (отстранување на шумот), агрегирање и анализа на собраните податоци. Поради ранливоста на уредите од потенцијални напади и неовластен пристап, на ова ниво може да се употребат и техники како шифрирање, автентикација и контрола на пристап [37,38].

2.1.2 Мрежно ниво

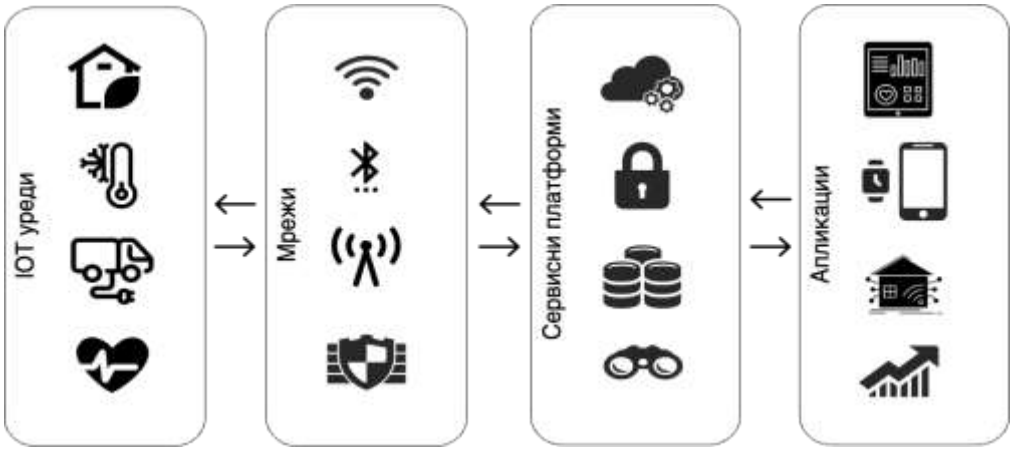
Мрежното ниво се грижи за преносот на податоци од нивото на уреди до нивото на сервисни платформи. Ова ниво се состои од уреди како што се прекинувачи, рутери (насочувачи) и точки на пристап (AP). Покрај овие уреди, нивото вклучува и протоколи кои овозможуваат комуникација помеѓу уредите и сервисните платформи. Овие протоколи се дизајнирани така да обезбедуваат сигурна комуникација помеѓу уредите и нивото на сервисни (услужни) платформи, како и доверлив пренос на податоците.

Главни карактеристики на ова ниво се рутирање (насочување) на податоците и преведување (транслација) на протоколи, како и управување со мрежната топологија и квалитет на услуга (QoS) [39,40]. Бидејќи IoT уредите може да бидат поврзани преку различни видови на жични и безжични мрежи, податоците мора да бидат насочувани до крајната дестинација со помош на физички уреди како рутери и прекинувачи. Исто така, поради различните комуникациски протоколи кои може да ги користат IoT уредите, мрежното ниво мора да биде способно да ги преведе овие протоколи со цел да овозможи непречена комуникација помеѓу различни уреди.

Под поимот мрежна топологија се подразбира поставеноста на уредите и на мрежните јазли. Мрежното ниво управува со мрежната топологија на тој начин што обезбедува оптимален проток на податоци и го минимизира мрежниот застој. Исто така, нивото може да даде приоритет на различни типови на мрежен сообраќај со цел да се осигура дека критичните податоци се испорачуваат со висока доверливост и минимално доцнење. Ова може да се изведе со користење на техники како што се обликување на сообраќајот и давање приоритет. Мрежното ниво е одговорно и за безбедноста на податоците кои се испраќаат меѓу уредите и повисоките нивоа. На ова ниво може да се воведат системи за детекција на аномалии, техники на шифрирање, како и огнени ѕидови (firewalls) [41-43].

2.1.3 Ниво на сервисни (услужни) платформи

Третото ниво на архитектурата на Интернет на нешта е наречено ниво на сервисни (услужни) платформи и е одговорно за процесирање и управување со податоците кои што се собрани од IoT уредите. Ова ниво содржи различни софтверски компоненти кои ја обезбедуваат потребната инфраструктура за обработка и анализа на податоците, како што се: брокери за пораки, складишта за податоци, системи за анализа на податоци во реално време итн. Нивото на сервисни платформи гарантира ефикасен и навремен начин на обработување и анализирање на податоците собрани од нивото на перцепција.



Слика 2.1. Архитектонскиот модел на IoT ја дефинира структурата со цел да се изгради сигурен и ефикасен систем. Овој модел обично се состои од четири слоја. Секој слој извршува специфична функција и работи заедно со другите за да создаде беспрекорни и скалабилни IoT системи.

За да се постигне горенаведеното, како главни карактеристики на ова ниво се јавуваат процесирање и анализа на податоци во облик на агрегација на податоци, филтрирање, трансформација, како и анализа со предвидување со помош на машинско учење. Покрај ова, ова ниво се занимава и со управување на уредите поврзани во IoT системот. Ова вклучува регистрирање на уредите, нивна конфигурација, софтверски надоградувања, како и следење на нивната работа и справување со можни проблеми [44].

Нивото на сервисни платформи обезбедува и интерфејси за програмирање на апликации (Application Programming Interface - API) и други интерфејси кои им овозможуваат на развивачите на софтвер да развиваат апликации над IoT системот. Освен тоа, ова ниво има можност за интеграција со други системи, како што се: системи за планирање на ресурси на претпријатието (Enterprise Resource Planning - ERP), системи за управување со односи со клиенти (Customer Relationship Management - CRM) и други IoT системи. Ова овозможува споделување на податоците помеѓу системите, обезбедувајќи поцелосен приказ на целокупниот систем и овозможувајќи понапредна анализа и донесување подобри одлуки. Како и претходните нивоа и ова ниво се грижи за безбедноста на IoT системите преку механизми како автентикација, авторизација и шифрирање, како и мониторирање на системот за безбедносни закани и превземање на активности за спречување на истите.

2.1.4 Апликациско ниво

Апликациското ниво е најгорното ниво во архитектурата на Интернет на нешта, прикажано како место на интеракција на корисниците со IoT системот. Ова ниво го дава интерфејсот преку кој корисниците може да пристапат до податоците и сервисите кои системот ги нуди. Апликациското ниво се грижи за доставување на вредноста на податоците до корисниците преку овозможување увид и давање корисни информации кои потекнуваат од податоците собрани од нивото на уреди и процесирани од нивото на сервисните платформи. Приказот се постигнува со интерфејси за визуализација на податоците како што се контролни табли, графикони и сл. Ова ниво е и лесно прилагодливо со цел да ги задоволи специфичните потреби за различните апликации за IoT.

Како друга важна карактеристика на ова ниво може да се нагласи можноста за испраќање на предупредувања и известувања. Апликациското ниво може да биде конфигурирано за да испраќа предупредувања и известувања на корисниците врз основа на претходно дефинирани правила и граници, како на пример: доколку некоја машина е пред откажување, да се испрати предупредување до техничар за одржување и сл. Ова ниво може да се искористи и за автоматизирање на уредите и процесите внатре во системот, како автоматско подесување на собна температура или светлина во зависност од податоци собрани преку нивото на перцепција. Дополнително, апликациското ниво им овозможува на корисниците да прават одлуки врз основа на собраните податоци преку прогнози и анализа на трендови [45-47].

Севкупно, слоевитата архитектура на Интернет на нешта обезбедува структуриран пристап за развој на нови и подобрување на постоечки IoT системи. Секое ниво од моделот е дизајнирано и оптимизирано за неговата специфична функција, што се пресликува во подобри перформанси и доверливост на системите. Со сегментирање на системите на IoT во различни нивоа, станува полесно да се развиваат и управуваат нови IoT системи.

2.2 Протоколи

Протоколите се збир на правила и стандарди кои ја регулираат комуникацијата помеѓу уредите и системите. Тие го дефинираат форматот и структурата на пораките, како и процесите за пренос и примање податоци. Протоколите се од суштинско значење за олеснување на комуникацијата меѓу уредите и системите, кои може да се наоѓаат на различни физички локации и да имаат различни способности за обработка и комуникација, како и за да се осигура меѓусебното комуницирање. Притоа, изборот на протокол зависи од специфичните барања на самата апликација или системот, како што се проток на податоци, доверливост, безбедност итн. Различни протоколи може да се користат во различни нивоа на протоколниот пласт, во зависност од специфичните барања на секое ниво.

Поради ограничувачките фактори на IoT уредите - мала процесирачка моќ, меморија и мал капацитет на батеријата, како и нивното распоредување во околини каде поврзаноста на мрежа е ограничена или периодична, се увидела потребата од создавање на нови протоколи кои би можеле да ги подржат уникатните барања на IoT комуникациите. Новите IoT протоколи треба да бидат во можност да ја минимизираат потрошувачката на моќ преку намалување на обемот на податоци кои се пренесуваат, како и намалување на фрекфенцијата на комуникација. Исто така, протоколите треба да бидат во можност да го оптимизираат преносот на податоци и да го намалат обемот на мрежен сообраќај. Бидејќи архитектурата на Интернет на нешта, како и сите досегашни компјутерски и мрежни архитектури е подложна на закани и напади, протоколите треба да бидат дизајнирани така што ќе обезбедат сигурна комуникација и пренос на податоци. Во продолжение на оваа глава се наоѓаат најчесто користените протоколи кои се специјализирани за Интернет на нешта.

2.2.1 ZigBee, LoRaWAN, BLE, Z-Wave и NB-IoT

ZigBee, LoRaWAN (Long Range Wide Area Network), BLE (Bluetooth Low Energy), Z-Wave и NB-IoT (Narrowband Internet of Things) се протоколи за безжична комуникација кои вообичаено се користат во мрежите од Интернет на нешта. Секој од овие протоколи има свои предности и недостатоци, а изборот на протокол зависи од специфичните барања на апликацијата или применливоста во IoT системот.

ZigBee

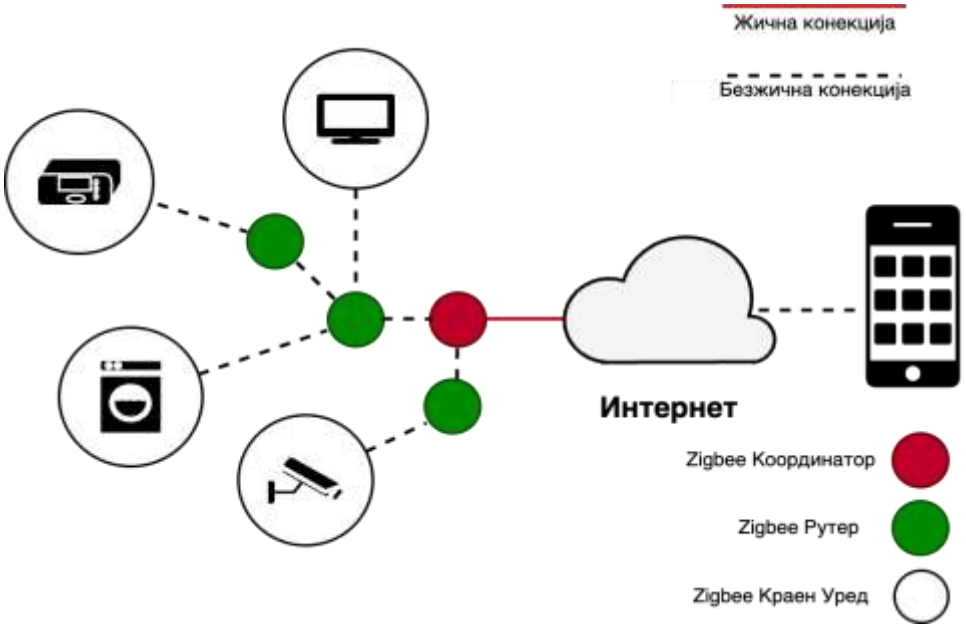
ZigBee [48] е безжичен протокол дизајниран за уреди со мала моќност и апликации со мал проток на податоци, како што се автоматизација во домот и контролни системи во индустријата. Се заснова на стандардот IEEE 802.15.4 [49] и користи испреплетена (mesh) мрежна топологија, која им овозможува на уредите да комуницираат едни со други директно или преку јазли-посредници, зголемувајќи го опсегот на мрежата и доверливоста. ZigBee поддржува три типа уреди: крајни уреди, рутери и координатори. Крајните уреди се уреди со најниска моќност кои комуницираат со мрежата преку рутер или координатор. Рутерите се задолжени за препраќање на податоците меѓу крајните уреди и координаторите. Координаторите се уреди со најголема моќност кои управуваат со мрежата и ја координираат комуникацијата помеѓу уредите (Слика 2.2.).

ZigBee работи во нелиценциран опсег од 2,4 GHz (достапен за употреба од секого без потреба од лиценца), што може да предизвика пречки со други безжични уреди кои работат во истиот опсег. Овој протокол го користи протоколот CSMA-CA (Carrier Sense Multiple Access with Collision Avoidance) за да избегне судири помеѓу преносите. Исто така, обезбедува AES шифрирање за безбедна комуникација. Најчесто, ZigBee се користи во системи за автоматизација на домот, индустриски контролни системи и безжични сензорски мрежи. Неговите предности вклучуваат: мала потрошувачка на енергија, ниска цена и можност за мрежно поврзување. Сепак, недостатоци на овој протокол се ограниченият опсег на мрежа, како и пречките од други безжични уреди.

LoRaWAN

LoRaWAN [50] е безжичен протокол дизајниран за мрежи со мала моќност кои се распространети на широка површина. Протоколот користи мрежна топологија „ѕвезда од ѕвезди“ , која им овозможува на уредите да комуницираат со една централна точка, овозможувајќи комуникација на долг домет до неколку километри. LoRaWAN работи во нелиценциран подгигахерцен опсег, кој обезбедува подобро продирање на сигналот во споредба со опсегот од 2,4 GHz што го користат ZigBee и другите протоколи. Ја користи техниката на модулација LoRa, која овозможува комуникација со мала моќност и долги растојанија, правејќи го идеален за IoT апликации кои комуницираат на долги растојанија.

Овој протокол поддржува три класи на уреди: Класа А, Класа Б и Класа С. Уредите од класа А се најчести и имаат најниска потрошувачка на енергија. Тие пренесуваат податоци кога имаат податоци за испраќање, а потоа се враќаат во режим на мирување за да заштедат енергија. Уредите од класа В и класа С имаат повеќе функции, но бараат поголема моќност. LoRaWAN најчесто се користи во апликации за паметни градови, следење на животната средина и прецизно земјоделство. Предностите на овој протокол вклучуваат комуникација на долги растојанија, мала потрошувачка на енергија и ниска цена. Сепак, неговите ограничувања вклучуваат ограничен пропусен опсег и ограничена брзина на податоци.



Слика 2.2. Протоколот ZigBee во акција. Со својата ниска потрошувачка на енергија и сигурен пренос, Zigbee им овозможува на уредите да комуницираат без напор, формирајќи еден вид на еластична мрежа. Како што податоците патуваат од јазол до јазол, протоколот Zigbee обезбедува ефикасно рутирање и способности за само-заздравување, обезбедувајќи непрекинато поврзување дури и во динамични средини.

BLE

BLE [51] е безжичен протокол кој е дизајниран за комуникација со мала моќност и кратки растојанија помеѓу уредите. Овој протокол работи во нелиценцираниот опсег од 2,4 GHz и обезбедува безбедна комуникација користејќи AES шифрирање. Ја користи техниката на модулација GFSK (Gaussian Frequency Shift Keying), која овозможува комуникација со мала моќност. BLE уредите работат во два режима: централен и периферен. Централниот уред отпочнува комуникација со периферниот уред и ја контролира врската. Периферниот уред одговара на барањата на централниот уред и испраќа податоци. Најчесто се користи во паметни уреди за носење (часовници, очила и сл.), како и системи за следење на здравјето и паметни домашни уреди. Неговите предности вклучуваат мала потрошувачка на енергија, ниска цена и комуникација со краток опсег. Како ограничувања се јавуваат ограниченият пропусен опсег, ограничената брзина на податоци и ограниченият опсег.

Z-Wave

Z-Wave [52] е безжичен протокол кој е дизајниран за автоматизација во домот и паметни домашни уреди. Работи во нелиценцираниот подгигахерцен опсег и користи испреплетена мрежна топологија, која овозможува комуникација помеѓу уредите на долги растојанија. Z-Wave уредите работат во два режима: управувач и “слуга” (slave). Управувачот започнува комуникација со слугата и ја контролира врската. Слугата одговара на барањата на управувачот и испраќа податоци. Овој протокол користи техника на модулација на распространет спектар со фреквентно скокање (FHSS), која обезбедува подобра отпорност на пречки и подобар опсег во споредба со другите протоколи. Исто така, обезбедува AES шифрирање за безбедна комуникација. Z-Wave

најчесто се користи во системи за домашна автоматизација, безбедносни системи и системи за управување со енергија. Неговите предности вклучуваат: комуникација на долги растојанија, мала потрошувачка на енергија, како и можност за мрежно поврзување. Сепак, неговите ограничувања вклучуваат ограничен пропусен опсег и ограничена брзина на податоци.

NB-IoT

Друг популарен протокол за безжична комуникација за IoT е NB-IoT [53]. NB-IoT е дизајниран за да овозможи широк спектар на уреди и услуги на IoT да се поврзат со мобилните мрежи. Работи на лиценцирани фреквентни опсези и обезбедува ниски податочни рати, долго траење на батеријата и продолжено покривање. NB-IoT е дизајниран да биде компатибилен со постоечките мобилни мрежи, што им го олеснува распоредувањето и одржувањето на операторите. Може да поддржува голем број уреди, што го прави погоден за апликации како паметни градови, индустриска автоматизација и паметно земјоделство. NB-IoT користи радио интерфејс со тесен опсег за да обезбеди сигурна комуникација со мало доцнење и мала потрошувачка на енергија. Исто така, поддржува двонасочна комуникација, овозможувајќи им на уредите да испраќаат и примаат податоци од мрежата. Овој протокол обезбедува и шифрирање и автентикација со цел да се обезбеди сигурна комуникација помеѓу уредите и мрежата. Недостатоците на овој протокол се ограничени податочни рати, ограничена флексибилност, повисока цена на имплементација, како и ограничувања во однос на покриеност.

Во Табела 2.1. е направена сумаризација на претходно разработените протоколи, нивниот фрекфенциски опсег, податочна рата, домет, како и предностите и недостатоците на секој поединечно.

Табела 2.1. Карактеристики, предности и недостатоци на ZigBee, LoRaWAN, BLE, Z-Wave и NB-IoT

Протокол	**Фрекфенциски опсег**	**Податочна рата**	**Домет**	**Предности**	**Недостатоци**
ZigBee [48,49]	2.4 GHz	250 Kbps	10 - 100м	ниска потрошувачка на енергија, ниска цена, надежен, безбеден, поддржува големи мрежи	Ограничен пропусен опсег, проблеми со интероперабилноста
LoRaWAN [50]	915 MHz (US), 868 MHz (EU)	до 27 Kbps	до 10 км	Долг дострел, мала потрошувачка на енергија, ниска цена, поддржува големи мрежи	Ограничен пропусен опсег, ограничена брзина на податоци, доцнење во преносот
BLE [51]	2.4 GHz	до 2 Mbps	до 100м	Ниска потрошувачка на енергија, ниска цена, лесен за користење, поддржува големи мрежи, интероперабилност со паметни телефони и други уреди	Ограничен пропусен опсег, подложност на пречки
Z-Wave [52]	908 MHz (US), 868 MHz (EU)	до 100 Kbps	до 100м	Ниска потрошувачка на енергија, безбеден, надежен, поддржува големи мрежи интероперабилност со паметни домашни уреди	Ограничен пропусен опсег, проблеми со интероперабилноста
NB-IoT [53]	700 MHz (US), 800 MHZ (EU)	до 250 Kbps	до 10км	Широка покриеност, висока податочна рата, сигурен, поддржува големи мрежи	Ограничено траење на батеријата, висока цена, ограничена флексибилност

2.2.2 6LoWPAN и RPL

6LoWPAN

6LoWPAN [54] е протокол кој дефинира како IPv6 пакетите може да се пренесат преку безжични мрежи со мала моќност со ограничен опсег и енергија, притоа минимизирајќи ја потрошувачката на енергија на IoT уредите. Ова се постигнува со компресија на заглавјата на IPv6, така што пакетите можат да се вклопат во малите димензии на рамката на безжични мрежи со мала моќност, при што помага за заштедување на пропусниот опсег. Исто така, овој протокол им овозможува на IoT уредите да формираат мрежа и директно да комуницираат едни со други, без да се потпираат на единствена централна точка. Ова го прави 6LoWPAN скалирачки протокол, бидејќи може да се додадат повеќе уреди на мрежата без да се предизвика метеж. 6LoWPAN е широко искористен во многу апликации за IoT, вклучувајќи: автоматизација на домот, индустриска автоматизација, паметни градови и паметно здравство и е од суштинско значење за успехот на IoT апликациите.

RPL

RPL [55] е рутирачки протокол специјално дизајниран за употреба во мрежи што се состојат од голем број наизменично поврзани уреди со ниска моќност и високи стапки на загуба на пакети, каде што традиционалните протоколи за рутирање, како што се оние што се користат на Интернет, можеби и не се соодветни поради ограничените ресурси и непредвидливата природа на мрежата. RPL е проактивен протокол, што значи дека гради и одржува рутирачки табели однапред, наместо да чека додека пакетот треба да се пренасочи до дестинацијата. Овој протокол поддржува повеќе функции, кои се користат за пресметување на оптималната патека за пренос на пакети врз основа на различни критериуми, како што се минимизирање на потрошувачката на енергија или минимизирање на доцнењето од крај до крај. Ова овозможува RPL да се оптимизира за различни типови на апликации и мрежни услови. Една клучна предност на RPL е неговата скалабилност - способност да поддржува мрежи од големи размери со голем број јазли. RPL е широко користен во многу апликации за IoT, како што се паметни домови, индустриска автоматизација и паметно здравство.

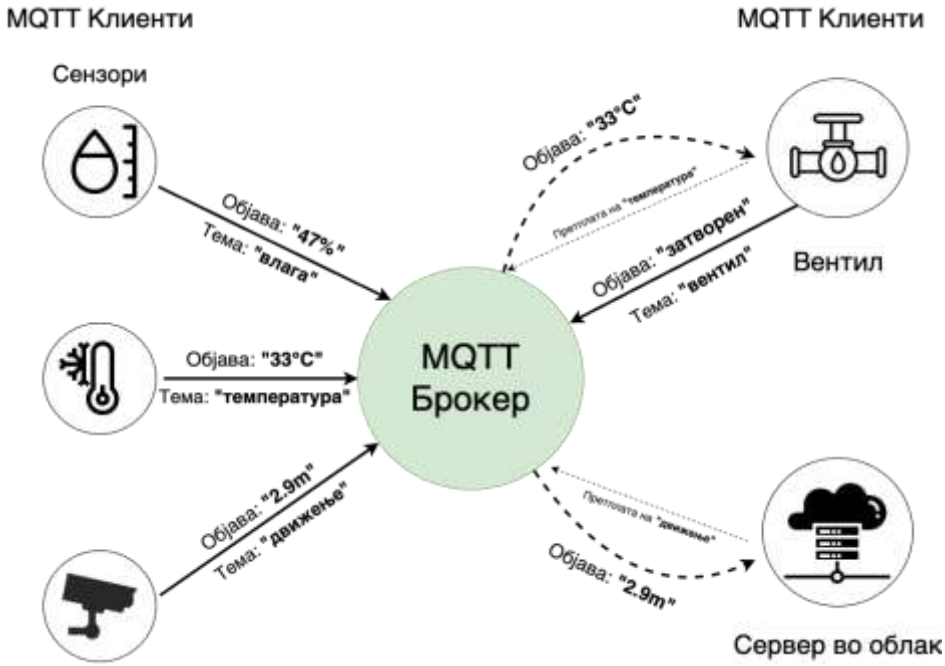
2.2.3 MQTT, CoAP, AMQP и DDS

MQTT

MQTT [56] е протокол со отворен стандард дизајниран за уреди со ограничени ресурси кој е широко користен во IoT поради неговата едноставност, приспособливост и доверливост. Протоколот користи шаблон за испраќање пораки за објава (publish) и претплата (subscribe), каде што клиентите се претплатуваат на тема (topic) и добиваат пораки објавени на таа тема. Брокерот за пораки е централна точка кој прима пораки објавени од издавачите и ги доставува до претплатниците. MQTT е широко користен во IoT апликации како што се: паметни домови, индустриска автоматизација и далечинско мониторирање.

MQTT се состои од три компоненти (прикажани на Слика 2.3.):
Клиенти: уреди кои објавуваат пораки или вршат претплата на одредени теми. Клиентите можат да бидат и објавувачи и претплатници.
Брокер: сервер кој прима и насочува пораки меѓу MQTT клиентите.
Теми: хиерархиска структура на која се објавуваат пораките и на која може да се изврши претплата.

Протоколот поддржува различни нивоа на QoS за да обезбеди сигурност на пораките:
QoS ниво 0 (најмногу еднаш): пораката се доставува најмногу еднаш и не се испраќа потврда за испорака до издавачот.
QoS ниво 1 (барем еднаш): Пораката се гарантира дека ќе биде испорачана барем еднаш, а потврда за испорака се испраќа до издавачот.
QoS ниво 2 (точно еднаш): Се гарантира дека пораката ќе биде доставена точно еднаш, а двонасочното ракување се користи за да се потврди испораката.



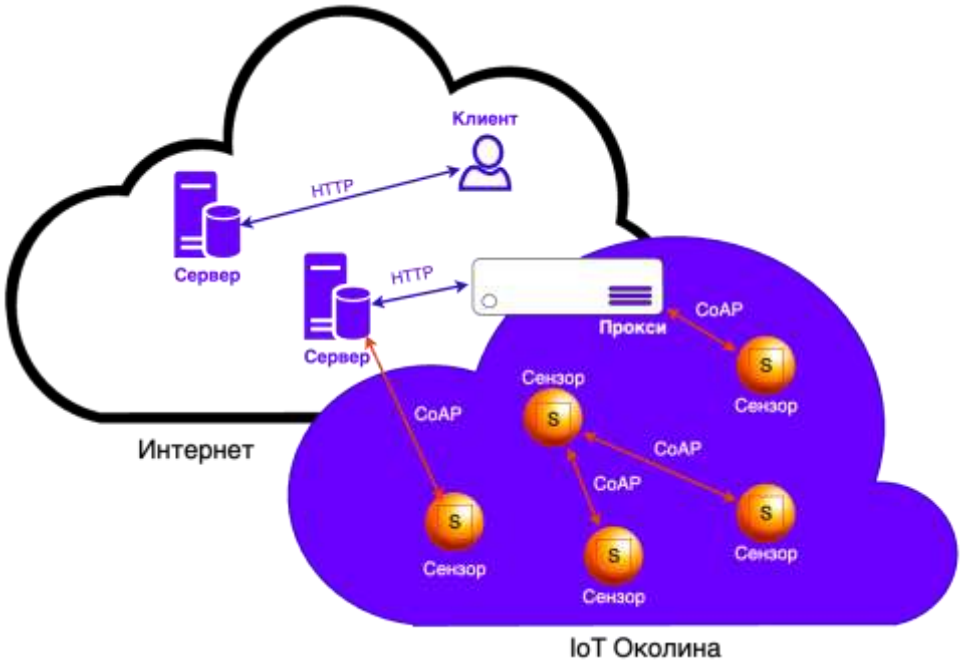
Слика 2.3. MQTT протокол во акција. MQTT делува како оркестратор, олеснувајќи го ефикасното испраќање пораки помеѓу издавачите и претплатниците. Пораките се објавуваат на одредени теми, а претплатниците со соодветни интереси ги добиваат тие пораки веднаш. Со малиот пропусен опсег и барањата за мала енергија, MQTT овозможува размена на податоци во реално време во средини со ограничени ресурси.

CoAP

CoAP [57] е апликациски протокол дизајниран да работи на уреди со ограничени ресурси, како што се микроконтролери и да поддржува безжични мрежи со мала моќност. Овој протокол користи клиент/сервер архитектура каде што клиентите испраќаат барања до серверите за превземање или менување на ресурсите. До ресурсите се пристапува со користење на стандардни HTTP методи, како што се GET, POST, PUT и DELETE. CoAP наоѓа примена во паметните домови, индустриската автоматизација, како и паметното здравство.

CoAP има четири главни компоненти (прикажани на Слика 2.4.):
Клиенти: уреди кои испраќаат барања до сервери за превземање или менување ресурси.
Сервери: уреди што ги хостираат ресурсите и одговараат на барањата од клиентите.
Прокси: уреди кои дејствуваат како посредници помеѓу клиентите и серверите.
Ресурси: податоци или услуги кои се достапни преку CoAP.

CoAP го користи UDP како протокол на транспортно ниво и поддржува четири типа пораки:
Потврдено (Confirmable - CON): порака која бара одговор од серверот.
Непотврдено (Non-Confirmable - NON): порака за која не е потребен одговор од серверот.
Потврда (Acknowledgement - ACK): Порака што го потврдува приемот на порака што може да се потврди.
Ресетирање (Reset - RST): Порака што покажува дека барањето не може да се обработи.



Слика 2.4. CoAP протоколот во акција. CoAP служи како столб кој им овозможува на уредите да комуницираат користејќи ја RESTful архитектурата. Со својата едноставност и минимални трошоци, CoAP го оптимизира искористувањето на ресурсите, што го прави добро прилагоден за ограничени мрежи и уреди со мала моќност. Поддржува интеракции со барање-одговор, како и механизми за објавување-претплати, овозможувајќи флексибилни комуникациски обрасци.

AMQP

AMQP [58] е протокол за пораки дизајниран да овозможи доверливо, безбедно, асинхроно и интероперабилно испраќање пораки помеѓу различни системи и апликации. Овозможува шеми за размена на пораки како што се точка до точка, барање/одговор и објавување/претплата. Една од клучните карактеристики на AMQP е неговата способност да обезбеди гаранции за квалитет на услугата дури и во предизвикувачки мрежни услови. Исто така, поддржува поставки за приоритет на пораката и време за живот (time to live), што овозможува поголема контрола врз нивната испорака. AMQP го одредува форматот и структурата на пораките на ниво на бајт. Ова овозможува интероперабилност помеѓу различни платформи и програмски јазици, се додека тие ја имплементираат истата верзија на AMQP. Овој протокол користи разменувачи и редици за насочување на пораките помеѓу крајните точки. Производителите испраќаат пораки до разменувачите, кои потоа ја насочуваат пораката до соодветната редица/и врз основа на правилата за рутирање. Потрошувачите треба да се претплатат на соодветната редица за да примаат пораки.

DDS

DDS [59] е протокол кој обезбедува стандардизиран начин за дистрибуираните апликации да разменат податоци во реално време. DDS е дизајниран да обезбеди високи перформанси, доверливост и приспособливост за дистрибуирани системи. Користи модел за објавување-претплата за размена на податоци, каде што издавачите испраќаат податоци до брокер, а претплатниците ги добиваат податоците за кои се заинтересирани од брокерот. DDS поддржува обрасци за комуникација еден-на-многу и многу-на-многу, што го прави погоден за широк опсег на дистрибуирани системи. DDS користи пристап фокусиран на податоци, што значи дека фокусот е на самите податоци, а не на крајните јазли. Овој пристап го олеснува додавањето или отстранувањето на крајните јазли од системот без да се влијае на целокупното работење. DDS поддржува и широк опсег на типови податоци, вклучувајќи структурирани податоци, неструктурирани податоци и мултимедијални податоци.

Овој протокол обезбедува голем број функции што го прават погоден за употреба во IoT апликации. Една од клучните карактеристики е неговата способност да ракува со големи количини на податоци во реално време. Ова се постигнува преку употреба на ефикасни техники за компресија на податоци и оптимизирана мрежна комуникација. Исто така, DDS обезбедува QoS, со што им овозможува на апликациите да ги специфицираат барањата за доверливост и навременост за размена на податоци. Ова е особено важно во апликациите за IoT, каде што податоците треба да се доставуваат веродостојно и во одреден временски период. Друга клучна карактеристика на DDS е неговата поддршка за динамично откривање и конфигурација. Ова значи дека крајните точки може да се додадат или отстранат од системот без рачна конфигурација. DDS, исто така, поддржува автоматско откривање на типови на податоци и нивните поврзани шеми, што го олеснува интегрирањето на новите извори на податоци во системот. DDS се користи во широк опсег на апликации, вклучувајќи системи за индустриска контрола, воени системи и системи за здравствена заштита. Неговата употреба во апликациите за IoT расте, особено во областите на паметни градови, управување со енергија и транспортни системи.

Глава 3 Безбедност во Интернет на нешта

3.1 Предизвици и барања во поглед на безбедноста

Безбедноста на IoT уредите е од суштинско значење за да им се овозможи доверба на поединците и организациите при користење на овие уреди, без грижа дека истите се ранливи на напади или дека податоците ќе им бидат пробиени. Иако придобивките од ваквите мрежи се многубројни, вклучувајќи зголемена ефикасност и практичност, важно е да се нагласат потенцијалните ризици поврзани со нивната употреба. Со цел да се ублажат овие ризици, важно е да се даде приоритет на безбедноста во дизајнот и имплементацијата на IoT уредите. Предизвиците [60] со кои секојдневно се соочуваат овие мрежи се:

Сложеност на мрежата: Навигацијата низ сложеноста на IoT мрежите претставува значителен предизвик поради големата разновидност на уреди и протоколи. Надгледувањето и безбедноста на овие мрежи стануваат напорни задачи бидејќи големата количина на податоци генерирани од IoT уредите придонесува откривањето и навременото справување со потенцијалните безбедносни закани да е сè поголем предизвик.

Недостаток на вградени безбедносни карактеристики: Многу IoT уреди се дизајнирани имајќи ја предвид функционалноста и леснотијата на користење, наместо безбедноста. Ова може да ги направи ранливи на широк опсег на напади, вклучувајќи злонамерен софтвер, DDoS (Distributed Denial of Service) напади итн. За да се одговори на овој предизвик, производителите на IoT уреди треба да дадат приоритет на безбедноста при самиот процес на дизајнирање. Ова може да вклучува имплементирање механизми за безбедно подигање и ажурирање на фирмверот (firmware), шифрирање на чувствителни податоци и вклучување на безбедни протоколи за автентикација. Дополнително, производителите треба редовно да ги ажурираат фирмверот и софтверот за да ги решат пропустите.

Приватност на податоците: IoT уредите дејствуваат како чувари на широк опсег на чувствителни податоци, вклучувајќи лични информации и податоци за локација, кои често се пренесуваат преку Интернет. Сепак, овие податоци стануваат подложни на пресретнување и експлоатација, што претставува значителен ризик, како кражба на идентитет или корпоративна шпионажа. Решавањето на овој критичен предизвик бара проактивен пристап во дизајнирањето на IoT уреди, со главен стремеж кон приватноста на податоците. Мерките како што се минимизирање на собирањето лични податоци, имплементирање робусни механизми со енкрипција за заштита на чувствителните информации и усвојување безбедни протоколи за комуникација мора да бидат вградени во архитектурата на уредот. Подеднакво важно е корисниците да имаат контрола врз нивните податоци, давајќи им овластување да ги бришат истите и да ја одредат нивната употреба, со што поединците стојат во првите редови за заштита на сопствената приватност.

Потешкотии со поправање на пропусти: За разлика од традиционалните софтверски апликации, IoT уредите може да се наоѓаат на тешко достапни локации, што го отежнува ажурирањето на нивниот фирмвер или софтвер. Покрај тоа, многу IoT уреди немаат процесорска моќ или меморија за да се справат со сложени безбедносни ажурирања. За

да се справат со овој предизвик, производителите на IoT уреди треба да дизајнираат уреди чии фирмвер и софтвер лесно се ажурираат. Ова може да вклучува имплементирање механизми за далечинско ажурирање и обезбедување дека уредите имаат доволно процесорска моќ и меморија за да се справат со ажурирањата. Покрај тоа, производителите исто така треба да обезбедуваат редовни ажурирања за да се решат познатите пропусти.

Недостаток на свест на самите корисници: Многу корисници можеби и не се свесни за безбедносните ризици поврзани со IoT уредите или можеби не знаат како да спроведат безбедносни мерки за да се заштитат. Ова може да ги направи ранливи на напади и навлегување во податоците. За да се справат со овој предизвик, производителите на IoT уреди треба да дадат приоритет на едукацијата и свесноста на корисниците. Ова може да вклучува обезбедување јасни упатства за обезбедување на уредите, како и совети за идентификување и избегнување на напади. Покрај тоа, производителите исто така можат да обезбедуваат редовни безбедносни ажурирања и известувања за корисниците.

Недостаток на IoT стандарди: Широкиот опсег на уреди и протоколи што се користат во IoT мрежите може да го отежнат воспоставувањето на заеднички безбедносен стандард. Ова може да им ја отежне работата на организациите за да имплементираат сеопфатни безбедносни мерки на сите нивни IoT уреди.

Исто така, IoT уредите треба да задоволуваат специфични безбедносни стандарди кои гарантираат за нивната безбедност, сигурност и доверливост. IoT стандардите за безбедност како што се ISO/IEC 30141:2018[172], ISO/IEC 27400:2022 series[173], NIST IoT Security Guidelines[174] and ENISA Baseline Security Recommendations for IoT[175] ги поставуваат критериумите за безбедност на IoT уредите. Овие стандарди покриваат различни аспекти, вклучувајќи принципи за безбедносен дизајн, енкрипција, автентикација, контрола на пристап, безбедни комуникациски протоколи, механизми за безбедно ажурирање и приватност на податоците. Со почитување на овие воспоставени стандарди, IoT уредите можат да ги исполнат потребните безбедносни критериуми [61], ублажувајќи ги потенцијалните ризици и обезбедувајќи доверливо и сигурно работење во рамките на екосистемите на IoT.

Доверливост: ова барање кажува дека пораките што се разменуваат помеѓу IoT уредите се достапни само за страните за кои се наменети. Доверливоста може да се постигне преку шифрирање, што гарантира дека пораките може да се читаат само од овластени страни кои го имаат клучот за дешифрирање. Дополнително, може да се користат и контроли за пристап и механизми за автентикација на корисникот за да се осигура дека само овластени корисници можат да пристапат до податоците генерирани од IoT уредите.

Интегритет: ова барање кажува дека пораките што се разменуваат помеѓу IoT уредите не се манипулирани. Тоа може да се постигне преку употреба на дигитални потписи, кои ја потврдуваат автентичноста на пораките и даваат гаранција дека тие не се изменети при транспортот. Кодовите за автентикација на пораки исто така може да се користат за да се обезбеди интегритет на пораките.

Автентикација: ова барање кажува дека страните што комуницираат се оние што тврдат дека се и со ова се спречуваат напади на имитирање. Автентикацијата може да се постигне преку употреба на дигитални сертификати, кои го потврдуваат идентитетот на страните што комуницираат. Покрај тоа, механизмите за автентикација со повеќе фактори, како што се употребата на биометриски податоци исто така може да се користат за да се обезбеди автентичноста на корисниците.

Достапност: ова барање кажува дека има услугата е непрекината и спречува напади на одбивање на услугата (DoS) што може да предизвикаат прекин на истата. Достапноста може да се постигне преку употреба на механизми кои осигуруваат дека постојат резервни системи во случај на дефект. Механизмите за балансирање на оптоварувањето може да се користат и за дистрибуција на сообраќај низ повеќе сервери и спречување на преоптоварување на поединечни сервери.

Овластување: ова барање кажува дека ентитетите имаат дозвола за извршување на доделените операции. Овластувањето може да се постигне преку употреба на листи за контрола на пристап, кои ги специфицираат дозволите што секој ентитет ги има за извршување на специфични операции. Контрола на пристап заснована на улоги може да се користи и за доделување дозволи врз основа на улогите и одговорностите на поединечните корисници.

Свежина на пораки: ова барање кажува дека старите пораки не можат да се репродуцираат, со што се спречуваат нападите за повторување. Свежината на пораки може да се постигне преку употреба на механизми за временски печати (timestamps), кои додаваат временски печат на секоја порака за да се осигураат дека таа е единствена и не била претходно користена.

Неодбивање: ова барање кажува дека субјектите не можат да ги отфрлат дејствата што ги извршиле. Неодбивањето може да се постигне преку употреба на дигитални потписи, кои обезбедуваат доказ за автентичноста и интегритетот на пораките и гарантираат дека испраќачот не може да негира дека ја испратил пораката.

Тајност по напуштање на мрежа (forward secrecy): Ова барање кажува дека комуникациите што се случуваат по заминувањето на јазолот не се препознаваат од истиот. Ова може да се постигне преку генерирање на нов клуч за секоја сесија, така што се гарантира дека безбедноста на другите сесии нема да биде загрозена иако сесискиот клуч е компромитиран.

Тајност при вклучување во мрежа (backward secrecy): ова барање кажува дека новите објекти кои се приклучуваат на мрежата немаат информации на претходните комуникации. Ова може да се постигне преку употреба на механизми за ротација на клучеви, кои гарантираат дека клучевите периодично се менуваат за да се спречат напаѓачите да добијат пристап до претходни комуникации.

3.2 Напади во Интернет на нешта

3.2.1 Напади на ниво на уреди

Физички напади

Физичките напади [62] на IoT уредите вклучуваат различни злонамерни активности како физичко манипулирање или компромитирање на уредите или нивната физичка околина. При овие напади се искористуваат пропустите во физичките аспекти на IoT системите за да се добие неовластен пристап, контрола или да се наруши нивното нормално функционирање. Физичките напади можат да бидат еден од следните видови:

Нападите на обратно инженерство (reverse engineering) вклучуваат процес на расклопување на уред или IoT сензор за да се добие увид во неговата внатрешна структура, функционирање и содржани информации. Овој метод им овозможува на напаѓачите да ги анализираат фирмверот, софтверот, хардверските компоненти и протоколите за комуникација на уредот. Со обратно инженерство на IoT уред, напаѓачот има за цел да ја разбере неговата внатрешна работа, да ги идентификува потенцијалните пропусти и да ги искористи за злонамерни цели.

Нападите со вбризгување се вид на физички напад во кој напаѓачот вбризгува (инјектира) малициозен софтвер или друг злонамерен код во еден IoT уред или сензор преку физичка порта или врска. Овој метод му овозможува на напаѓачот да ги искористи пропустите во влезните механизми на уредот и да ја загрози неговата безбедност и функционалност. Нападите со вбризгување може да се случат преку различни физички порти или конекции, вклучувајќи: USB конекции, мрежни конекции и сериски или UART конекции.

Напади на страничен канал

Нападите на страничен канал (side channel attacks) [63] ги искористуваат физичките карактеристики на IoT уредот или сензорот за да извлечат чувствителни информации. Овие напади го користат непланираниот проток на информации што се случува за време на работата на уредот, наместо да ги таргетираат неговите логички безбедносни мерки. Нападите на страничните канали може да вклучуваат:

Анализа на моќност е тип на напад кој се фокусира на анализирање на моделите за потрошувачка на енергија на уредот или сензорот за извлекување чувствителни податоци, како што се криптографски клучеви. Со внимателно следење на флуктуациите во моќноста и нивото на потрошувачка на IoT уредот за време на неговото работење, напаѓачот може да добие увид во внатрешните процеси и пресметки што се случуваат во самиот уред. Потрошувачката на енергија на еден IoT уред може да открие вредни информации за неговите операции и може да се користи за да се извлечат чувствителни податоци.

Анализа на времето е напад кој се фокусира на анализирање на временските шеми и варијациите на операциите извршени од еден IoT уред или сензор за да се извлечат чувствителни податоци. Со внимателно следење на прецизното време на операциите, напаѓачот може да добие увид во внатрешните процеси, контролниот тек или обрасците на податоци во уредот. Нападите за анализа на времето го користат фактот дека

различните операции или влезови може да резултираат со варијации во времето на извршување на операциите на IoT уредот. Со внимателно мерење и анализа на овие временски шеми, напаѓачот може потенцијално да заклучи чувствителни информации, вклучувајќи криптографски клучеви, влезни вредности или дури и внатрешна логика на уредот.

Напади со манипулирање на софтверот/фирмверот

Нападите за манипулација на софтвер/фирмвер [64] се малициозни активности чија цел е менување на софтверот или фирмверот на IoT уред или сензор, со цел да се добие неовластен пристап или контрола врз него. Овие напади ги искористуваат пропустите присутни во софтверот или фирмверот на IoT уредите, што на крајот ја загрозува нивната безбедност и функционалност. Напаѓачите често се фокусираат на специфични слабости за да ги постигнат своите цели. Притоа, примарната цел се врти околу стекнување неовластен пристап до IoT уредот. Со модифицирање на софтверот или фирмверот на уредот, напаѓачите се обидуваат да ги заобиколат механизмите за автентикација, да ја ослабат постоечката безбедносна контрола или да воведат задни врати што им обезбедуваат привилегиран пристап. Откако ќе се добие неовластен пристап, напаѓачите можат да извлечат чувствителни податоци, да манипулираат со операциите на уредот или да го искористат компромитираниот уред за да извршат дополнителни напади во IoT екосистемот. Нападите за манипулирање може да вклучуваат:

Ажурирањата на фирмверот играат клучна улога во одржувањето на безбедноста и функционалноста на IoT уредите. Меѓутоа, во контекст на напади со манипулација на софтвер/фирмвер, напаѓачите може да инсталираат злонамерно ажурирање на фирмверот на IoT уред или сензор. Со тоа тие имаат за цел да стекнат неовластена контрола над уредот или да украдат чувствителни податоци. При извршување на напад за ажурирање на фирмверот, напаѓачот може да употреби различни техники за да го загрози интегритетот на процесот на ажурирање, како замена на самиот фирмвер или phishing и употреба на социјално инженерство.

Вбризгување на код е техника што ја користат напаѓачите за да вметнат злонамерен код во IoT уред или сензор со цел да се добие контрола над уредот или да се украдат чувствителни податоци. Овој тип на напад цели кон ранливости во слојот на софтверот или фирмверот, дозволувајќи му на напаѓачот да манипулира со однесувањето на уредот за злонамерни цели. При извршување на напад со вбризгување код, напаѓачот ги користи безбедносните слабости во софтверот или фирмверот на уредот за да го инјектира и изврши сопствен код.

Напади на сензори

Нападите на сензори [65] се однесуваат на малициозни активности кои ги искористуваат ранливостите во сензорите или поврзаниот софтвер со намера да добијат неовластен пристап или контрола врз IoT системот. Овие напади се фокусираат на компромитирање на компонентите одговорни за собирање и пренос на податоци од физичката средина до IoT уредот или мрежата. При напад на сензори, напаѓачите ги идентификуваат слабостите на сензорите или нивниот софтвер за извршување на нивните злонамерни цели. Со искористување на овие пропусти, тие се здобиваат со способност да манипулираат или да ги нарушат податоците од сензорот, што може да доведе до тешки последици во рамките на IoT системот. Нападите на сензорите може да вклучуваат:

Лажирање (измама) на сензори е техника што ја користат напаѓачите за да манипулираат со читањата произведени од сензорот, што доведува до генерирање лажни податоци или влијае на работата на целиот IoT систем. Со измама на сензори, напаѓачите можат да воведат неточности, погрешни информации или дури и целосни измислици, што може да има значителни последици врз функционалноста и доверливоста на IoT системот. Кога напаѓачите се впуштаат во напади на измама на сензори, тие имаат за цел да го измамат IoT системот со менување на податоците обезбедени од сензорот.

Заглавување на сигналот е злонамерна техника што ја користат напаѓачите за да го нарушат правилното функционирање на сензорот со намерно мешање во неговите сигнали. Овие пречки може да доведат до неможност на сензорот да работи правилно или да произведе лажни отчитувања. Со заглавување на сигналите, напаѓачите имаат за цел да ја загрозат точноста, доверливоста и севкупните перформанси на сензорот и поврзаниот IoT систем.

Прислушување на сензори (Слика 3.1.) е специфичен тип на сензорски напад каде напаѓачот ја пресретнува комуникацијата помеѓу IoT сензорите и мрежата или други уреди. Овој напад има за цел да собере информации пренесени од сензорите за неовластен пристап, кражба на чувствителни податоци или манипулација со податоците од сензорите за злонамерни цели. Во напад на прислушување со сензор, напаѓачот тајно ја следи комуникацијата помеѓу IoT сензорите и мрежната инфраструктура. Со пресретнување на податоците што ги пренесуваат сензорите, напаѓачот добива пристап до вредни информации за работата на IoT системот, отчитувањата на сензорите и потенцијално чувствителните податоци што се разменуваат.

Напади на напојување

Нападите на напојувањето [66] се класа на напади кои се насочени кон ранливости во напојувањето или управувањето со енергијата на уред или сензор. Овие напади ги искористуваат недостатоците во енергетската инфраструктура за да добијат неовластен пристап или контрола врз целниот систем. Со манипулирање или компромитирање на напојувањето, напаѓачите можат да ги постигнат своите злонамерни цели. Нападите со моќност опфаќаат различни техники кои ги искористуваат слабостите во аспектите поврзани со моќноста на уредот или сензорот. Напаѓачите ги користат овие недостатоци за да ја загрозат безбедноста на системот. Овие напади може да вклучуваат:

Анализа на енергијата е техника што ја користат напаѓачите за да ги анализираат моделите за потрошувачка на енергија на уред или сензор за IoT со цел да извлечат чувствителни податоци. Со внимателно следење на флуктуациите и варијациите на напојувањето за време на работата на уредот, напаѓачот може да добие увид во внатрешните процеси, криптографските операции или чувствителните информации што ги обработува уредот. Напаѓачите користат анализа на моќноста за да го искористат истекувањето на информациите што се случува преку варијации во потрошувачката на енергија. Со внимателно набљудување и анализа на овие модели на моќ, тие можат да извлечат вредни сознанија и потенцијално да извлечат чувствителни податоци.

Пречки во напојувањето е злонамерна техника што ја користат напаѓачите за да создадат прекини во напојувањето или флуктуации со намера да го нарушат нормалното функционирање на IoT уред или сензор. Со манипулирање со напојувањето или со воведување нарушувања во електричниот систем, напаѓачите можат да ја загрозат функционалноста, доверливоста и достапноста на насоченият IoT уред или сензор.

Напаѓачите може да користат различни методи за да предизвикаат прекини во напојувањето или флуктуации на IoT уред или сензор, како што се прекини на струја, скокови и падови на напонот, нарушувања на фреквенцијата или бучава од напојувањето.

Напади со манипулација со податоци со сензори

Нападите со манипулација со сензорски податоци [67] се однесуваат на малициозни активности насочени кон менување на податоците генерирани од сензорите во рамките на IoT системот. Примарната цел на овие напади е да произведат лажни отчитувања или да влијаат на работата на системот за измамнички или злонамерни цели. Со манипулирање со податоците од сензорите, напаѓачите можат да го загрозат интегритетот, доверливоста и функционалноста на IoT системот. Еве неколку такви подтипови на напади:

Нападите на температурниот сензор вклучуваат намерна манипулација со температурен сензор во рамките на IoT систем со цел да се генерираат лажни отчитувања кои можат значително да влијаат на работата на системот. Напаѓачите ги искористуваат пропустите во хардверот, софтверот или механизмите за калибрација на сензорот за температура за да го измамат IoT системот, што потенцијално ќе доведе до штетни последици. Во нападите на сензорот за температура, напаѓачите користат различни техники за да манипулираат со читањата на сензорот и да влијаат на однесувањето на системот, како што се лажни отчитувања на температурата, манипулација со прагот или манипулација со околината.

Нападите на сензорот за движење вклучуваат манипулација со сензор за движење во рамките на IoT систем за да произведе лажни аларми или да ја попречат способноста на сензорот прецизно да детектира движење. Напаѓачите ги искористуваат пропустите во хардверот, софтверот или факторите на околината на сензорот за движење за да го измамат IoT системот и да ги постигнат своите злонамерни цели. Некои од техниките што се користат вклучуваат: генерирање лажен аларм, заглавување на сензорот за движење, манипулација со околината или манипулација со калибрација на сензорот.

Нападите на сензорите за гас вклучуваат манипулација со сензор за гас во системот IoT за да се генерираат лажни отчитувања, што доведува до лошо однесување на системот или ја попречува способноста на сензорот да открие истекување на гас. Напаѓачите ги искористуваат пропустите во хардверот, софтверот или факторите на животната средина на сензорот за гас за да го измамат IoT системот, потенцијално предизвикувајќи значителни безбедносни ризици или нарушувања. Некои од техниките што се користат вклучуваат: лажни отчитувања на гас, заглавување на сензорот за гас, манипулација со калибрација на сензорот или манипулација со околината.

Слика 3.1. Демонстрација на напад на прислушкување на сензор. Злонамерната страна вешто ја пресретнува и дешифрира комуникацијата помеѓу сензорите, искористувајќи ги слабостите во нивните безбедносни протоколи. Со прецизност, напаѓачот добива неовластен пристап до чувствителните податоци што ги пренесуваат сензорите, со што се нарушува доверливоста и интегритетот на мрежата.

3.2.2 Напади на мрежно ниво

Напади на одбивање на услуга

Нападите на одбивање на услугата [68] (Слика 3.2.) се вообичаен тип на напад што може да се случи и во IoT мрежите. При DoS напад, напаѓачот ја преплавува мрежата со сообраќај, преоптоварувајќи ја и предизвикувајќи таа да стане недостапна за корисниците. Ова може да биде особено штетно во IoT, бидејќи овие мрежи често се потпираат на комуникација во реално време за да донесат временски чувствителни одлуки. Постојат неколку различни типови на DoS напади кои можат да се случат во IoT:

Нападите базирани на пропусен опсег се малициозни активности кои имаат за цел да ја нарушат мрежата со тоа што ќе ја преплават со огромен обем на сообраќај, надминувајќи го достапниот пропусен опсег и правејќи го недостапен за легитимните корисници. Напаѓачите ги користат пропустите во мрежната инфраструктура за да ги иницираат овие напади, што доведува до значителни прекини на услугите и го попречува нормалното функционирање на целната мрежа.

Нападите базирани на протокол се однесуваат на малициозни активности кои ги искористуваат пропустите во комуникациските протоколи што ги користат IoT уредите. Овие напади имаат за цел да го нарушат нормалното функционирање на IoT уредите преку искористување на слабостите во имплементацијата на протоколот, што доведува до исклучување или неактивност на уредот. На пример, напаѓачите може да ги користат пропустите за прелевање на баферот во ракувањето со протоколот со пакетите со податоци, предизвикувајќи уредот да падне или да не реагира или да внесат неочекуван податок во протоколниот пласт, а потоа да ги користат програмските

грешки или несоодветното ракување со грешки, компромитирајќи ја стабилноста на уредот.

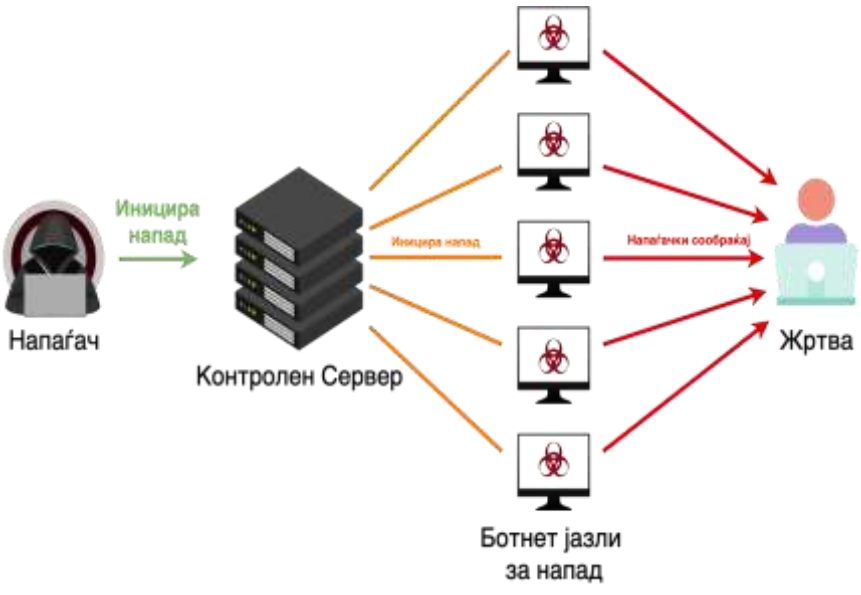
Нападите базирани на ресурси вклучуваат малициозни активности кои ги таргетираат ресурсите на IoT уредите, како што се централната процесорска единица или меморијата со намера да ги исцрпат. Овие напади имаат за цел да го надминат капацитетот на уредот, што резултира со неодговорност, деградирани перформанси или дури и целосен дефект на системот.

Напади со човек во средината

Нападите со човек во средината (MitM) [69] (Слика 3.3.) се уште еден вообичаен тип на напад кој може да се случи и во IoT мрежите. Во MitM, напаѓачот пресретнува комуникација помеѓу учесниците во комуникацијата (уреди, сервери, итн.) и манипулира или краде податоци. Ова може да направи многу штета во IoT мрежите, бидејќи податоците собрани од уреди често се чувствителни и вклучуваат лични информации. Постојат неколку различни типови на MitM напади кои можат да се случат во такви мрежи:

Нападите за пресретнување се малициозни активности кои вклучуваат прислушување на комуникацијата помеѓу IoT уредите и пресретнување на податоците што се пренесуваат. Напаѓачите ги искористуваат пропустите во комуникациските канали, протоколите или механизмите за шифрирање за да добијат неовластен пристап до чувствителни информации или податоци разменети помеѓу IoT уредите. Нападите од пресретнување претставуваат значителни ризици за безбедноста и приватноста на IoT системите и податоците со кои ракуваат.

Напади со измена на комуникациите се однесува на малициозни активности кои вклучуваат промена на комуникацијата помеѓу IoT уредите со намера да се манипулира со податоците што се пренесуваат. Напаѓачите ги искористуваат пропустите во комуникациските канали, протоколите или механизмите за шифрирање за да ги изменат податоците што се разменуваат помеѓу IoT уредите, што потенцијално ќе доведе до неовластени дејства, манипулација со податоци или компромис на системот.



Слика 3.2. Демонстрација на напад на одбивање на услугата. Нападот се отпочнува на системот од интерес преку преоптоварување со злонамерни барања и сообраќај. Последиците стануваат очигледни кога системот не реагира, нарушувајќи ја неговата планирана функционалност, правејќи го истиот ефективно недостапен за легитимните корисници.

Напади со рутирање

Нападите со рутирање [70] се друг тип на напад што може да се случи во IoT. Во нападот за рутирање, напаѓачот манипулира со информациите за рутирање на мрежата, предизвикувајќи испраќање на податоците до погрешна дестинација или нарушување на достапноста на мрежата. Ова може да биде особено штетно во IoT мрежите, бидејќи многу уреди се потпираат на централизирана мрежна архитектура, што ги прави ранливи на такви напади. Постојат неколку различни типови на напади со рутирање во овие мрежи:

Нападите на рутирачката табела вклучуваат малициозни активности кои имаат за цел да манипулираат со рутирачките табели во мрежната инфраструктура за да го пренасочат сообраќајот на податоци кон неточни дестинации. Напаѓачите ги искористуваат пропустите во протоколите за рутирање, мрежните уреди или административните конфигурации за да ги променат патеките за рутирање, што потенцијално ќе доведе до пресретнување на податоци, неовластен пристап или прекини на услугите. На пример, напаѓачите може да внесат лажни или манипулирани информации за рутирање во рутирачките табели на мрежата. Со испраќање фалсификувани ажурирања за рутирање или искористување на пропусти во протоколите за рутирање, тие можат да влијаат на одлуките за рутирање на мрежата. Ова може да резултира со пренасочување на сообраќајот на неовластени или компромитирани дестинации под контрола на напаѓачот. Напаѓачите, исто така, може да користат киднапирање на рути или други механизми за да ја нарушат мрежата.

Нападите со инјектирање на пакети се однесуваат на малициозни активности каде што напаѓачите инјектираат лажни или фалсификувани пакети во мрежа, нарушувајќи го нормалниот проток на податоци и потенцијално предизвикувајќи дефект на уредите. Овие напади ги искористуваат пропустите во мрежните протоколи, софтверот или уредите за да манипулираат со мрежниот сообраќај, што доведува до прекини, неовластен пристап или компромитиран интегритет на системот. На пример, напаѓачите може да фалсификуваат пакети со фалсификувани изворни адреси, создавајќи илузија дека пакетите потекнуваат од доверлив извор или легитимен уред во мрежата. Со инјектирање на овие лажни пакети, напаѓачите можат да ги измамат мрежните уреди, апликациите или безбедносните механизми, што доведува до неовластен пристап, заобиколување на безбедносните контроли или предизвикување конфузија во мрежата.

Нападите за повторување на контекстот за рутирање се малициозни активности кои вклучуваат пресретнување и реемитување на пакети со податоци во контекстот на рутирање на IoT системот. Напаѓачите го фаќаат мрежниот сообраќај, вклучувајќи информации за рутирање. Со зафаќање на овие информации, тие добиваат увид во тековниот контекст на рутирање и патеките што ги следат пакетите со податоци во рамките на IoT системот, а подоцна повторно ги репродуцираат за да манипулираат со одлуките за рутирање, што потенцијално ќе доведе до грешки, неточности или пречки во рамките на IoT систем.

Слика 3.3. Демонстрација на напад со човек во средината. Напаѓачот лукаво се позиционира меѓу две страни кои комуницираат, пресретнувајќи и манипулирајќи со протокот на информации. Како што напредува нападот, напаѓачот суптилно ги менува разменетите пораки, вбризгувајќи ја сопствената злонамерна содржина или само ги набљудува чувствителните податоци што се пренесуваат.

Напади со повторување

Нападите со повторување [71] се друг вид мрежен напад кој вклучува пресретнување и реемитување на мрежниот сообраќај. Во контекст на IoT, нападите со повторување може да се случат кога напаѓачот пресретнува пакет со податоци испратен помеѓу два IoT уреди и го реемитува пакетот по одредено време, потенцијално предизвикувајќи грешки или неточности во системот. Постојат различни подтипови на напади со повторување што може да се појават во IoT мрежите:

Едноставен напад со повторување е најосновниот облик на напад со повторување, каде што напаѓачот повторно пренесува претходно снимена порака до целниот уред без да прави никакви измени. Овој тип на напад има за цел да го измами целниот уред со репродукција на оригиналната порака, потенцијално предизвикувајќи несакани последици или неовластени дејства.

Нападот со сериско повторување е напредна форма на напад со повторување, каде што напаѓачот пресретнува и повторно пренесува повеќе пораки во серија или низа. Овие пораки може да бидат дел од специфична комуникациска сесија, низа наредби или збир на инструкции релевантни за работата на системот. Со репродукција на серија снимени пораки, напаѓачот има за цел да манипулира со целниот IoT систем, потенцијално предизвикувајќи позначајна штета или прекини.

Нападот за повторување Man-in-the-Middle е софистицирана форма на напад со повторување каде што напаѓачот ја пресретнува комуникацијата помеѓу два уреди и ги модифицира пресретнатите пораки пред да ги препрати до целниот уред. Со вметнување меѓу легитимниот испраќач и примач (ова може да се постигне со искористување на пропусти во мрежната инфраструктура, загрозување на мрежен уред или извршување напади како ARP измама или труење со DNS), напаѓачот може да манипулира со содржината на пресретнати пораки, како и потенцијално да доведе до неовластени дејства, оштетување на податоците или компромис на системот.

Нападот со рефлексија е тип на напад каде што напаѓачот испраќа копија од пресретната порака назад до оригиналниот испраќач, што потенцијално ќе доведе до конфузија, прекини или несакани последици во системот. Овој напад ги искористува пропустите во мрежните конфигурации или протоколи за да манипулира со протокот на комуникација и да го искористи однесувањето на испраќачот. Нападите со рефлексија може да имаат засилено влијание кога дејствата што ги активира испраќачот како одговор на рефлектираната порака предизвикуваат несакани последици или дополнителни нарушувања во системот. Ова може да доведе до дополнителни комуникациски јамки, деградација на перформансите или дури и дефекти на системот.

Нападот со слепо повторување е тип на напад каде што напаѓачот пресретнува и повторно пренесува порака без да знае за нејзината содржина. Во овој напад, напаѓачот има за цел да предизвика нарушување или оштетување на системот со повторување на пресретнатата порака, дури и без да го разбере нејзиното значење или цел. Нападите со слепо повторување може да имаат засилено влијание поради непредвидливата природа на повторното пренесување на непознати пораки. Во зависност од специфичните пропусти или карактеристики на системот, слепото повторување на пораките може да доведе до каскадни неуспеси, ненамерни однесувања или компромитиран интегритет на податоците.

Напади со заглавување на мрежата

Нападите со заглавување на мрежата [72] се вид на DoS напади кои вклучуваат нарушување на комуникацијата помеѓу IoT уредите со преплавување на комуникацискиот канал со голем волумен на сигнали за радиофреквенција. Ова предизвикува IoT уредите да ја изгубат нивната способност за ефективно комуницирање, па дури и целосно да ја прекинат мрежата. Постојат различни типови на напади со заглавување, вклучувајќи:

Континуираното заглавување е тип на напад каде напаѓачот континуирано емитува радиофреквенциски сигнал со висока моќност, ефикасно блокирајќи ја целата комуникација на целната фреквенција или фреквенции. Овој напад со заглавување има за цел да ја наруши безжичната комуникација во засегнатата мрежа, правејќи ја недостапна или значително деградирајќи ги нејзините перформанси. Континуираните напади на заглавување на мрежата може да имаат колатерално влијание надвор од целната мрежа. Исто така и блиските мрежи кои работат на слични фреквенции може да доживеат пречки, што ќе доведе до пошироко нарушување на безжичната комуникација во близина. Ова може да влијае на повеќе IoT системи, мобилни уреди или други услуги зависни од безжично поврзување.

Случајното заглавување е вид на напад каде што напаѓачот наизменично емитува сигнал со ниска радиофреквенција (RF) во случајни интервали. За разлика од континуираните напади на мрежно заглавување, случајното заглавување на мрежата има за цел да ја наруши комуникацијата на посуптилен и непредвидлив начин, што го прави предизвик за ефикасно откривање и ублажување. Случајните напади со заглавување на мрежата се дизајнирани да работат тајно, избегнувајќи откривање со мешање со нормален RF шум или пречки во заднина. Природата со мала моќност на сигналот за заглавување, во комбинација со случајно тајмирање, ги намалува шансите лесно да се открие од системите за следење или анализаторите на спектарот.

Заглавувањето со измама е софистициран тип на заглавувачки напад каде напаѓачот намерно испраќа сигнал што го имитира легитимниот мрежен сообраќај. Во овој напад, лажниот сигнал содржи грешки, доцнења или други аномалии, што ги тера IoT уредите да го перцепираат како вистинска комуникација. Ова однесување ги спречува уредите да воспостават сигурни врски или ефективно да комуницираат едни со други.

Sybil напади

Sybil нападите [73] претставуваат тип на мрежен напад каде што напаѓачот создава повеќе лажни идентитети (познати како „Sybils“) со цел да стекне контрола над мрежата или да влијае на однесувањето на другите јазли во мрежата. Во контекст на IoT, овие напади можат да бидат особено штетни бидејќи можат да ја загрозат критичната инфраструктура и да предизвикаат значително нарушување на системот. Постојат различни подтипови на Sybil напади кои можат да се појават во контекст на IoT. Тие вклучуваат:

Нападите со исцрпување на ресурси вклучуваат создавање на бројни лажни идентитети од страна на напаѓачот за да се исцрпат ресурсите на целниот систем. Напаѓачот ги користи овие лажни идентитети за да троши критични ресурси како што се процесорската моќ, меморија или пропусниот опсег. Како што се исцрпуваат ресурсите на системот, легитимните корисници и процесите во целниот систем се лишени од потребните ресурси за извршување на нивните наменети функции. Ова може да резултира со бавно време на одговор, паѓање на системот или одбивање на условите за услуга, што го прави системот недостапен или несигурен. Нападите со исцрпување на ресурси имаат директно влијание врз достапноста и перформансите на целниот систем.

Sybil нападите со рутирање се специфичен тип на Sybil напад каде што напаѓачот создава повеќе лажни идентитети за да манипулира со одлуките за рутирање донесени од системот. Со контролирање на овие фабрикувани идентитети, напаѓачот се обидува да влијае на рутирачките патеки во мрежата. Преку манипулација со одлуките за рутирање, напаѓачот има за цел да го пренасочи протокот на податоци преку компромитирани или малициозни јазли што ги контролираат. Со погрешно насочување на сообраќајот на податоци низ овие јазли, напаѓачот може потенцијално да ги прислушува, модифицира или блокира пренесените податоци, загрозувајќи ја нивната доверливост, интегритет или достапност.

Нападите на репутација се тип на напад каде што напаѓачот создава повеќе лажни идентитети за да манипулира со репутацијата на одредени јазли во системот. Со контролирање на овие лажни идентитети, напаѓачот се обидува да ги измами другите јазли да им веруваат на компромитирани или злонамерни јазли, што може да доведе до несакани последици. Откако ќе се воспостави репутацијата на компромитирани јазли и ќе им веруваат други јазли, напаѓачот може да ја искористи оваа доверба за да пропагира малициозни активности. Ова може да вклучува ширење на лажни информации, вбризгување на злонамерен код или малициозен софтвер или манипулација со податоци и трансакции за лична корист или за предизвикување штета во системот.

Нападите со затемнување се тип на напад каде напаѓачот создава повеќе лажни идентитети или јазли во мрежата за да изолира одреден целен јазол или група јазли од остатокот од мрежата. Со опкружување на таргетираниот јазол со овие малициозни ентитети, напаѓачот има за цел да ја контролира или прекине неговата комуникација со легитимните јазли, што потенцијално ќе доведе до разни штетни ефекти како што се

започнување дистрибуирани DoS напади или стекнување неовластена контрола врз изолираните јазли.

3.2.3 Напади на ниво на сервисни платформи

Напади со користење на злонамерен софтвер

Злонамерниот софтвер [74] се однесува на тип на софтвер специјално дизајниран да нанесува штета, да ја загрози безбедноста, да украде податоци или да го наруши нормалното функционирање на системот. Во контекст на IoT системите, нападите со малициозен софтвер претставуваат значителна закана за слојот на услужната платформа. Злонамерниот софтвер е намерно создаден од злонамерни актери со различни цели, како што се финансиска добивка, шпионажа, саботажа или едноставно предизвикување хаос. Специфичните намери зад нападите на малициозен софтвер може да се движат од крадење чувствителни податоци, загрозување на интегритетот на системот, оневозможување на критичните функции или користење на компромитирани системи за понатамошни злонамерни активности. Некои вообичаени напади со малициозен софтвер што можат да се користат против овој слој се:

Ransomware [75] е специфичен тип на малициозен софтвер кој е дизајниран да ги шифрира податоците на јазолот на жртвата, што ги прави недостапни за сопственикот или корисникот. Напаѓачот зад откупниот софтвер потоа бара исплата на откуп, обично во криптовалути, во замена за обезбедување на клучот за дешифрирање или алатката неопходна за враќање на пристапот до шифрираните податоци. Напаѓачот може да се закани дека трајно ќе ги избрише податоците или ќе ги објави на јавноста доколку не се плати откупот. Во IoT системите, овој тип на малициозен софтвер може да го таргетира слојот на услужната платформа, што ги прави IoT уредите недостапни за комуникација со услугите базирани на облак. Исто така, овој тип на напад може да се примени на слојот на апликации.

Тројанците [76], познати и како Тројански коњи, се вид на малициозен софтвер кој се маскира како легитимна програма или датотека. Овој тип на малициозен софтвер ги мами корисниците со тоа што изгледа како безопасен или пожелен, додека неговата вистинска намера е да врши злонамерни активности. Тие често доаѓаат маскирани како легитимни датотеки, софтвер, па дури и прилози за е-пошта. Тие може да се дистрибуираат преку различни канали, како што се малициозни веб-локации, заразени преземања или измамнички прилози за е-пошта, што ги наведува корисниците несвесно да ги инсталираат или извршуваат. Тројанците може да се користат за кражба на чувствителни податоци, нарушување на перформансите на системот или обезбедување неовластен пристап. Во IoT системите, тројанците можат да го таргетираат слојот на услужната платформа со инфицирање на услуги базирани на облак. Исто така, овој тип на напад може да се примени на слојот на апликации.

Ботнет [77] е мрежа на компромитирани IoT уреди кои се инфицирани со малициозен софтвер и контролирани од централизиран сервер за команда и контрола (C&C). Терминот „ботнет“ е изведен од зборовите „робот“ и „мрежа“, нагласувајќи ја автоматизираната природа и меѓусебната поврзаност на овие компромитирани уреди. Ботнетите служат како моќна алатка за сајбер-криминалците, овозможувајќи им да вршат различни видови напади, вклучувајќи DDoS напади и кражба на податоци (Слика 3.4). Во IoT системите, ботнетите можат да го таргетираат слојот на услужната

платформа со инфицирање на услуги базирани на облак. Овој тип на напад може да се примени и на слојот на апликации.

IoT ботнетите се специфичен тип на ботнет кој цели на IoT уреди како што се: рутери, камери, други паметни уреди итн. Овие ботнети ги искористуваат пропустите во безбедносните карактеристики на IoT уредите и може да се користат за различни малициозни активности, како што се DDoS напади, злонамерно ископување криптовалути и ширење злонамерен софтвер. Некои примери на IoT ботнети се:

Mirai [78] е еден од најпознатите IoT ботнети. Тој е одговорен за еден од најголемите DDoS напади во историјата, кој се случил во 2016 година. Идејата зад нападот била да се таргетира давателот на услуги DNS Dyn и да се нарушат услугите за популарните веб-локации како Twitter, Netflix и Spotify. Mirai инфицира уреди користејќи стандардни параметри за најавување или со искористување на познати пропусти.

Reaper [177], исто така познат како IoTroop, е значаен IoT ботнет кој се појавил во 2017 година. Го привлекол вниманието на јавноста поради неговата способност да лансира DDoS напади и да заразува IoT уреди со малициозен софтвер, истакнувајќи се како пософистициран и еволуиран ботнет во споредба со неговиот претходник, Mirai. Reaper претставува значаен напредок во развојот и софистицираноста на IoT ботнетите, првенствено поради неговата модуларна архитектура и способноста за ажурирање на неговите карактеристики и методи на напад.

VPNFilter [178] е забележителен IoT ботнет кој конкретно ги таргетира мрежните рутери и другите уреди кои поддржуваат протоколи за безбедносен слој на приклучоци (SSL) и за безбедност на транспортниот слој (TLS). Овој ботнет е од интерес поради неговите напредни способности, кои вклучуваат кражба на податоци, следење на мрежниот сообраќај и способност за отпочнување DDoS напади. Овој ботнет претставува значајна закана за безбедноста и интегритетот на IoT мрежите и чувствителните податоци што се пренесуваат преку нив. VPNFilter конкретно се фокусира на компромитирачки мрежни рутери и слични уреди кои се основата на многу IoT мрежи.

Satori [179], варијанта на озлогласениот ботнет Mirai, се појавил кон крајот на 2017 година, воведувајќи нови способности и проширувајќи ги методите за напад што ги користи неговиот претходник. Додека Satori задржува некои од методите за напад на Mirai, тој вклучува дополнителни функции, како што е способноста да се шири преку Secure Socket Shell (SSH) и протоколите Telnet. Покрај тоа, Satori ја демонстрира способноста да започне DDoS напади и да пропагира малициозен софтвер за ископување криптовалути, дополнително истакнувајќи го неговиот потенцијал за широко распространето нарушување и финансиска добивка.

Persirai [180] е значаен ботнет кој конкретно се фокусира на таргетирање на IP камери користејќи децентрализирана мрежна архитектура, позната како Peer-to-Peer (P2P) модел. Првпат станал познат во 2017 година и споделува сличности со Mirai во однос на стратегијата за таргетирање, искористување на уреди со слаби или стандардни шеми за најавување. Штом IP-камера ќе се зарази со Persirai, ботнетот добива можност не само да ги шпионира корисниците туку и да започне разорни DDoS напади.

BrickerBot [180] се издвојува како уникатен и деструктивен IoT ботнет кој работи со посебна цел: да ги направи ранливите IoT уреди неупотребливи. За разлика од традиционалните ботнети кои компромитираат уреди за злонамерни цели, BrickerBot презема поагресивен пристап со искористување на познатите пропусти за трајно да ги оштети насочените IoT уреди. Тоа го постигнува со бришење на фирмверот или меморијата на уредот, што го отежнува или оневозможува враќањето на уредот во неговата оперативна состојба.

Напади преку експлоатација на софтвер

Нападите за експлоатација на софтвер [79] се однесуваат на искористување на ранливости во софтверските компоненти што се користат во IoT системите. Овие напади конкретно ги таргетираат слабостите во софтверската инфраструктура, апликациите или комуникациските протоколи кои овозможуваат функционирање на IoT уредите и интеракцијата меѓу нив. Овој тип на напад може да се примени и на слојот на апликации. Некои вообичаени софтверски експлоатации што можат да се лансираат против слојот на услужната платформа се:

Нулти-ден експлоатација ги користи предностите на пропустите во софтверот што се непознати за производителот или развивачот. Терминот „нулти ден“ означува дека пропустот бил откриен од напаѓач или злонамерен ентитет пред продавачот на софтверот да стане свесен за тоа или да објави поправка. Во контекст на IoT системите, нулта-дневните експлоатации претставуваат значајна закана, особено кога се таргетираат услуги базирани на облак што се користат од IoT уредите.

Софтверските експлоатации преку далечинско извршување на кодот (Remote Code Execution) се однесуваат на специфичен тип на напад што му овозможува на напаѓачот далечински да изврши произволен код на системот на жртвата. Овој тип на експлоатација е особено опасен бидејќи му дава на напаѓачот целосна контрола над компромитираниот систем, овозможувајќи му да извршува злонамерни команди, да инсталира дополнителен малициозен софтвер или да добие неовластен пристап до чувствителни податоци. Во IoT системите, RCE експлоатирањата може да ги таргетираат услугите базирани на облак што ги користат IoT уредите.

Софтверска експлоатација со прелевање на баферот [80] е тип на напад што конкретно ги насочува недостатоците во управувањето со меморијата на програмата. Ги користи предностите на пропустите што му овозможуваат на напаѓачот да прелева бафер во меморијата на програмата, предизвикувајќи таа да ги презапише соседните мемориски локации со злонамерен код или податоци. Во контекст на IoT системите, експлоатирањата за прелевање на баферот може да се насочат кон услуги базирани на облак што ги користат IoT уредите.

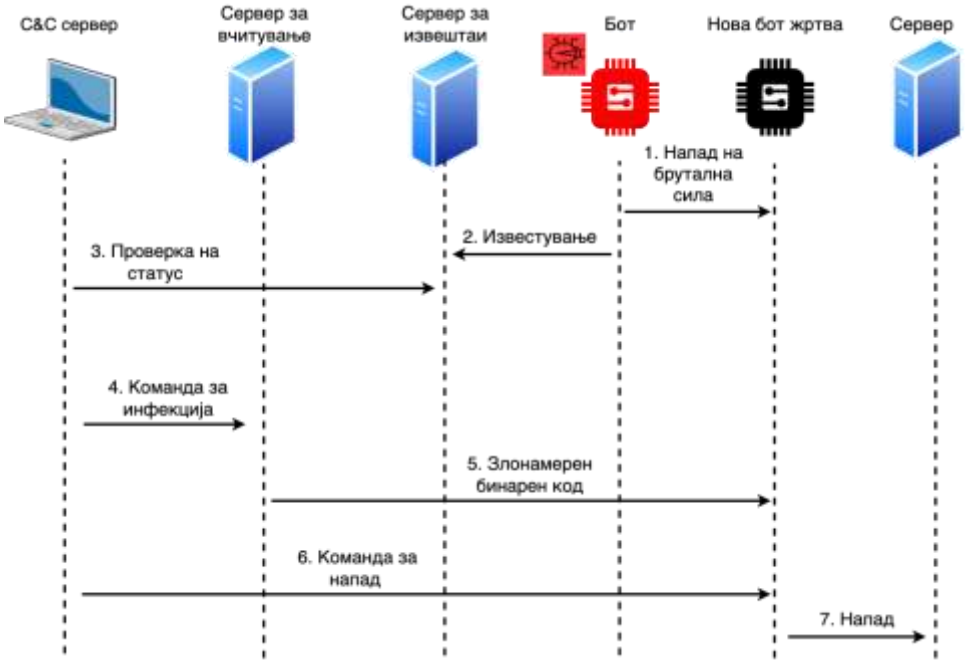
Напади на API

Нападите на API [81] се вообичаен тип на напад на слојот на сервисни платформи. IoT уредите често се потпираат на услуги базирани на облак за различни функционалности, како што се складирање податоци, аналитика, далечинско управување или ажурирања на софтвер. API-њата делуваат како интерфејс помеѓу IoT уредите и овие облак услуги, овозможувајќи беспрекорна комуникација и интеракција. Сепак, ранливостите во овие API може да бидат искористени од напаѓачите за да добијат неовластен пристап до услугите. Нападите на API ги користат слабостите во дизајнот, имплементацијата или конфигурацијата на API што ги користат IoT уредите. Овие пропусти може да

произлезат од несоодветни механизми за автентикација, небезбеден пренос на податоци, недоволна валидација на влезните податоци или недостатоци во ескалацијата на привилегиите. Некои вообичаени напади на API кои можат да бидат лансирани против слојот на сервисни платформи се:

Нападите за вбризгување код во структурен јазик за пребарување (SQL) ги користат предностите на пропустите во API-њата кои прифаќаат внесување на корисникот и конструираат SQL пребарувања за да комуницираат со базите на податоци. Напаѓачите ги идентификуваат недостатоците во валидацијата на влезот во API, дозволувајќи им да инјектираат злонамерен SQL-код во рамките на API барањето. Со вбризгување на внимателно изработен SQL код, напаѓачите можат да манипулираат со однесувањето на API и да извршат неовластени операции на базата на податоци. Вообичаените техники вклучуваат додавање дополнителни барања, модифицирање на постоечките барања или заобиколување на механизмите за автентикација за да се добие неовластен пристап до чувствителни податоци. Овој тип на напад може да има сериозни последици за IoT системите, особено кога се таргетираат услуги базирани на облак што ги користат IoT уредите.

Нападите на Cross Site Scripting (XSS) ги користат ранливостите во API-њата што не правилно го потврдуваат влезот на корисникот. Овие пропусти им овозможуваат на напаѓачите да инјектираат малициозни скрипти, обично напишани во JavaScript, во API барањата. Со инјектирање на малициозни скрипти, напаѓачот манипулира со однесувањето на API и го принудува веб-прелистувачот на целниот корисник да го изврши инјектираниот код. Злонамерните скрипти можат да украдат чувствителни информации, како што се ингеренциите за најавување, токени за сесија или лични податоци или да вршат неовластени дејства во име на корисникот. Нападите на XSS насочени кон услугите базирани на облак во IoT системите може да имаат значителни последици. Со искористување на пропустите во овие услуги, напаѓачите можат да ја загрозат доверливоста и интегритетот на податоците складирани или обработени во облакот. Тие, исто така, можат да манипулираат со однесувањето на услугите, што потенцијално ќе доведе до неовластен пристап, прекршување на податоците или прекини на услугите.



Слика 3.4. Фази на извршување на ботнет напади. Првата фаза вклучува регрутирање на ранливи уреди, обично преку искористување на пропусти на софтверот или техники на социјално инженерство. Откако ќе се состави огромна војска на ботнет, втората фаза, позната како фаза на команда и контрола (C&C), стапува во игра. Овде, напаѓачот воспоставува централизирана контролна инфраструктура за управување и координација на компромитирани уреди. Со ботнетот под нивна команда, напаѓачот продолжува во третата фаза: започнување координирани напади.

Напади на процесот на автентикација

Во нападите на процесот на автентикација [82], напаѓачите ги искористуваат пропустите во механизмот за автентикација што го користат API за да добијат неовластен пристап до услугите базирани на облак. Овие напади се особено релевантни во IoT системите, каде што услугите базирани на облак интензивно се користат од IoT уредите. Дополнително, постојат неколку подваријанти на напади за автентикација кои можат дополнително да ги искористат слабостите во процесот на автентикација.

Нападите со лозинка се вид на напад каде напаѓачот се обидува да ја погоди или пробие лозинката на корисникот за да добие неовластен пристап до систем, сметка или услуга. Во контекст на IoT системите, нападите со лозинки конкретно ги таргетираат услугите базирани на облак кои се користат од IoT уредите, што претставува значителен безбедносен ризик.

Нападите со брутална сила се тип на напад во кој напаѓачот систематски ги пробува сите можни комбинации на лозинки за да добие неовластен пристап до систем или сметка. Тие вклучуваат автоматизиран или рачен процес на систематско тестирање на сите можни комбинации на лозинки додека не се открие точната лозинка. Овој метод се потпира на претпоставката дека точната лозинка постои во одреден опсег или шема на знаци. Нападите со брутална сила можат да бидат интензивни за ресурси и одземаат многу време, бидејќи вклучуваат обиди за огромен број комбинации на лозинки. Напаѓачите користат различни техники, како што се напади со речник (пробување зборови од речник) или исцрпни повторувања засновани на знаци, за да ги максимизираат нивните шанси за успех. Во контекст на IoT системите, нападите со брутална сила можат конкретно да ги таргетираат услугите базирани на облак што ги користат IoT уредите.

Нападите на социјално инженерство вклучуваат манипулација со поединци за да ги измамат да откријат чувствителни информации, како што се лозинки или други доверливи податоци. Тие ги користат техниките на психолошка манипулација за да ги измамат и манипулираат поединците. Напаѓачите може да имитираат доверливи субјекти, како што се персоналот за техничка поддршка, давателите на услуги или колегите, за да ја добијат довербата на целниот корисник. Тие често се потпираат на тактики како што се кражба на е-пошта, телефонски повици или имитирање преку платформите на социјалните медиуми за да ги измамат корисниците да ги откријат нивните лозинки или други чувствителни информации. Во контекст на IoT системите, нападите од социјалниот инженеринг можат конкретно да ги таргетираат корисниците на услуги базирани на облак што се користат од IoT уредите.

3.2.4 Напади на апликациско ниво

Нападите со соголување на SSL/TLS [83] се форма на MitM напад кој ги искористува пропустите во протоколите SSL или TLS. Овие напади имаат за цел да ја намалат безбедноста на комуникацијата со пресретнување и мешање на слојот на каналот. Со пресретнување на комуникацијата помеѓу IoT уредот и услугата базирана на облак, напаѓачот ја принудува врската да користи нешифриран или слабо шифриран канал, отстранувајќи ја безбедноста обезбедена од SSL/TLS. Во IoT системите, нападите за одземање на SSL/TLS можат конкретно да ја таргетираат комуникацијата помеѓу IoT уредите и услугите базирани на облак, загрозувајќи ја доверливоста и интегритетот на податоците.

Нападите со измама на DNS [84], познати и како труење со кешот на DNS или напади на киднапирање на DNS, се форма на напад на MitM каде напаѓачот манипулира со DNS за да ја пренасочи комуникацијата помеѓу IoT уредите и услугите базирани на облак кон лажна веб-локација. Како резултат на тоа, IoT уредите несвесно воспоставуваат врски со серверот на напаѓачот наместо со предвидената легитимна услуга. Ова му овозможува на напаѓачот да пресретнува, следи, менува или ја блокира комуникацијата помеѓу IoT уредите и услугите базирани на облак, што потенцијално ја загрозува доверливоста, интегритетот и достапноста на податоците.

ARP измамнички напади [85], исто така познати како труење на кешот на ARP или напади на труење со ARP, се форма на напад на MitM што го таргетира протоколот за резолуција на адреса (ARP) за да ја пренасочи комуникацијата помеѓу IoT уредите и услугите базирани на облак кон системот на напаѓачот. Во IoT системите, ARP нападите конкретно ја таргетираат комуникацијата помеѓу IoT уредите и услугите базирани на облак, компромитувајќи ја доверливоста, интегритетот и достапноста на комуникацијата. Напаѓачот испраќа лажни ARP пораки за да ги измами IoT уредите да ја поврзат MAC адресата на напаѓачот со IP адресата на легитимната услуга базирана на облак. Како резултат на тоа, IoT уредите несвесно го испраќаат својот мрежен сообраќај до системот на напаѓачот, наместо до наменетата дестинација. Ова му овозможува на напаѓачот да прислушува чувствителни податоци, да менува команди или одговори, па дури и да инјектира злонамерна содржина во протокот на комуникација.

Напади врз процесот на виртуелизација

Нападите врз процесот на виртуелизација [86] се тип на напад каде што напаѓачот ги искористува пропустите во слојот за виртуелизација на услугата базирана на облак за да добие неовластен пристап до системот домаќин или други виртуелни машини. Во IoT системите, нападите за виртуелизација можат да ги таргетираат услугите базирани на облак што ги користат уредите за IoT за складирање и обработка на податоци. Некои вообичаени подтипови на напади за виртуелизација во IoT системи се:

Нападите врз хипервизорот вклучуваат искористување на ранливости во хипервизорот, кој е софтверски слој кој овозможува виртуелизација на ресурсите и управување со виртуелни машини на физички систем домаќин. Откако ќе влезат во хипервизорот, напаѓачите можат да таргетираат други виртуелни машини кои работат на истиот физички хост. Тие можат да ги искористат ранливостите или слабостите во механизмите за изолација на хипервизорот за да излезат од нивната почетна виртуелна машина и да добијат контрола над другите виртуелни машини, потенцијално компромитирајќи го интегритетот и доверливоста на податоците складирани во тие виртуелни машини.

Нападите на хипервизорите во IoT системите може да имаат значителни последици за услугите базирани на облак кои се справуваат со IoT податоци. Со загрозување на инфраструктурата на хипервизорот, напаѓачите можат да пристапат и да манипулираат со чувствителни податоци складирани во облакот, да ја нарушат достапноста на услугите или да го компромитираат интегритетот на податоците обработени од IoT уредите.

Нападите врз виртуелни машини (VM) вклучуваат искористување на пропусти во виртуелната машина за да се добие неовластен пристап до системот домаќин или други виртуелни машини. Ако напаѓачите успешно ја компромитираат виртуелната машина, тие може да се обидат да ги искористат слабостите во околината за виртуелизација за да добијат пристап до други VM хостирани на истиот физички сервер. Со пробивање на првичните VM граници, напаѓачите можат да се движат странично во виртуелизираната околина, потенцијално компромитувајќи ја безбедноста на другите VM и податоците што ги содржат. Нападите на VM во IoT системите може да имаат значителни импликации за услугите базирани на облак кои се справуваат со IoT податоци. Ако напаѓачот добие неовластен пристап до виртуелните машини што ги извршуваат услугите на облак, тој може да ги манипулира податоците, да извлече чувствителни информации, да ја наруши достапноста на услугите или да изврши секундарни напади насочени кон IoT уредите или самата инфраструктура на облак.

Напади со хакирање на контејнерите [87] - Контејнерите се лесна форма на виртуелизација која вообичаено се користи во IoT околини за да се обезбедат изолирани средини за самите апликации. Напаѓачите може да ги таргетираат пропустите во софтверскиот пласт на контејнерите, вклучувајќи го времето на траење на контејнерот, рамки за оркестрација или самите слики од контејнерот. Овие пропусти може да произлезат од погрешни конфигурации, небезбедни слики од контејнери или застарени софтверски компоненти. Со искористување на пропустите на контејнерите, напаѓачите имаат за цел да добијат неовластен пристап до системот домаќин на кој се распоредени контејнерите. Откако ќе влезат во компромитираниот контејнер, напаѓачите може да се обидат да ги ескалираат привилегиите, како и да извршат произволен код за да добијат подлабок пристап до основниот систем на домаќинот. Нападите со хакирање на контејнер може да имаат значителни импликации за IoT системите. Со загрозување на контејнеризираните апликации или услуги, напаѓачите можат да добијат неовластен пристап до чувствителни податоци, да манипулираат со однесувањето на IoT уредите, да ја нарушат достапноста на услугите, па дури и да започнат секундарни напади насочени кон други делови од IoT инфраструктурата.

Напади врз деловната логика

Нападите врз деловната логика [88] се тип на напад кој го таргетира апликативниот слој на IoT системите, кој е одговорен за логиката и функционалноста на самата апликација. Притоа, напаѓачите можат да ги искористат дупките во деловната логика на IoT апликацијата за да извршат лажни активности или да добијат неовластен пристап до чувствителни податоци или ресурси. Нападите од овој тип можат да имаат различни форми и нивното откривање/ублажување е чест предизвик. Општо земено, овие напади ги искористуваат недоследностите помеѓу дизајнот и имплементацијата на логиката на IoT апликацијата, дозволувајќи им на напаѓачите да манипулираат со однесувањето на апликацијата во своја полза. Еве неколку вообичаени типови на напади врз деловната логика во IoT системите:

Нападите за валидација на влез се тип на напад каде што напаѓачот доставува неважечки или злонамерен влез до апликацијата со цел да ги заобиколи безбедносните механизми или да изврши неовластени дејства на апликацијата. Нападите за валидација на влезот можат да ги искористат пропустите во логиката на IoT апликацијата, како што е недоволната валидација на влезовите од корисникот или некои недостатоци во самите податоци. На пример, напаѓачот може да поднесе специјално изработен влез за апликација за да ги заобиколи проверките за автентикација или овластување или да внесе злонамерен код.

Нападите за киднапирање на сесии (Слика 3.5.) се тип на напад каде што напаѓачот „киднапира“ сесија на активен корисник за да добие неовластен пристап до ресурсите или податоците на IoT апликацијата. Овој тип на напад може да ги искористи пропустите во логиката за управување со сесиите на апликацијата, како што се слаби ID на сесии или фиксирање на сесиите. На пример, напаѓачот може да го украде идентификаторот на сесијата на корисникот и да го користи за да го имитира корисникот и да врши неовластени дејства во апликацијата.

Слика 3.5. Фази на извршување на напад на киднапирање на сесија. Почетната фаза вклучува следење и собирање клучни информации за целната сесија, како што се идентификатори на сесии или токени за автентикација. Вооружен со ова знаење, напаѓачот продолжува до фазата на пресретнување, каде што ја пресретнува и ја прислушкува комуникацијата помеѓу жртвата и серверот. Ова му овозможува да добие неовластен пристап до сесијата и да извлече чувствителни податоци или да манипулира со нејзината содржина. Кога сесијата е компромитирана, напаѓачот потоа ја презема контролата, преземајќи ги привилегиите и можностите на жртвата.

Глава 4 Системи за детекција на аномалии во мрежен сообраќај

За аномалија или упад во мрежи на Интернет на нешта може да се смета било каква невообичаена активност или однесување на уредите/сообраќајот која отскокнува од секојдневните операции што се извршуваат. Аномалиите може да се појават во различни форми, како на пример: неочекуван мрежен сообраќај (ненадеен пораст на сообраќајот или отстапување од очекуваното однесување), невообичаено однесување на самиот IoT уред (пренос на податоци во чудни временски рамки, користење на повеќе енергија од вообичаеното), неовластени обиди за пристап (повеќекратни обиди за најавување со неточни податоци), како и инфекции со злонамерен софтвер или вируси.

4.1 Типови на системи за детекција на аномалии

Системите за детекција на аномалии (системи за детекција на упад - IDS) претставуваат тип на безбедносна технологија чија задача е откривање на потенцијални безбедносни пропусти во мрежата или системот на домаќинот. Најчесто, ваквите системи генерираат предупредувања при откривање на потенцијални безбедносни закани, дозволувајќи им на организациите да преземат брзи активности, како што е блокирање на мрежниот сообраќај или ставање уреди во карантин. Известувањата обично вклучуваат информации за природата на заканата, изворот и дестинацијата на нападот и степенот на сериозност на нападот (Слика 4.1.). Системите за откривање аномалии може да се категоризираат во четири главни категории: тип на архитектура, локација, метод за откривање што се користи и степен на интеграција.

4.1.1 Системи за детекција на аномалии според архитектурата

Host-базирани IDS

Системите за детекција на упад базиран на домаќин (Host IDS) се тип на IDS кои ги следат поединечните уреди, како што се серверите или работните станици со цел да детектираат знаци на упад. За разлика од IDS базирани на мрежа, кои го следат мрежниот сообраќај за знаци на упад, HIDS се фокусираат на активноста на поединечните уреди. Вообичаено, HIDS се инсталира на уредот што се следи и ги следи системските датотеки, нивниот интегритет и други аспекти на оперативниот систем за откривање знаци на неовластен пристап или друга сомнителна активност. На пример, HIDS може да ги следи промените на критичните системски датотеки или неовластени обиди за најава на самиот уред. HIDS може да обезбеди подетални информации за потенцијални упади отколку IDS базиран на мрежа, бидејќи има пристап до поспецификни информации за уредот што се следи. Исто така, тој може да детектира напади кои можеби не се видливи во мрежниот сообраќај, како што се напади кои ги искористуваат пропустите во одредени апликации или оперативни системи.

Еден од недостатоците на HIDS е побарувачката на ресурси, бидејќи истиот мора да работи на секој уред што се следи и може да троши значителни системски ресурси. Дополнително, тој можеби нема да биде ефективен против нападите што се случиле пред

да се инсталира HIDS или против нападите што ја таргетираат мрежата наместо поединечните уреди. И покрај овие ограничувања, HIDS може да биде ефикасен начин за дополнување на IDS базиран на мрежа и обезбедување дополнителна видливост на потенцијалните безбедносни закани. Овој тип систем најчесто се користи во владини агенции и/или финансиски институции, каде што цената на нарушување на безбедноста е особено висока.

IDS базирани на мрежа

Системите за детекција на упад базирани на мрежа (Network IDS) се тип на IDS кои го следат мрежниот сообраќај за знаци на упад. NIDS работат со анализа на податоците присутни во мрежата, барајќи шаблони кои укажуваат на потенцијални безбедносни закани. Овој тип на систем може да се распореди на различни точки во мрежата, вклучувајќи ги и крајните точки како што се серверите и работните станици. Најчеста стратегија за распоредување е NIDS да се постави на самиот периметар (граница) на мрежата, од каде што ќе може да го следи целиот влезен и излезен сообраќај и да детектира потенцијални закани пред истите да влезат во мрежата. NIDS може да користи различни методи за детекција за да ги идентификува потенцијалните безбедносни закани, вклучувајќи детекција заснована на потпис (signature), детекција на аномалии и анализа на протоколи. Во принцип, ваквиот систем генерира предупредувања штом се откријат потенцијални закани, дозволувајќи им на безбедносните тимови да ја испитаат заканата и адекватно да реагираат на истата. Покрај тоа, некои NIDS се способни да преземат и автоматизирани дејства, како што е блокирање на сообраќајот од одреден извор или ставање во карантин на одреден компромитиран уред.

NIDS може да биде ефикасен начин за детекција на широк опсег на потенцијални безбедносни закани и обезбедување сеопфатна покриеност на мрежата. Сепак, ваквите системи често можат да генерираат голем број лажно позитивни одлуки (false positive) и можеби нема да бидат толку ефикасни против напади кои користат шифрирање или други методи за да ја прикријат нивната активност. И покрај овие ограничувања, NIDS се суштинска компонента на безбедносните стратегии на многу организации и вообичаено се користат заедно со други безбедносни мерки, како што се заштитни ѕидови и антивирусен софтвер, со цел да обезбедат повеќе слоеви на заштита од потенцијални безбедносни закани.

Хибридни IDS

Хибридните системи за детекција на упад (Hybrid IDS) ги комбинираат придобивките од IDS базирани на мрежа и базирани на домаќин за да обезбедат сеопфатно безбедносно решение. Хибридниот IDS го следи мрежниот сообраќај, како и активноста на поединечни системи и предупредува за сомнителни активности што се случуваат на тие системи/мрежа. Со вакво комбинирање, овој тип на IDS може да обезбеди поцелосно покривање на мрежата и системите на организацијата, откривајќи потенцијални закани што можеби би ги пропуштиле NIDS и HIDS.

Дополнително, хибридниот IDS може да обезбеди повеќе информации околу предупредувањата, овозможувајќи им на безбедносните тимови побрзо и попрецизно да го идентификуваат изворот на потенцијалната закана и да преземат соодветни мерки. На пример, ако се детектира потенцијална закана која потекнува од одредена IP адреса, хибридниот IDS може да го идентификува конкретниот систем кој е извор на заканата и да обезбеди дополнителни информации за природата на оваа активност. Хибридниот IDS може да биде покомплексен и поскап за распоредување во споредба со другите IDS,

бидејќи бара и мрежни и сензори базирани на домаќин, како и инфраструктура за собирање и анализа на податоци од двата типа сензори. Сепак, за организации со високо ниво на безбедносен ризик, хибридниот IDS може да обезбеди дополнителен слој на заштита од потенцијални закани.

4.1.2 Системи за детекција на аномалии според локацијата

Централизираните IDS ги испраќаат податоците од мрежните сензори до централна локација за анализа и носење на одлуки. Централниот систем собира и корелира податоци од повеќе сензори, што го олеснува идентификувањето и одговарањето на заканите низ целата мрежа. Централизираните IDS обично се полесни за управување, но тие може да бидат помалку отпорни на напади чија цел е централната локација.

Дистрибуираните IDS имаат сензори кои се распоредени на повеќе локации во мрежата и секој од нив врши сопствена анализа и одлучување. Дистрибуираните IDS често се користат во големи и сложени мрежи, каде што е тешко или речиси невозможно да се централизира целото собирање и анализа на податоци. Дистрибуираните IDS се поотпорни на напади бидејќи се дистрибуирани низ мрежата, но можат да бидат поголем предизвик за управување поради високиот број на вклучени сензори и извори на податоци.

Хибридните IDS комбинираат елементи на централизирани и дистрибуирани IDS архитектури. Во хибриден IDS, сензорите се распоредени на различни точки во мрежата, а некои податоци се испраќаат до централна локација за анализа и донесување одлуки, додека другите податоци се анализираат и за нив се одлучува локално. Хибридните IDS можат да го обезбедат најдоброто од двата света, комбинирајќи ја еластичноста на дистрибуираните IDS со управливоста на централизираните IDS.

4.1.3 Системи за детекција на аномалии според методот на детекција

IDS базирани на потпис

Системите за детекција на упад базирани на потпис (Signature IDS) претставуваат тип на IDS кои користат база на податоци со познати потписи за напад со цел да ги идентификуваат и да ги предупредат безбедносните тимови за потенцијални закани. Во суштина, овие потписи претставуваат шаблони на познато злонамерно однесување кои биле претходно идентификувани и запишани во базата на SIDS. Кога мрежниот сообраќај поминува низ SIDS, примените пакети се анализираат според базата на податоци за потписи за да се утврди можно совпаѓање со некои познати злонамерни потписи.

IDS базирани на потпис се релативно едноставни за имплементација и работа и истите може да бидат ефективни во идентификување на познати закани. Тие можат прецизно да детектираат напади што се совпаѓаат со постоечките потписи, што ги прави особено корисни за идентификување на добро познат и широко користен злонамерен софтвер. Исто така, овие типови системи може да се користат за детекција на специфични типови напади насочени кон одредени апликации или оперативни системи. Сепак, IDS базирани на потпис имаат некои ограничувања. Тие се ефикасни само во детекција на напади што се совпаѓаат со познати потписи, што значи дека не се ефикасни против zero-day напади

или напади кои користат нови и претходно непознати тактики. Дополнително, напаѓачите можат лесно да ги менуваат своите напади за да ги заобиколат IDS базирани на потпис со менување на кодот или однесувањето на начини што не се совпаѓаат со постоечките потписи. За да се решат овие ограничувања, IDS базирани на потпис често се користат заедно со други видови IDS, како што се IDS базирани на однесување. Оваа комбинација на методи за детекција може да ја подобри севкупната ефикасност на IDS со зголемување на опсегот на закани што може да се откријат и намалување на бројот на лажно позитивни отчитувања.

IDS базирани на аномалија

Системите за детекција на упад базирани на аномалија (Anomaly IDS) се тип на IDS кои користат статистичка анализа и алгоритми од машинско учење за да воспостават првична основа за нормално однесување на мрежниот сообраќај, а потоа го следат сообраќајот за отстапувања од таа основа. Аномалиите во мрежниот сообраќај, како што се нагло зголемување на обемот на сообраќајот или невообичаени шаблони при пренос на податоци се означуваат како потенцијални безбедносни закани. AIDS се особено корисни за идентификување на претходно непознати напади и zero-day напади, бидејќи тие можат да детектираат шаблони на однесување кои не се вклучени во базата на податоци со познати потписи за напад. Исто така, тие се ефикасни во откривањето на внатрешни закани, бидејќи со нив може да се идентификува некое нестандардно однесување на вработен во мрежата.

IDS базиран на аномалии може да биде посложен за имплементација од IDS базиран на потпис, бидејќи тој бара создавање на основа за нормално однесување на мрежата, како и тековно следење и анализа на сообраќајот со цел за идентификација на можните отстапувања. За олеснување на овој процес често се користат алгоритми од машинско учење, бидејќи тие можат да бидат обучени да препознаваат модели на однесување што укажуваат на безбедносни закани. Еден од предизвиците на IDS базиран на аномалии се лажно позитивните отчитувања, бидејќи понекогаш дури и легитимниот мрежен сообраќај може погрешно да се идентификува како нестандардно однесување. Ова може да создаде преоптовареност на безбедносните тимови со предупредувања и други потенцијални недостатоци при реални безбедносни закани.

IDS базирани на однесување

Системите за детекција на упад базирани на однесување (Behaviour IDS) се фокусираат на откривање на нелогично однесување во мрежата или системот. Тие го следат однесувањето на корисниците и компонентите на системот за да идентификуваат активност што е надвор од нормалниот опсег на однесување. Целта на IDS базиран на однесување е да ги идентификува безбедносните закани кои не се откриени со други методи, како што е IDS базиран на потпис. BIDS користи различни техники за да идентификува нелогично однесување. Овие техники вклучуваат статистичка анализа, машинско учење и ископување на податоци (data mining). Статистичката анализа вклучува собирање и анализа на податоци за однесувањето на корисниците и системите со цел да се идентификуваат моделите на однесување што може да укажат на закана за безбедноста. Техниките за машинско учење се користат за да се изградат модели на нормално однесување и да се идентификуваат отстапувањата од овие модели. Техниките за ископување на податоци се користат за да се идентификуваат скриените шаблони во големите групи на податоци.

Кај BIDS постојат неколку предности во однос на другите типови IDS. Една од нив е тоа што тие можат да детектираат zero-day напади. Понатаму, тие можат да се прилагодат на промените во мрежното опкружување, како што се промени во однесувањето на корисниците или нови типови на напади. Сепак, IDS базирани на однесување имаат и некои ограничувања. Тие можат да генерираат голем број на лажно позитивни отчитувања ако моделите на нормално однесување не се внимателно изработени. Исто така, тие може да бидат помалку ефикасни во откривањето на напади кои се насочени и дизајнирани да избегнуваат детекција. За да се решат овие ограничувања, BIDS често се користат заедно со други видови IDS, како што се IDS базирани на потпис. Оваа комбинација на методи за детекција може да ја подобри севкупната ефективност на овој тип IDS со намалување на лажно позитивните отчитувања и зголемување на опсегот на закани што може да се откријат.

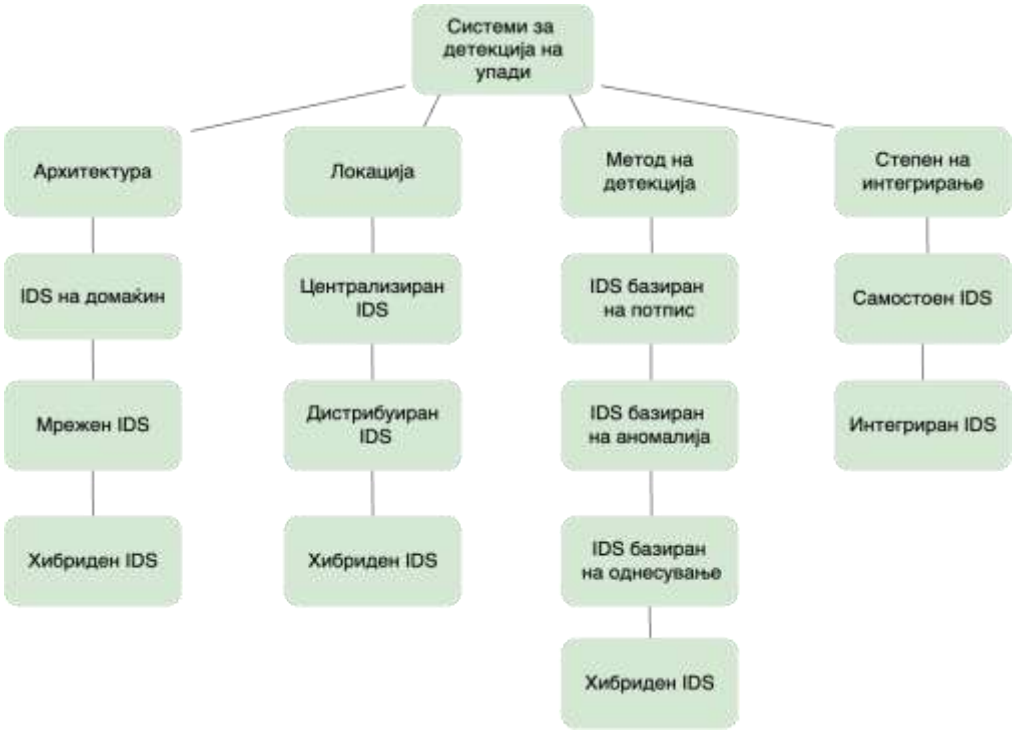
Хибридни IDS за детекција на аномалии

Хибридните системи за детекција на аномалии комбинираат повеќе методи за детекција, вклучувајќи ги и пристапите базирани на потпис и пристапи базирани на однесување, за да ја подобрат точноста на детекција на упад. Со искористување на предностите на секој пристап, хибридните системи можат да обезбедат посеопфатно покривање на потенцијалните безбедносни закани. Целта на хибридниот систем за детекција на аномалии е да открие широк опсег на обиди за упад, вклучувајќи познати и непознати напади, притоа минимизирајќи ги лажно позитивните отчитувања. Ова се постигнува со идентификување на познатите потписи за напад преку користење на детекција базирано на потпис, а потоа со користење на детекција базирано на однесување за да се идентификуваат претходно непознати или zero-day напади кои не се совпаѓаат со познатите шаблони. Хибридните системи за детекција на аномалии можат да бидат особено корисни во откривањето на насочени напади, каде што напаѓачите користат софистицирани техники за да избегнат откривање од традиционалните IDS. Со користење на комбинација од методи за детекција, хибридните системи можат попрецизно и побрзо да ги откријат овие типови напади. Сепак, имплементирањето на хибриден систем за детекција на аномалии може да биде предизвик, бидејќи истиот бара интеграција на повеќе технологии за детекција и воспоставување на робусна основа на нормално однесување. Дополнително, системот мора постојано да се следи и ажурира за да ги земе предвид промените во однесувањето на мрежата и новите типови на напади.

4.1.4 Системи за детекција на аномалии според степенот на интегрирање

Самостојните IDS претставуваат уреди или софтверски апликации кои работат независно од другите безбедносни технологии. Тие може да имаат свои управувачки конзоли и може да се распоредат on-premises или во облак. Самостојните IDS се лесни за поставување и управување и можат да обезбедат длабок увид во мрежниот сообраќај или активностите на домаќинот. Сепак, тие не се толку ефикасни како интегрираните IDS во детекција и спречување напредни или постојани закани. Тоа е затоа што на самостојните IDS им недостасува способноста да се интегрират со други безбедносни технологии, како што се заштитни ѕидови, антивирусен софтвер или системи за управување со безбедносни информации и настани (SIEM) за да обезбедат поцелосен пристап кон безбедноста.

Интегрираните IDS се системи кои најчесто се интегрирани со други безбедносни технологии, како што се заштитни ѕидови, антивирусен софтвер или SIEM. Тие може да споделуваат управувачки конзоли или да работат заедно како дел од поголема безбедносна целина. Интегрираните IDS можат да им помогнат на организациите да откријат и поефикасно да одговорат на напредните или постојаните закани, бидејќи можат да ги искористат можностите на другите безбедносни технологии за да обезбедат покоординиран одговор. Сепак, овие системи може да бидат посложени за поставување и управување, бидејќи бараат интеграција со други безбедносни технологии и може да бараат понапредна техничка експертиза.



Слика 4.1. Видови системи за откривање аномалии. Изборот на одреден IDS зависи од специфичните безбедносни барања, мрежната инфраструктура и нивото на видливост потребно за ефективно откривање на упад.

4.2 Машинско учење во системи за детекција на аномалии кај IoT

Машинското учење (ML) претставува технологија која дава можност за понатамошен развој на системите за откривање на упад (IDS) за IoT бидејќи може да помогне да се надминат некои од ограничувањата на традиционалните пристапи базирани на потписи и однесување. Во IoT мрежите, огромниот обем на податоци генерирани од поврзаните уреди претставува проблем за безбедносните тимови кои работат на идентификација и одговор на потенцијалните закани. Ова е проблем каде што едно од најдобрите можни решенија би било користењето на алгоритми од машинско учење. ML алгоритмите може да бидат обучени за автоматско откривање на шаблони и аномалии во мрежниот сообраќај и активноста во системот што може да укаже на потенцијален упад. ML процесите може да содржат (но не се ограничени на овие) неколку чекори [176]:

- *Собирање податоци:* ML алгоритмите имаат потреба од големи количини на податоци од кои може да учат. Во поглед на IDS, ова може да вклучува записи од мрежен сообраќај, системски датотеки и други релевантни извори на податоци.

- *Претходна обработка на податоците:* Откако податоците ќе се соберат, тие мора да бидат претходно обработени за да се осигура дека се во соодветниот формат за анализа. Ова може да вклучува чистење на податоците, нивно трансформирање во соодветен формат и нормализирање за тие да бидат конзистентни.

- *Извлекување на карактеристики:* Со цел да се идентификуваат шаблоните и аномалиите во податоците, ML алгоритмите најпрвин мора да ги идентификуваат релевантните карактеристики или атрибути. Ова може да вклучува избор на одредени полиња од записи на мрежниот сообраќај или идентификување на специфични системски повици што може да укажат на напад.

- *Процес на обука (тренирање):* Откако податоците ќе бидат обработени и ќе се извлечат избраните карактеристики, ML алгоритмот може да се обучи со користење на надгледувани или ненадгледувани техники за учење. Во надгледуваното учење, алгоритмот се тренира со користење на означени податоци кои веќе се класифицирани како нормални или злонамерни. Во учењето без надзор, алгоритмот мора да идентификува шаблони и аномалии во податоците без претходно знаење за тоа што претставува нормално однесување.

- *Процес на валидација:* Откако алгоритмот е обучен, тој мора да се тестира и да се потврди неговата точност со користење на посебен сет на податоци. Ова помага да се добие сигурност дека алгоритмот е способен точно да ги идентификува нормалниот сообраќај но и потенцијалните закани, а дополнително и да ги намали лажните позитивни и лажните негативни отчитувања.

- *Распоредување во реална околина:* Кога сите претходни процеси ќе бидат комплетирани, ML алгоритмот може да се распореди во средина во која ќе може постојано да го следи мрежниот сообраќај и системската активност и да ги предупредува безбедносните тимови за потенцијални закани или напади.

Една од клучните придобивки на ML во IDS е неговата способност за учење и прилагодување на новите закани со текот на времето. За разлика од пристапите засновани на потпис кои се потпираат на однапред дефинирани правила или потписи за да ги идентификуваат заканите, ML алгоритмите можат да учат од новите податоци и да ги ажурираат своите модели за да ги откријат претходно невидените закани. Дополнително, ML алгоритмите можат да помогнат да се намалат лажните позитивни и лажните негативни отчитувања во IDS. Со анализа на големи количини на податоци и идентификување на суптилни шаблони, овие алгоритми можат попрецизно да ги идентификуваат потенцијалните закани и да го намалат бројот на лажни предупредувања. Друга предност на ML во IDS е неговата способност да открие zero-day напади.

Повеќето од истражувањата за откривање аномалии во IoT мрежи се фокусираат на надгледувано ML извршено на податочни множества собрани од IoT уреди во контролирана средина. Истражувањата покажуваат успешна примена на ML во

справувањето со проблемот со големи податоци во мрежната безбедност, со методи кои варираат од надгледувано до ненадгледувано ML, од кои некои се фокусираат на примена во реално време.

Многу истражувачи ги испитуваат перформансите на популарните ML алгоритми како што се векторски машини за поддршка (Support Vector Machines), наивен Баесов класификатор (Naïve Bayes), дрво на одлука (Decision Tree) и алгоритам на случајна шума (Random Forest) [89-91]. Најдобри перформанси генерално се добиени од алгоритмите SVM [92] и XGBoost (eXtreme Gradient Boosting) [93], двата со точност нешто под 99%. Сахид и сор. [94] предлагаат IDS базиран на надгледувано ML учење за аномалии во IoT мрежи. По нормализацијата и намалувањето на димензијата на податочното множество обучени се шест различни ML модели. Сите модели имаат прецизност во откривањето од околу 99%.

Некои пристапи, како Голмах [95] и Танпуре [96] предлагаат хибриден IDS со класификациски модел за откривање на нестандардно мрежно однесување. Првиот користи SVM, а вториот користи K-means кластерирање и NB за групирање и класификација на податоците. Може да се види дека со ваквата комбинација на алгоритми се подобрува точноста на откривање аномалии. Во [97], предложен е хибриден пристап кој користи шума за изолација, самоорганизирана мапа (Self-Organizing Map), SVM со една класа и алгоритам на Гаусово моделирање на мешавини (Gaussian Mixture Modelling) за откривање аномалии, постигнувајќи точност од 98%.

Исто така, ML модели базирани на DT се предложени за откривање напади во IoT мрежите, постигнувајќи 100% перформанси во сите параметри при извршување на MitM напад на нивното тест множество. Авторите во [98] и [99] тврдат дека според нивните истражувања DT има највисока стапка на точност од 98,23% во споредба со останатите класификатори кои се вклучени во истражувањето (SVM, NB и Adaboost). Нобакхт [100] и Хосеин и сор. [101] користат алгоритми на логистичка регресија (Logistic Regression), SVM и алгоритам на вештачки имунолошки систем (Artficial Immune System) за откривање само на еден тип закани (DoS, ботнет напади итн.).

Авторите во [102] ја оценуваат можноста за користење на различни ML алгоритми (DT, SVM и K-means кластерирање) за откривање на безбедносни напади во медицински IoT уреди. Резултатите покажуваат дека алгоритмите засновани на дрво на одлуки постигнуваат најголема точност на откривање (~90%). Интелигентното и динамичко откривање на ширењето на ransomware во медицинските сајбер физички системи е тема на интерес во [103]. Во ова истражување, два различни ML модели се успешни во откривање и класифицирање на овие типови напади, при што NB има точност од 99,99%. SVM е избраниот алгоритам и во истражувањето на Вернер и Бутвиник [104], каде што податоците од сензор за гликоза во крвта се следат во обид да се детектираат случајни упади со модификација на податоците.

Нова методологија за избор на карактеристики наречена CorrAUC, базирана на мерката површина под ROC кривата (Area Under the ROC Curve) е предложена во [105], постигнувајќи над 96% точност со имплементација на NB, RF и SVM на податочно множество со повеќе класи. Во [106], алгоритмот за оптимизација на случаен рој (Particle Swarm Optimization) е искористен за избор на карактеристики за дизајнирање на напреден IDS за IoT мрежи, постигнувајќи 98% точност за бинарна класификација и 83% за класификација со повеќе класи на податочното множество IoTID20[107].

Во [108] предложена е безбедносна архитектура која користи софтверски дефинирани мрежи (Software Defined Network) и овозможувачи на виртуелизација на мрежни функции (Network Function Virtualization) за намалување на ранливостите на IoT мрежите, имплементирајќи и IDS за откривање на аномалии на сензорски податоци. Во овој случај, SVM покажува прецизност на откривање повисока од 98% за повеќето предложени комбинации на податоци. IDS за индустриски Интернет на нешта (Industrial IoT) е имплементиран во [109] со користење на RF и генетски алгоритам (Genetic Algorithm), постигнувајќи точност на тест множеството од 87,61% за бинарна класификација со 16 карактеристики и точност на тест множеството од 77,64% за класификација со повеќе класи, со евалуација на податочното множество UNSW-NB15[110], притоа користејќи 17 атрибути.

Некои од истражувањата го разработуваат и проблемот на класна нерамнотежа во податочните множества. Во [111] авторите ги споредуваат перформансите на различни ML алгоритми врз податочното множество CICIDS2017 [112]. Со цел да се балансира податочното множество, авторите ја користат техниката за прекумерно земање примероци од синтетички малцинства (Synthetic Minority Oversampling TEchnique) [113], што ги подобрува перформансите за напади на малцинската класа. Сепак, постојат докажани случаи каде што SMOTE може да предизвика одредени проблеми со податоците [114].

4.3 Длабоко учење во системи за детекција на аномалии кај IoT

Длабокото учење (DL) дополнително ја проширува способноста на IDS за IoT преку обезбедување напредни алгоритми од ML кои можат ефективно да се справат со големиот обем и сложеност на податоци генерирани од IoT уредите. Традиционалните алгоритми од ML често бараат рачна екстракција на карактеристики, која може да одземе драгоцено време и истовремено да претставува предизвик, особено за податоците на IoT, кои може да бидат од различен тип и структура. Од друга страна, алгоритмите на DL се способни за автоматско учење и извлекување корисни функции од необработени податоци, што ги прави идеален избор за IDS во IoT. Суштината на DL е да се користат вештачки невронски мрежи (Artificial Neural Networks) за учење на шаблони и карактеристики директно од податоците, без потреба од експлицитна екстракција на карактеристики.

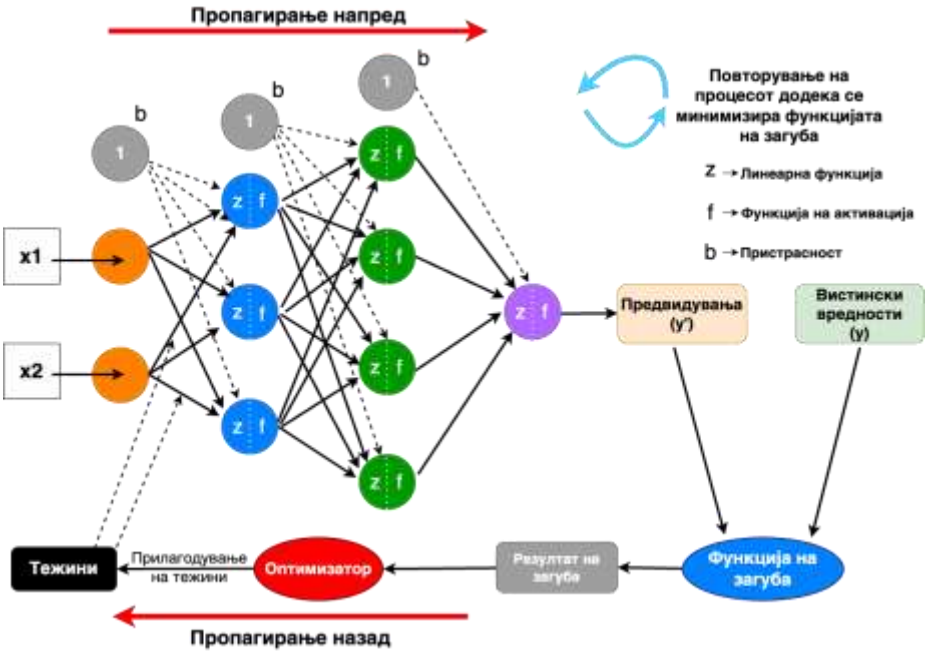
ANN се инспирирани од структурата и функцијата на биолошките неврони во човечкиот мозок. Тие се состојат од слоеви на меѓусебно поврзани јазли, наречени неврони, кои обработуваат влезни податоци и даваат предвидување или класификација. Секој неврон зема влез од претходниот слој, применува математичка операција на него и го пренесува резултатот на следниот слој (Слика 4.2.). Во типична невронска мрежа, влезниот слој прима необработени податоци, како слика или текст. Потоа, скриените слоеви во мрежата вршат низа нелинеарни трансформации на влезните податоци за да научат карактеристики и шаблони кои се релевантни за задачата што се анализира. На крај, излезниот слој произведува предвидување или класификација врз основа на научените карактеристики. ANN се карактеризираат со нивната длабочина, што значи дека најчесто имаат повеќе слоеви на неврони. Со зголемување на бројот на слоеви во мрежата, моделите за DL можат да учат сè покомплексни шаблони и карактеристики од податоците. Меѓутоа, подлабоките мрежи бараат и повеќе пресметковни ресурси и

подолго време за обука. За време на процесот на обука, мрежата ги прилагодува своите тежини (weights) за да ја минимизира функцијата на загуба (loss function), која ја мери разликата помеѓу предвидениот излез и вистинскиот излез. Овој процес на пропагирање назад (backward propagation) и овозможува на мрежата да учи од податоците и да ги подобри своите предвидувања со текот на времето.

Една од главните предности на DL во IDS за IoT е неговата способност да открие и да одговори на претходно невидени упади. Моделите за DL можат да научат да идентификуваат шаблони и аномалии во податоците кои можеби не се експлицитно дефинирани. Понатаму, моделите за DL, можат да се приспособат и да учат од нови податоци, што ги прави поробусни и способни за справување со нови закани. Друга предност на DL во IDS за IoT е неговата способност да се справи со обработка на податоци од големи размери. Алгоритмите за DL, како што се конволуциски невронски мрежи (Convolutional Neural Network) и рекурентните невронски мрежи (Recurrent Neural Network), можат да обработуваат и анализираат големи количини на податоци на скалабилен и ефикасен начин, што го олеснува следењето и откривањето на заканите низ бројни IoT уреди и мрежи.

Сепак, постојат и некои предизвици поврзани со користењето на DL во IDS за IoT. Еден од главните предизвици е потребата од значителна количина на означени податоци за ефективно да се обучуваат моделите за DL. Ова може да биде тешко да се постигне во пракса, особено за нови закани или напади кои сè уште не се идентификувани, или напади за кои не постојат јавно достапни податоци за обучување на модели. Понатаму, моделите за DL можат да бидат подложни на противнички напади, каде што напаѓачот намерно создава податоци за да ги доведе во заблуда предвидувањата на моделот. Ова е особено загрижувачко во контекст на IDS за IoT, каде што напаѓачите може да се обидат да ги искористат пропустите во мрежата или уредот за да добијат пристап до чувствителни податоци или да предизвикаат штета. И покрај овие предизвици, DL се карактеризира со модели кои идентификуваат комплексни шаблони во податоците и со тоа овозможува подобрување на ефикасноста на IDS за IoT, а со тековните истражувања продолжуваат да развиваат нови и уште поробустни алгоритми за справување со нив.

Во истражувањата за IDS базирани на DL може да се напомене искористеноста на најразлични методи како: CNN, мрежи за длабоко верување (Deep Belief Network), ограничени Болцман машини (Restricted Boltzmann Machine), RNN, длабоки невронски мрежи (Deep Neural Network) и сл. Трудот [115] претставува IDS заснован на DBN за IoT мрежа која детектира аномалии во мрежен сообраќај. Методот е евалуиран со користење на необработени податоци за мрежниот сообраќај. Во [116] авторите предлагаат IDS базиран на CNN користејќи достапно, добро воспоставено податочно множество. Пристапот користи CNN за да ги извлече карактеристиките на мрежниот сообраќај, пред да примени надгледувано учење за откривање на упадите. Постојат и IDS пристапи базирани на DNN кои користат трансформација и нормализација на податоците пред да ги внесат истите во самиот модел [117,118]. Во [119], авторите предлагаат нов IDS што користи CNN како модел за детекција на аномалии. Се покажува дека моделот ја зголемува точноста на класата додека истовремено ја намалува стапката на лажни тревоги.



Слика 4.2. Процес на длабоко учење. Мрежната архитектура се конфигурира така што се специфицираат бројот на слоеви, функции за активирање и други хиперпараметри. Следно, моделот се тренира со користење на алгоритам за оптимизација, како што е SGD, кој ги прилагодува тежините и пристрасноста на мрежата за да ја минимизира функцијата на загуба. Овој процес вклучува итеративно напојување на податоците за обука преку мрежата, пресметување на градиенти и ажурирање на параметрите. По обуката, перформансите на моделот се оценуваат со помош на податоци за валидација и може да се изврши оптимизација со цел подобрување на резултатите.

Отум и сор. [120], како и истражувачите во [121] исто така предлагаат IDS решенија базирани на DL кои се справуваат со аномалиите во IoT мрежи. Првото истражување користи пајак-мајмун алгоритам за оптимизација (Spider Monkey Optimization) и Stacked-Deep полиномна мрежа (Stacked Deep Polynomial Network), постигнувајќи точност од 99,02%, додека второто користи длабока невронска мрежа базирана на Q-учење со метод на зачувување на приватноста (Deep Q - Neural Network with Privacy Preservation) и постигнува точност од 93,74%. Шакил П и сор. [122] ја користат истата методологија на Deep-Q-Network, каде медицинскиот IoHT систем е анализиран од страна на DNN со цел да се откријат и елиминираат сите напади на злонамерен софтвер.

DeL-IoT [123] претставува рамка (framework) за откривање на упади кое што со помош на длабоки и stacked автоенкодери извлекува карактеристики од податоците. Како резултат, постигната е висока доверливост од 99,5%-99,9%, како и 91,04%-99,95% MCC. Аверсано и сор. [124] предлагаат пристап за откривање аномалии во IoT користејќи техники на DL. Како градежен блок за обука на DNN е искористен концептот на автоенкодери и применето е намалување на карактеристиките со цел подобрување на точноста на откривање. Абусита и сор. [125] предлагаат систем за откривање на упад во облак со помош на модифицирана верзија на автоенкодер наречен автоенкодер за отстранување на шум (denoising autoencoder) за ефикасно тренирање на DNN. Притоа, може да се забележи дека автоенкодерот за отстранување на шум ја подобрува точноста на откривање дури и при нецелосни информации. Покрај тоа, Абусита и сор. [126] предлагаат и проактивен колаборативен систем за откривање на злонамерен софтвер

користејќи го автоенкодерот за отстранување на шум за да се подобри точноста на класификацијата и откривањето во реално време.

Мустафа и сор. [127] предлагаат сеопфатна рамка за откривање на упади која интегрира различни алгоритми за машинско учење како што се NB, DT и ANN. Нивната рамка е специјално дизајнирана да открие напади на ботнет против протоколи поврзани со IoT, вклучувајќи MQTT, DNS и HTTP. Кале и сор. [128] предлагаат унифицирана рамка која интегрира три техники, вклучително и K-means кластерирање, GANomaly и CNN, за откривање аномалии во IoT мрежи. Слично на претходните две истражувања, Сарма [129] предлага рамка во две фази за откривање напади во IoT (фаза на екстракција и фаза на класификација на карактеристики).

Истражувањето во [130] има за цел да направи разлика помеѓу нормално и необично мрежно однесување преку користење на DL за да се минимизираат лажните аларми врз податочното множество NSL-KDD [131]. Притоа, постигната е стапка на точност од 87,2%. [132] и [133] дизајнираат IDS заснован на DL за откривање аномалии во сообраќај со повеќе класи користејќи ги податочните множества UNSW-NB15, CICIDS2017, Bot-IoT [134] и постигнувајќи точност од 96,69%-99,998%. [135] предлагаат користење на CNN1D, CNN2D и CNN3D длабоко учење за ефикасно откривање на повеќе напади во IoT мрежи со користење на неколку податочни множества, вклучувајќи ги: Bot-IoT, IoT-DS-2, IoT-23[136] и MQTT-IoT-IDS2020[137]. Во [138], авторите предлагаат шема за откривање на упади заснована на DL за IoT мрежи, чии резултати се блиску до 99% во сите метрики за евалуација и во бинарна и во класификација со повеќе класи.

Во [139] се предлага комбинација од RNN и CNN модели за безбедност на IoT мрежи. Притоа, постигнати се 96,32% точност, 95,74% F1 резултат, како и 95,43% прецизност. Улах и сор. [140] користат RNN напојувана со долга краткорочна меморија (Long Short Term Memory) и затворена повторлива единица (Gated Recurrent Unit) за да имплементираат систем за откривање на аномалии во IoT мрежи. Тие исто така користат CNN за да ги анализираат влезните карактеристики поврзани со IoT со цел за да задржат важни информации. Колошњаји и др. [141] комбинираат CNN со мрежи за LSTM за да ги истражат семејствата на упади.

Сакс и Берлин [142] предлагаат DNN за откривање на упади со дизајнирање на IDS напојуван со повеќеслојни перцептрони (Multi Layer Perceptron). Тие, исто така, користат Баесов модел за да ги проценат сомнителните активности и да идентификуваат вистински упади. Џоу и сор. [143] предлагаат интегриран модел за DL напојуван од генеративни противнички мрежи (Generative Adversarial Networks) наречени DB-CGAN. Тие користат противничка обука за да ја подобрат точноста на класификацијата и самото откривање.

4.4 Анализа на податочни множества за детекција на аномалии кај IoT

Во овој дел извршена е длабинска анализа на различни податочни множества за откривање на упади, притоа истакнувајќи ги нивните уникатни карактеристики, силни страни, како и ограничувања. Овие податочни множества служат како непроценливи ресурси за обука и тестирање на IDS моделите и обезбедуваат средства за подобрување на точноста на откривање и робусноста на нивните системи.

4.4.1 KDD CUP 1999

Податочното множество KDD Cup 1999 [131] е широко користено податочно множество во областа на откривање на упади и мрежна безбедност. Создадено е како дел од натпревар за рударење на податоци организиран од конференцијата ACM SIGKDD во 1999 година. Ова множество е наменето да обезбеди реална средина за тестирање и оценување на IDS.KDD Cup 1999 се состои од збир на податоци за мрежниот сообраќај, кој вклучува различни видови напади, како и нормален мрежен сообраќај. Всушност, множеството потекнува од податочното множество за евалуација за откривање на упад DARPA [144] од 1998 година, создадено од MIT Lincoln Labs. Ова множество содржи вкупно 4,9 милиони записи од симулирана средина која имитира воена мрежа, поделени на два дела: множество за обука и множество за тестирање. Множеството за обука содржи 2,5 милиони записи, додека множеството за тестирање содржи 1,4 милиони записи. Останатите записи се резервирани за други цели, како што се валидација или евалуација на перформанси. Податочното множество вклучува 42 различни карактеристики кои можат да се користат за дизајнирање на модели за откривање на упади, како што се: искористениот протокол, IP-адресите на изворот и одредиштето, броевите на изворните и одредишните порти и разни други карактеристики на ниво на мрежа и на транспортно ниво. Секој запис во податочното множество е означен како нормален или како еден од четирите типови на напади како што се: DoS, корисник до корен (User to Root), далечински до локален напад (Remote to Local) и напад со проби.

Иако KDD Cup 1999 е широко користено множество за евалуација на IDS, тоа е критикувано поради тоа што е премногу ограничено во опсегот и не е репрезентативно во однос на современите мрежни безбедносни закани. Исто така, податочното множество има неколку ограничувања. Едно од главните ограничувања е тоа што се заснова на симулирана мрежна средина, која можеби нема точно да ја одрази сложеноста на реалните мрежни средини. Дополнително, KDD Cup 1999 вклучува само ограничен број на типови на напади, кои можеби не го претставуваат целиот опсег на напади што може да се сретнат во сценарија од реалниот свет.

4.4.2 NSL-KDD

Податочното множество NSL-KDD [130] претставува обработена и рафинирана верзија на податочното множество KDD Cup 1999 и истото е развиено за да се надминат некои од ограничувањата на оригиналното податочно множество, бидејќи преку истражувања се откриле некои проблеми, вклучувајќи дупликати и ирелевантни записи и некои неточности во означувањето на типовите на напади. NSL-KDD е развиено од страна на истражувачите на Универзитетот во Њу Бранзвик и објавено во 2009 година. Податочното множество содржи 22 различни типови на напади, кои се класифицирани

во четири главни категории: DoS, напади со проба, U2R и R2L. Ова множество содржи вкупно 41 карактеристика, каде се вклучени атрибути и на ниво на мрежа и на транспортно ниво, како што се времетраење на врската, IP-адресите на изворот и одредиштето, броевите на изворната и одредишната порта и бројот на бајти и пакети разменети во секоја насока. Секој запис во множеството е означен како нормален или како еден од 22-те типови на напади.

Едно од главните подобрувања на NSL-KDD во однос на KDD Cup 1999 е тоа што се решава проблемот со преклопување и редундантни записи. Множеството вклучува и подмножество на записи кои се независни од множеството за обука, што може да се користи за тестирање на способноста за генерализирање на моделите за откривање на упад. NSL-KDD е широко користено како референтно податочно множество за евалуација на IDS и се користи во бројни истражувачки студии во областа на мрежната безбедност. Сепак, важно е да се забележи дека како и секое податочно множество, NSL-KDD има и ограничувања и можеби нема целосно да ја опфати сложеноста и разновидноста на заканите за безбедноста на мрежата од реалниот свет. Затоа, треба да се користи со претпазливост при оценување на перформансите на системите за откривање на упад во сценарија од реалниот свет.

4.4.3 USNW-NB15

Податочното множество UNSW-NB15 [110] е множество за откривање на мрежни упади развиено од истражувачи од Универзитетот во Нов Јужен Велс (UNSW) во Австралија во 2015 година. Дизајнирано е да биде сеопфатено и реално податочно множество за евалуација на IDS. UNSW-NB15 содржи вкупно 2,5 милиони записи за мрежен сообраќај, собрани од симулирана мрежна средина во период од пет месеци. Во него е вклучен широк опсег на различни типови на напади, како и нормален мрежен сообраќај. Податочното множество вклучува вкупно 49 карактеристики за секој запис, вклучувајќи атрибути на ниво на мрежа и транспорт, како што се користениот протокол, IP-адресите на изворот и одредиштето, броевите на изворната и одредишната порта и бројот на разменети бајти и пакети во секоја насока. Исто така, UNSW-NB15 вклучува и метаподатоци за секој запис, како што се време на поврзување и времетраење на врската. Типовите на напади во податочното множество се класифицирани во девет главни категории: Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Shell Code, Reconnaissance и Worms. UNSW-NB15 вклучува и неколку подкатегории за секој тип на напад, што го прави порепрезентативен за заканите за безбедност на мрежи од реалниот свет.

Една од уникатните карактеристики на UNSW-NB15 е тоа што вклучува разновидни различни типови на напади за секоја категорија на напади, што овозможува посеопфатна евалуација на IDS. Исто така, податочното множество вклучува и подмножество на записи кои се независни од множеството кое се користи за обука, што може да се користи за тестирање на способноста за генерализирање на моделите за откривање на упад. UNSW-NB15 е популарно и референтно податочно множество за евалуација на IDS и се користи во бројни истражувачки студии во областа на мрежната безбедност. Сепак, важно е да се забележи дека како и секое податочно множество, и UNSW-NB15 има свои ограничувања и можеби нема целосно да ја опфати сложеноста и различноста на заканите за безбедноста на мрежата од реалниот свет.

4.4.4 BoT-IoT

Податочното множество Bot-IoT [134] е множество за откривање на мрежни упади развиено од истражувачи од Универзитетот во Нов Јужен Велс (UNSW) во Австралија. Објавено е во 2019 година и е дизајнирано да им помогне на истражувачите да ја проценат ефективноста на IDS за IoT. Bot-IoT вклучува 7 милиони записи од мрежен сообраќај собрани од 5 IoT уреди (вклучувајќи паметни термостати, паметен фрижидер, паметно осветлување) кои симулираат типична мрежна IoT околина. Сообраќајот вклучува и нормален сообраќај и сообраќај од различни типови на напади, кои вообичаено се користат при напади на IoT уреди.

Податочното множество вклучува низа различни типови на напади, класифицирани во 3 главни категории: DoS, собирање на податоци и крадење на податоци. Исто така, Bot-IoT вклучува неколку подкатегории за секој тип на напад, што го прави порепрезентативен за заканите за безбедноста на IoT уредите од реалниот свет. Една од уникатните карактеристики на Bot-IoT е тоа што вклучува податоци од низа различни IoT уреди. Податочното множество е популарно и референтно евалуација на IDS за IoT уреди и е искористено во бројни истражувачки студии во областа на безбедност на IoT. Севкупно, Bot-IoT обезбедува вреден ресурс за истражувачите кои се заинтересирани да ја проценат ефективноста на IDS за IoT. Како и да е, како и секое податочно множество, Bot-IoT има свои ограничувања и можеби нема целосно да ја опфати сложеноста и разновидноста на заканите за безбедноста во IoT.

4.4.5 ToN-IoT

Податочното множество ToN-IoT [145] е множество за откривање на мрежни упади развиено од истражувачи од Универзитетот во Нов Јужен Велс (UNSW) во Австралија. Објавено е во 2020 година и е дизајнирано да им помогне на истражувачите да ја проценат ефективноста на IDS за IoT. ToN-IoT вклучува телеметриски податоци од различни извори како што се поврзани уреди, записи од Linux и Windows оперативните системи, како и индустриски IoT мрежен сообраќај. Податочното множество е означено и содржи голем број на напади и нивни подкатегории: DoS, DDoS, вбризгување податоци, ransomware, напад со лозинка, напад со користење на задна врата, скриптирање меѓу страници, скенирање и MITM. Една од уникатните карактеристики на ToN-IoT е тоа што вклучува реални податоци од низа различни IoT уреди. Податочното множество е популарно и референтно евалуација на IDS за IoT уреди и е искористено во бројни истражувачки студии во областа на безбедност на IoT. Севкупно, ToN-IoT обезбедува вреден ресурс за истражувачите кои се заинтересирани да ја проценат ефективноста на IDS за IoT.

4.4.6 N_BaIoT

N-BaIoT [146] е податочно множество создадено од група на истражувачи од Универзитетот во Калифорнија, Ирвин во САД. Ова податочно множество го покрива недостатокот на јавни ботнет податочни множества, особено за IoT доменот. Составено е од податоци од мрежен сообраќај во реално време собрани од девет комерцијални IoT уреди, вклучувајќи монитор за бебиња, безбедносни камери, веб-камера, ѕвончиња и термостат, кои се заразени од најчестите семејства на ботнет: Mirai и Bashlite. N-BaIoT се состои од 7.062.606 записи со 115 различни карактеристики, кои понатаму се поделени на 10 категории на напад: gafgyt_combo, gafgyt_junk, gafgyt_scan, gafgyt_tcp,

gafgyt_udp, mirai_ack, mirai_raudp, mirai_rai_ , како и бенигна категорија, која го содржи нормалниот проток на сообраќај на набљудуваните уреди.

4.4.7 CIC-IDS-2017

Податочното множество CIC-IDS2017 [147] е множество за откривање на упади развиено од страна на истражувачи од Канадскиот институт за сајбер безбедност. Објавено во 2017 година, CIC-IDS2017 е дизајнирано да им помогне на истражувачите да ја проценат ефективноста на IDS за мрежен сообраќај. Податочното множество вклучува податоци за мрежниот сообраќај снимени од низа симулирани сајбер напади, вклучувајќи DDoS напади, напади на ботнет и разни видови малициозен софтвер (Brute Force FTP, Brute Force SSH, DoS, Heartbleed, Web Attack, Infiltration, Botnet and DDoS). Сообраќајот е снимен со помош на мрежа за тестирање дизајнирана да симулира реална мрежна средина. CIC-IDS2017 содржи вкупно 80 карактеристики за секој запис, вклучувајќи атрибути на ниво на мрежа и транспортно ниво, како што се протоколот што се користи, IP-адресите на изворот и одредиштето, броевите на изворната и одредишната порта и бројот на бајти и пакети разменети во секоја насока. Исто така, множеството вклучува карактеристики поврзани со сообраќајот на апликациско ниво, како што се заглавјата на HTTP барања и одговори.

4.4.8. CSE-CIC-IDS2018

Податочното множество CSE-CIC-IDS2018 [148] е множество за откривање на упади развиено од истражувачи од Канадскиот институт за сајбер безбедност. Објавен е во 2018 година и е дизајниран да им помогне на истражувачите да ја проценат ефективноста на IDS за мрежниот сообраќај. CSE-CIC-IDS2018 вклучува податоци за мрежниот сообраќај снимени од низа симулирани сајбер напади, вклучувајќи DDoS напади, напади на ботнет и разни видови малициозен софтвер. Множеството содржи вкупно 79 карактеристики за секој запис, вклучувајќи атрибути на ниво на мрежа и транспортно ниво, како што се протоколот што се користи, IP-адресите на изворот и одредиштето, броевите на изворната и одредишната порта и бројот од бајти и пакети разменети во секоја насока. исто така, вклучува карактеристики поврзани со сообраќајот на слојот на апликацијата, Исто така, множеството содржи карактеристики поврзани со апликациското ниво, како што се заглавјата на HTTP барања и одговори. Нападите во CSE-CIC-IDS2018 се класифицирани во неколку категории, вклучувајќи DoS напади, скенирање на порти, веб-напади, инфилтрација и напади на ботнет. Податочното множество вклучува и неколку подкатегории за секој тип на напад, што го прави порепрезентативно за заканите за безбедност на мрежата од реалниот свет.

4.4.9 IoT-23

Податочното множество IoT-23 [136] е множество за откривање на упади развиено од Stratosphere Laboratory, AIC group, FEL, CTU University од Чешка. Објавено е во 2020 година година и е дизајнирано да им помогне на истражувачите да ја проценат ефективноста на IDS за мрежниот сообраќај во IoT. Истражувањето е подржано од Avast Software со седиште во Прага. IoT-23 се состои од дваесет и три сценарија на различен мрежен IoT сообраќај. Овие сценарија се поделени на дваесет сценарија со заразени IoT уреди (запишани со име на примерокот од злонамерен софтвер извршен на секое сценарио) и три мрежни сценарија на мрежниот сообраќај од вистински IoT уреди (кои го носат името на уредите). Важно е да се спомене дека овие три IoT уреди се вистински хардвер и не се симулирани.

4.4.10 IoTID20

Податочното множество IoTID20 [107] претставува типична средина на паметен дом, која вклучува камера, паметен телефон, домашен звучник и неколку компјутери. Со следење на мрежниот сообраќај и следење во различни временски периоди, извлечени се 83 мрежни карактеристики од уредите со користење на Wireshark [149]. Дистрибуцијата помеѓу нормалниот мрежен сообраќај и различните типови на аномалии е следна: категоријата нормален сообраќај претставува околу 28%, во споредба со остатокот од податочното множество. Остатокот претставуваат напади кои вообичаено може да се случат во IoT околина, како што се: поплави на Mirai UDP (183.554 примероци или 22,5%), Mirai Hostbruteforce (121.181 примероци или 14,8%), DoS SYN Flooding (7.3%), Mirai HTTP Flooding (55.818 примероци или 6,8%), Mirai ACK Flooding (55.124 примероци или 6,8%), Scan Port OS (53.073 примероци или 6,5%), MITM ARP Spoofing (4,3%), Scan Hostport (22.192 примероци или 2,8%) и Bot (1.966 примероци или 0,2% од целото множество).

Процесот на селекција на податоци за развој на пристап за откривање на аномалии во мрежниот сообраќај во IoT мрежите вклучуваше избор на две податочни множества: IoTID20 и N-BaIoT. Овие податочни множества беа специјално избрани за целите на самите експерименти. Подетални информации за процесот на селекција може да се најдат во Глава 6, каде истиот е дополнително објаснет и детално дискутиран.

Табела 4.1. Збирна табела која ги прикажува податочните множества употребени во развојот на IDS. За секое податочно множество се вклучени: годината на објавување, целта, големината, бројот на карактеристики, како и видот на базата на податоци, методот на собирање и видовите на напади.

Податочно множество	**Година**	**Цел**	**Големина**	**Тип**	**Метод на собирање на податоци**	**Број на карактеристики**	**Типови напади**
KDD Cup 1999	1999	Мрежен упад	4,898,431	Мрежен сообраќај	симулирано	42	DoS, User to Root, Remote to Local и Probe attack
NSL-KDD	2009	Мрежен упад	125,973	Мрежен сообраќај	симулирано	41	22 различни типови напади, класифицирани како DoS, напади за вежбање, U2R & R2L
UNSW-NB15	2015	Мрежен упад	2,540,045	Мрежен сообраќај	симулирано	49	Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Shell Code, Reconnaissance & Worms
BoT-IoT	2019	IoT безбедност	7,000,000	IoT сообраќај	симулирано	34	Различни типови напади класифицирани во три главни категории:DoS, data harvesting и data theft.

TON-IoT	2020	IoT безбедност	22,000,000	IoT сообраќај	реално	46	DoS, DDoS, data injection, ransomware, password attack, backdoor attack, cross-site scripting, scanning и MitM.
N-BaIoT	2018	IoT безбедност	7,062,606	IoT сообраќај	лабораторија	115	gafgyt_combo, gafgyt_junk, gafgyt_scan, gafgyt_tcp, gafgyt_udp, mirai_ack, mirai_raudp, mirai_rai_ , како и benign категорија
CIC-IDS2017	2017	Мрежен упад	2,242,376	Мрежен сообраќај	симулирано	80	Brute Force FTP, Brute Force SSH, DoS, Heartbleed, Web Attack, Infiltration, Botnet & DDoS
CSE-CIC-IDS2018	2018	Мрежен упад	20,000,000	Мрежен сообраќај	симулирано	80	Опсег од симулирани напади, вклучувајќи DDoS attacks, botnet attacks и различни типови на malware
IoT-23	2020	IoT безбедност	варијабилен	IoT сообраќај	реално	варијабилен	Дваесет сценарија на инфицирани IoT уреди (означени со името на примерокот за малициозен софтвер извршен во

							секое сценарио) и три мрежни сценарија на реален мрежен сообраќај од IoT уреди (именувани по уредите).
IoTID20	2020	IoT сообраќај	816816	IoT сообраќај	лабораторија	83	Mirai UDP Flooding, Mirai Hostbruteforce , DoS SYN Flooding, Mirai HTTP Flooding, Mirai ACK Flooding ,Scan Port OS, MITM ARP Spoofing, Scan Hostport & Bot

Глава 5 Федеративно машинско учење

Машинското учење и длабокото машинско учење се две техники кои направиле револуција во начинот на решавање на комплексни проблеми. Тие се потпираат на големи количини на податоци за обука на модели кои можат да прават прецизни предвидувања или одлуки. Сепак, традиционалните ML и DL методи претпоставуваат дека сите податоци се достапни на централизирана локација, што може да создаде значителни грижи за приватноста и безбедноста кога се работи со чувствителни податоци. Федеративното учење (FL) [150] ги решава овие проблеми со овозможување на дистрибуирани модели на машинско учење кои можат да се обучуваат за локални податоци без да се загрози приватноста на корисниците.

FL е децентрализиран ML пристап кој вклучува модели за обука на повеќе уреди наместо централизирање на сите податоци на еден сервер. Овој пристап вклучува централен сервер кој управува со целокупниот процес на учење, но вистинската обука се одвива на поединечни уреди, како што се паметни телефони или други IoT уреди. Главната предност на FL е тоа што ја зачувува приватноста на корисниците додека овозможува агрегирање на ажурирања на модели од повеќе уреди за да се создаде попрецизен глобален модел. Ова го прави особено корисен во ситуации кога приватноста на податоците е критична, како на пример во здравството, финансиите и другите индустрии кои се занимаваат со чувствителни податоци. Една од главните разлики помеѓу FL и традиционалните ML и DL методи е дистрибуцијата на податоците. Во традиционалните ML и DL методи, податоците се централизирани и сите модели се обучуваат на истото податочно множество. Во FL, податоците се наоѓаат на повеќе уреди, а моделите се обучуваат локално на секој уред. Ова создава неколку придобивки, како што е можноста за учење од поразновидни податочни множества и способноста за обука на модели на податоци што не се достапни на централизирана локација.

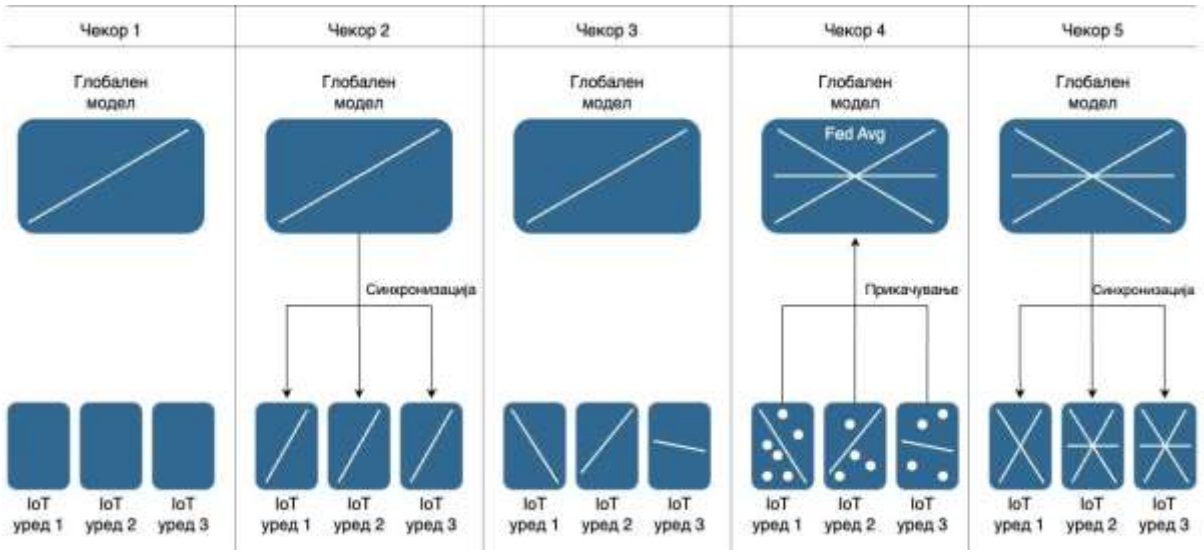
Друга придобивка на FL е тоа што може да го намали количеството на податоци што треба да се пренесат преку мрежата, што ја намалува количината на потребната пропусност и да ги подобрува севкупните перформанси на системот. Наместо да се испраќаат сите податоци на централен сервер, се пренесуваат само ажурирања на моделот, што може значително да ја намали количината на податоци што треба да се пренесат. FL исто така нуди неколку придобивки во однос на безбедноста и приватноста. Со чување податоци на локалните уреди, FL може да спречи пробивање на податоците и други безбедносни закани кои можат да произлезат од централизираното складирање. Исто така, FL гарантира дека податоците остануваат приватни и безбедни, дури и за време на процесот на обука, бидејќи се пренесуваат само ажурирањата на моделот, а не самите необработени податоци.

Сепак, FL има и некои ограничувања и предизвици на кои мора да им се обрне внимание со цел овој метод да биде остварлива алтернатива на традиционалните методи за ML и DL. Еден предизвик е хетерогеноста на уредите и податоците во системот. Бидејќи податоците се дистрибуираат на повеќе уреди, самите уреди може да имаат различни хардверски и софтверски конфигурации. Тоа значи дека ќе биде комплицирано да се обезбеди гаранција дека дека моделите се обучуваат доследно на сите уреди. Друг предизвик е потенцијалот за пристрасност во FL системите. Бидејќи податоците се дистрибуираат на повеќе уреди, може да биде тешко да се осигура дека истите се

репрезентативни за целата популација. Ова може да доведе до пристрасност во податоците за обуката, што може да резултира со пристрасни модели. И покрај овие предизвици, FL има потенцијал да стане моќна алатка за решавање на сложени проблеми додека ја зачувува приватноста и безбедноста на корисниците.

5.1 Процес на федеративно учење

На Слика 5.1. детално е прикажан процесот на федеративното учење, како и неговите составни компоненти и поединечни процеси.



Слика 5.1. Федеративен процес на учење. Наместо централизирање на податоците на една локација, моделите се обучуваат локално на секој уред користејќи локални податоци. Овие модели специфични за уредите потоа заеднички се агрегираат во глобален модел од централен сервер. Овој итеративен процес им овозможува на уредите да учат едни од други додека ја зачувуваат приватноста на поединечните податоци. Со минимизирање на преносот и обработката на податоци на серверот, федеративното учење ги намалува трошоците за комуникација и ги решава проблемите со приватноста.

IoT уреди - клиенти: Ова се уредите што учествуваат во FL системот, како што се паметни телефони, лаптопи или IoT уреди. Секој клиент има свои локални податоци и обучува модел за машинско учење на нив.

Федеративен сервер: Ова е централен сервер кој го координира целокупниот процес на учење. Тој управува со ажурирањата на моделот и обезбедува агрегирање на локалните модели од клиентските уреди на едно место со цел да се создаде глобален модел.

Алгоритам за федеративно учење: Алгоритмот за федеративно учење е јадрото на FL системот, бидејќи одредува како локалните модели од уредите на клиентите се собираат за да се создаде еден глобален модел. Алгоритмот мора да биде дизајниран така што глобалниот модел е точен и репрезентативен за локалните податоци од уредите на клиентите.

Локална обука на модели: Обуката за модели е процес на обука на модел за машинско учење на локалните податоци од секој клиентски уред. Секој клиент уред тренира модел локално, а моделите потоа се агрегираат за да се создаде глобален модел.

Агрегација на модели: Агрегација на модели е процес на комбинирање на локалните модели од уредите на клиентот за да се создаде глобален модел. Овој процес мора да биде внимателно дизајниран за да се осигура дека глобалниот модел е точен и репрезентативен за локалните податоци од уредите на клиентите.

Локалната обука на моделите, како и агрегацијата на моделите се процеси кои се повторуваат во повеќе рунди (rounds). Во секоја рунда, уредите вршат обука на локалните модели, генерираат ажурирања на моделот и ги испраќаат до централниот сервер за агрегација. Во текот на рундите, глобалниот модел постепено се подобрува со агрегирање на знаењето од различните IoT уреди.

Евалуација на моделот: Евалуацијата на моделот вклучува тестирање на глобалниот модел за да се осигура дека е точен и репрезентативен за податоците. Ова вклучува користење на податоци за валидација за тестирање на перформансите на глобалниот модел и нивно споредување со перформансите на локалните модели на секој клиентски уред.

5.2 Диференцијална приватност

Диференцијалната приватност [151] претставува моќна алатка за решавање на проблемите со приватноста во контекст на FL. Со интегрирање на диференцијални техники за приватност во FL, овој пристап нуди принципиелен и ригорозен начин за заштита на чувствителните податоци на корисниците, притоа овозможувајќи ефективна обука за колаборативен модел. Диференцијалната приватност обезбедува присуството или отсуството на податоци на поединецот да не влијае значително на резултатите или заклучоците извлечени од обучениот модел. Обезбедува математички квантитабилна гаранција за приватност со воведување контролирана бучава (noise) или нарушувања на ажурирањата на моделот или збирните резултати за време на федеративниот процес на учење.

Една од клучните предности на диференцијалната приватност во FL е способноста да ја заштити приватноста на индивидуалните корисници. Во традиционалното FL, уредите или клиентите ги доставуваат своите локални податоци до централниот сервер за обука на моделот, што предизвикува загриженост за приватноста на податоците и потенцијалното истекување на чувствителни информации. Диференцијалната приватност ги ублажува овие грижи со вбризгување на внимателно калибриран шум во ажурирањата на моделот или збирните резултати. Овој шум ги замаглува индивидуалните придонеси, што ја прави идентификацијата на конкретни кориснички податоци во обучениот модел исклучителен предизвик, со што се зачувува приватноста на корисникот. Исто така, диференцијалната приватност нуди мерка на робусност против различни напади на приватност. Со гарантирање на приватноста дури и кога противникот има знаење за податоците за обука или пристап до самиот модел, диференцијалната приватност обезбедува дополнителен слој на заштита од потенцијални прекршувања на приватноста. Понатаму, диференцијалната приватност во FL го олеснува зголеменото споделување на податоци и соработка меѓу повеќе организации или ентитети. Таа им овозможува на организациите да соработуваат додека

ги заштитуваат нивните индивидуални податоци, поттикнувајќи учество и колективно учење без загрозување на приватноста.

Сепак, вклучувањето на диференцијалната приватност во FL доаѓа со некои резерви. Воведувањето шум за заштита на приватноста може да влијае на корисноста или перформансите на обучените модели. Балансирањето на приватноста и точноста на моделот станува критичен компромис и потребно е внимателно подесување на параметрите за да се постигне соодветна рамнотежа. Дополнително, пресметковните трошоци поврзани со диференцијалните техники за приватност може да бидат значајни. Механизмите за вбризгување бучава и за зачувување на приватноста ја зголемуваат пресметковната сложеност и времето за обука.

5.3 Алгоритми за федеративно учење

Областа на федеративното учење е сеуште нова и забрзано се развива, а истражувачите континуирано истражуваат нови алгоритми и техники со цел да се справат со уникатните предизвици што ги поставуваат IoT мрежите. Притоа, изборот на соодветен алгоритам ќе зависи од фактори како што се: мрежната архитектура, ограничувањата на ресурсите, барањата за приватност и специфичните случаи на употреба на IoT апликациите. Во прилог се наведени неколку алгоритми кои би можеле да одговорат на ваквите предизвици. Со истите ќе биде темелно експериментирано со цел да се развие пристапот за откривање аномалии во IoT мрежите:

5.3.1 Federated Averaging

Federated Averaging (FedAvg) [150] е еден од најпознатите и најчесто користените алгоритми во федеративното учење. Една значајна предност на користење на овој алгоритам лежи во неговата комуникациска ефикасност. Со пренесување само ажурирања на моделот наместо необработени податоци, се минимизираат трошоците за комуникација помеѓу IoT уредите и централниот сервер. Ова не само што ги намалува барањата за пропусен опсег, туку и заштедува енергија, што го прави добро прилагоден за околини ограничени со ресурси, како што се IoT. Дополнително, овој алгоритам е флексибилен и овозможува прилагодување на различни сценарија. Хиперпараметрите, како што се стапката на учење и процентот на клиенти кои се користат при процесот на обука може да се подесуваат за да се оптимизираат перформансите. Покрај тоа, предложени се адаптивни методи за прилагодување на хиперпараметрите врз основа на карактеристиките на уредот, со што се приспособуваат варијациите во дистрибуцијата на податоците и достапноста на ресурсите.

Сепак, важно е да се забележи дека FedAvg не е без ограничувања. Овој алгоритам директно не вклучува механизми за приватност. Но, тука може да се користат дополнителни техники како безбедно собирање со помош на криптографски протоколи за да се обезбеди приватноста на поединечните ажурирања на моделите за време на собирањето или пак да се користи диференцијална приватност. Исто така, FedAvg претпоставува доверливи и секогаш поврзани уреди, што можеби нема да важи во динамичните IoT мрежи. Ракувањето со хетерогеноста на уредите, нееднаквите пресметковни способности и несигурното поврзување остануваат отворени предизвици за FedAvg и федеративното учење како целина. Покрај овој алгоритам, предложени се и негови екстензии и варијации за справување со специфични предизвици.

5.3.2 Federated Median

Алгоритмот Federated Median (FedMedian) [152] е иновативен пристап во доменот на FL кој воведува нов механизам за агрегација наречен федеративна медијана. Овој алгоритам има за цел да се справи со ограничувањата на традиционалните методи на собирање, како што е федеративниот просек (federated averaging) и нуди подобрувања во робусноста и зачувувањето на приватноста. Во традиционалното FL, најчесто користениот метод на собирање е FedAvg, кој го пресметува просекот на ажурирањата на моделот од уредите што учествуваат. Сепак, овој метод е чувствителен на исклучоци (outliers), бидејќи дури и еден уред со значително различни податоци или уред кој не функционира може да има непропорционално влијание врз глобалниот модел. Оваа чувствителност на исклучоци (outliers) може да доведе до помалку робустен и точен глобален модел. За да се справи со ова, FedMedian го воведува концептот на федеративна медијана како алтернативен механизам за агрегација. Наместо да го пресметува просекот, тој ја пресметува медијаната на ажурирањата на моделот добиени од клиентите. Со преземање на медијаната, FedMedian го намалува влијанието на екстремните или злонамерните ажурирања, што доведува до поцврст глобален модел.

Една од клучните предности на FedMedian е неговото подобрено зачувување на приватноста во споредба со FedAvg, кој предизвикува загриженост во поглед на приватноста бидејќи централниот сервер има пристап до поединечните ажурирања, што потенцијално ги изложува чувствителните информации. Спротивно на тоа, FedMedian бара само секој уред да го сподели своето средно ажурирање, кое ги крие вистинските вредности и обезбедува понадежни гаранции за приватност. FedMedian исто така нуди придобивки во сценарија каде што податоците не се IID (неидентично и независно дистрибуирани) низ уредите што учествуваат. Во федеративното учење, уредите може да имаат различна дистрибуција на податоци поради варијации во корисничките параметри или други фактори. Традиционалните методи на агрегација, како што е FedAvg, може да се мачат да ракуваат со податоци се не-IID бидејќи претпоставуваат дека податоците се рамномерно распределени. FedMedian, со својата робусност кон оддалечените, може подобро да се справи со хетерогеноста во распределбата на податоците и да произведе попрецизен и порепрезентативен глобален модел.

Сепак, и алгоритмот FedMedian доаѓа со свои недостатоци. Пресметковните трошоци за пресметување на медијаната може да бидат повисоки од оние на просекот, особено кога се работи за модели од големи размери или голем број уреди кои учествуваат. Дополнително, алгоритмот претпоставува дека поголемиот дел од уредите кои учествуваат се искрени и обезбедуваат сигурни ажурирања. Ако значителен број уреди се злонамерни или намерно обезбедуваат противнички ажурирања, FedMedian може да биде подложен на нивното влијание.

5.3.3 Federated Proximal

Federated Proximal (FedProx) [153] алгоритмот е екстензија на FedAvg и нуди подобрувања во справувањето со предизвиците на не-IID податоци во FL за IoT мрежи. Една од клучните разлики помеѓу FedAvg и FedProx лежи во нивниот пристап кон ракување со податоци кои се не-IID. Додека FedAvg ги собира ажурирањата на локалните модели преку правење на едноставен просек, FedProx воведува термин за регулација што ги казнува отстапувањата од глобалниот модел. Со поттикнување на локалните модели да бидат слични на глобалниот модел, FedProx го ублажува влијанието на уредите со значително различни дистрибуции на податоци и ја подобрува

целокупната конвергенција на федеративниот процес на учење. Степенот на регулација може да биде прилагоден за да одговара на специфичните барања на IoT апликациите. Оваа флексибилност му овозможува на FedProx да се справува со сценарија каде што преовладуваат не-IID податоци и ги подобрува можностите за генерализација на наученият глобален модел.

Друга разлика помеѓу FedAvg и FedProx е ефектот врз брзината на конвергенција на федеративниот процес на учење. Поради терминот за регулација, FedProx има тенденција да конвергира побрзо во споредба со FedAvg. Ова е особено поволно во околини ограничени со ресурси, каде што минимизирањето на бројот на комуникациски рунд и заштедата на енергија се од големо значење. Понатаму, FedProx ги одржува придобивките од агрегацијата за зачувување на приватноста што е присутна во FedAvg. Ажурирањата на моделот сè уште се пренесуваат од IoT уредите до централниот сервер, обезбедувајќи приватност и безбедност на податоците. Тука може да се користат техники како безбедна агрегација за да се заштити доверливоста на поединечните ажурирања на моделите за време на процесот на агрегација. Додека FedProx нуди значителни подобрувања во однос на FedAvg во ракувањето со податоци кои се не-IID, тој не е без ограничувања. Изборот на терминот за регулација тежина бара внимателно разгледување и подесување за да се постигне саканата рамнотежа помеѓу перформансите на глобалниот модел и локалното вклопување на моделот. Дополнително, FedProx исто така претпоставува присуство на сигурни и секогаш поврзани уреди, што можеби нема да биде случај во динамичните IoT мрежи.

5.3.4 Federated Yogi

Federated Yogi (FedYogi) [154] е алгоритам за оптимизација дизајниран да го подобри популарниот Адам оптимизатор. Тој ги комбинира техниките за адаптивна стапка на учење со придобивките од методите за оптимизација базирани на моментум. FedYogi има за цел да постигне побрза конвергенција и подобра генерализација во обуката на ANN. FedYogi ги прилагодува стапките на учење по параметар, слично на Адам. Одржува приспособливи стапки на учење по параметар врз основа на историјата на ажурирања на градиент и градиенти во квадрат. Ова му овозможува на FedYogi динамично да ги прилагодува стапките на учење на специфичните барања на секој параметар за време на тренингот.

Покрај адаптивните стапки на учење, FedYogi го вклучува механизмот на импулсот за забрзување на конвергенцијата. Тој воведува адаптивна регулација за да ги земе предвид варијациите во скалата на градиентите. Ја прилагодува силата на регуларизација со разгледување на историските големини на градиентите и нивните квадратни градиенти. Оваа адаптивна регулација има за цел да ја балансира важноста на регуларизацијата низ различни параметри и да ја прилагоди врз основа на карактеристиките на процесот на оптимизација. FedYogi користи механизам за поправен просек за време на ажурирањата на параметрите за да се спречи прекумерното распаѓање на стапката на учење. Го поправа процесот на просечно мерење со разгледување и позитивни и негативни промени во вредностите на параметарот. Оваа исправка му овозможува на Јоги да одржува подобро следење на состојбата на параметрите и соодветно да ги приспособи стапките на учење.

Овој оптимизатор нуди неколку предности во однос на традиционалните алгоритми за оптимизација. Со комбинирање на адаптивните стапки на учење со адаптација на импулсот, FedYogi го олеснува побрзото конвергенција за време на процесот на обука. Адаптивната регулација на FedYogi и поправениот просек придонесуваат за подобра генерализација во обуката за длабока невронска мрежа. Адаптивната регулација ја прилагодува јачината на регуларизација врз основа на скалата на градиенти, овозможувајќи поурамнотежена регулација низ различни параметри. Поправениот механизам за просечно одредување го спречува прекумерното распаѓање на стапката на учење, што може да доведе до прекумерна регулација и да ја попречи генерализацијата. FedYogi покажува робусност за избор на хиперпараметри, што го прави помалку чувствителен на специфичните вредности на хиперпараметрите како што се брзината на учење и силата на регулација. Оваа робусност го олеснува товарот на подесување на хиперпараметри и обезбедува постабилни перформанси во различни задачи и архитектури.

5.3.5 Federated Adagrad

Оптимизаторот Adagrad (Adaptive Gradient) [154] е алгоритам за оптимизација специјално дизајниран за задачи за машинско учење. Познат е по својата способност автоматски да ја приспособи стапката на учење за секој параметар во модел базиран на историски градиенти. Adagrad добро функционира во сценарија каде што податоците имаат ретки градиенти или различни фреквенции на карактеристики. Adagrad одржува збир на квадратни градиенти за секој параметар за време на процесот на обука. Тоа го прави со собирање на квадратот на секој градиент во однос на елементот. Оваа акумулација им дава поголема тежина на параметрите со ретки ажурирања или големи градиенти. Оптимизаторот ги прилагодува стапките на учење врз основа на акумулираните градиенти на квадрат. Ја дели брзината на учење со квадратниот корен од збирот на квадрати градиенти за секој параметар. Ова резултира со поголема стапка на учење за параметри со помали градиенти и помала стапка на учење за параметри со поголеми градиенти.

Механизмот за адаптивна стапка на учење му овозможува на Adagrad ефикасно да се справува со ретки градиенти. Со Adagrad, секој параметар во моделот има своја стапка на учење, динамички прилагодена врз основа на неговите историски градиенти. Оваа адаптација специфична за параметарот му овозможува на Adagrad автоматски да доделува соодветни стапки на учење на различни параметри, подобрувајќи ја севкупната ефикасност на процесот на оптимизација. Со акумулирање на квадратни градиенти, Adagrad обезбедува имплицитна форма на регулација. Параметрите кои имаат големи акумулирани квадратни градиенти ќе доживеат распаѓање на стапката на учење, што ефективно ја намалува стапката на учење како што напредува обуката. Оваа карактеристика помага да се спречи моделот да го надмине оптималното решение. Механизмот за адаптивна стапка на учење на Adagrad ја елиминира потребата за рачно подесување на хиперпараметарот на стапката на учење, односно ги прилагодува стапките на учење на основа по параметар.

Сепак, Adagrad има некои ограничувања - акумулацијата на квадратни градиенти може да доведе до распаѓање на стапката на учење со текот на времето, што може да предизвика процесот на учење да стане премногу бавен, особено во архитектурите на длабоки невронски мрежи. За да се реши ова, развиени се алтернативни оптимизатори како што се Adadelta и RMSprop. Како што Adagrad акумулира квадратни градиенти со

текот на времето, барањата за меморија се зголемуваат. Ова може да стане проблем кога се тренираат модели од големи размери со милиони параметри, бидејќи може да ги исцрпи достапните ресурси за меморија.

5.3.6 Federated Averaging со толеранција на грешки

FedAvg со толеранција на грешки (FTFedAvg) [155] е важен напредок во областа на федеративно учење кој ги адресира предизвиците што ги поставуваат несигурните и неисправните уреди во една околина за FL. Овој алгоритам обезбедува робустен и еластичен пристап за собирање на ажурирања на моделот од уредите кои учествуваат, дури и во присуство на неуспеси или противничко однесување. Една од клучните предности на FTFedAvg е неговата способност ефикасно да се справува со неисправни уреди. Во федеративното учење, уредите може да доживеат различни проблеми, како што се дефекти на мрежата, ниско ниво на батерија или дефекти во софтверот, што може да резултира со несигурни или погрешни ажурирања на моделот. FTFedAvg користи техники за толеранција на грешки за да го ублажи влијанието на таквите неисправни уреди. Тоа го постигнува со вклучување на механизми за редундантност и за откривање грешки за време на процесот на агрегација. Со разгледување на повеќекратни непотребни ажурирања и откривање и филтрирање на погрешни ажурирања, FTFedAvg може да го задржи интегритетот и точноста на глобалниот модел. Механизмите за толеранција на грешки во FTFedAvg ја подобруваат робусноста на федеративните системи за учење. Со идентификување и исклучување на неисправни ажурирања, алгоритмот го намалува потенцијалот за оштетени или погрешни информации да влијаат на процесот на агрегација.

FTFedAvg исто така го адресира и безбедносниот аспект на FL. Во сценарија каде што противниците може да се обидат да манипулираат или да внесат злонамерни ажурирања во процесот на агрегација, механизмите за толеранција на грешки на FTFedAvg играат клучна улога. Со инкорпорирање на техники за откривање грешки и редундантност, алгоритмот може да открие и да го ублажи влијанието на непријателското однесување, а со тоа да ја подобри безбедноста и интегритетот на федеративниот систем за учење. Понатаму, FTFedAvg придонесува за ефикасноста и приспособливоста на федеративното учење. Техниките за толеранција на грешки употребени во алгоритмот овозможуваат приспособливи и динамични ажурирања, овозможувајќи му на системот да ракува со голем број уреди и да приспособи уреди со различни нивоа на доверливост. Оваа приспособливост осигурува дека федеративното учење може ефективно да се распореди во сценарија од реалниот свет со различни карактеристики на уредот и мрежни услови. Сепак, важно е да се забележи дека FTFedAvg не е алгоритам без недостатоци. Дополнителните пресметковни трошоци потребни за спроведување на механизмите за толеранција на грешки може да влијаат на целокупното време за обука и искористување на ресурсите. Балансирањето на толеранцијата на грешки со ефикасноста е критичен аспект што треба да биде земен во предвид при примена на FTFedAvg во практични услови.

5.3.7. Federated Averaging со диференцијална приватност

Алгоритмот FedAvg со диференцијална приватност (Differential Privacy FedAvg) [156] е екстензија на традиционалниот алгоритам FedAvg кој вклучува техники за диференцијална приватност за да обезбеди гаранции за приватност за индивидуалните кориснички податоци во поставките за FL. FedAvg има за цел да обучи глобален модел

со агрегирање на ажурирања на локални модели од повеќе децентрализирани уреди, притоа зачувувајќи ја приватноста на податоците. Меѓутоа, без дополнителни механизми за приватност, при процесот на агрегација може ненамерно да протечат чувствителни информации од индивидуалните податоци на клиентот, што претставува ризик за приватноста. Диференцијалната приватност се справува со овие проблеми со воведување контролиран шум за време на процесот на агрегација на модели. Овој шум помага да се обезбеди дека присуството или отсуството на какви било конкретни податоци на корисникот нема значително влијание врз исходот на збирниот модел, обезбедувајќи математички квантитабилна гаранција за приватност. Диференцијалната приватност се постигнува со додавање на шумот на ажурирањата на моделот пред агрегација. Овој шум ги маскира индивидуалните придонеси на клиентот, што го прави предизвик да се извлечат конкретни податоци за корисникот од збирниот модел. Исто така, диференцијалната приватност овозможува соработка меѓу повеќе учесници додека се одржува приватноста. Организациите или субјектите можат да ги дадат своите податоци за модел на обука без да ги споделуваат вистинските податоци, да поттикнуваат соработка и колективно учење.

Сепак, додавањето шум за заштита на приватноста може да влијае на корисноста и перформансите на обучениот модел. Потребно е внимателно подесување на параметрите за да се постигне рамнотежа помеѓу приватноста и точноста на моделот. Потоа, воведувањето на механизми за зачувување на бучавата и приватноста ја зголемува комплексноста на пресметките и времето за обука, што потенцијално влијае на приспособливоста на алгоритмот. Неопходни се оптимизации и ефикасни алгоритми за да се ублажат овие предизвици.

5.4 Федеративно учење во системи за детекција на аномалии кај IoT

Во овој дел е истражен потенцијалот на иновативната комбинација на FL и IDS, кој претставува револуција во областа на безбедноста во IoT мрежите. Со искористување на моќта на децентрализираното ML и откривање закани во реално време, овој пристап нуди ветувачки решенија за заштита на уредите и податоците на IoT. Синергијата помеѓу FL и IDS во IoT мрежите нуди неколку предности. FL ја подобрува прецизноста на моделот со користење на податоци од широк опсег на IoT уреди, вклучувајќи ги и уредите со ограничени ресурси. Понатаму, FL обезбедува приватност на податоците со минимизирање на потребата за нивно споделување. Сепак, предизвиците како што се трошоците за комуникација, хетерогеноста на уредите и обезбедувањето правичност и робусност на моделот кај различни уреди мора да се решат со цел успешно распоредување на IDS базирани на FL во IoT мрежите.

Авторите во [157] како решение го предлагаат LocKedge, кој претставува IDS базиран на FL за IoT мрежи, кој открива аномалии на работ (edge) од мрежата. Сепак, кога се оценува на податочното множество BoT-IoT, може да се воочи дека FL моделот постигнува пониски перформанси од DL моделот со кој се споредува. Рахман и сор. [158] се обидуваат да ја задржат приватноста на податоците, предлагајќи нов систем базиран на FL за откривање на упади во IoT мрежи. Сепак, процесот на евалуација на податочното множество NSL-KDD покажува осцилирачка точност од околу 83,09%, што не е значајна точност кога станува збор за IDS во реално време.

Авторите во [159] користат хомоморфно шифрирање, како и CNN за развој на дистрибуиран IDS систем базиран на FL. Моделот е прилагоден да го анализира и блокира само DDoS сообраќајот на сателитско-копнените мрежи. Исто така, авторите во [160] развиваат FL модел за откривање DDoS напади во реално време во IoT мрежи, заснован на концептот на затворена повторлива единица (Gated Recurrent Unit). Притоа, двата модели [159,160] покажуваат прецизност при детекцијата, но се приспособени само за специфичен тип на напад и немаат присутна применливост. Пристапот DIoT [161] користи федеративно учење за да ги агрегира профилите на однесувањето на IoT мрежата. По евалуацијата во реални услови, овој пристап не пријавува лажни аларми.

Некои од пристапите користат комбинација на FL и блокчејн [162,163] со цел да се откријат аномалии во IoT мрежите. Во [164], федеративниот модел за длабоко учење создаден за zero-day ботнет напади на IoT уреди ги надминува традиционалните децентрализирани пристапи, како и локализираното длабоко учење и дистрибуираните DL методи. Во [165], воведен е нов FL модел на приватност по дизајн кој користи модел на LSTM за справување со откривање аномалии во паметни згради. Резултатите покажуваат двојно побрза конвергенција за време на процесот на обука во споредба со централизиран LSTM.

Хуонг и сор. [166] предлагаат откривање на аномалии во системи за индустриски ИОТ. Тие користат архитектура составена од VAE-енкодер, единица LSTM и VAE-декодер. Главното експериментирање се врши со користење на податоци од Фабрика за гасоводи, стекнати од SCADA системи. Резултатите покажуваат дека пристапот добива 97,9% F1-резултат. Ал-Мари и сор. [167] прават спој од предностите на FL и мимичкото учење (mimic learning) за да создадат дистрибуирани IDS што постигнуваат 98,11% точност на откривање обучени врз податочното множество NSL-KDD. Мимичкото учење е техника

која користи два модели и две податочни множества за обука кои споделуваат информации по процесот на обука. Цетин и сор. [168] предлагаат безжичен IDS кој користи stacked автоенкодери и FedAvg за детекција на аномалии обучен врз податочното множество AWID [169].

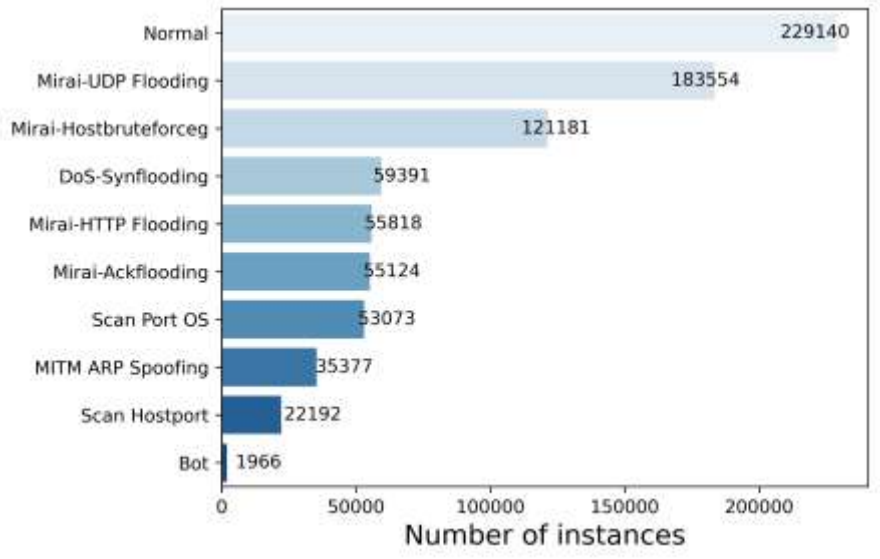
Авторите во [170] предлагаат пристап кој кој користи FL за откривање на напади во сензорски мрежи на возила (VSN). Пристапот овозможува споделување на компјутерски ресурси и заеднички услуги меѓу уредите. За да се подобри ефективноста на откривањето напади, предложениот пристап комбинира RF со множество од GRU. Ду и сор. [171] во своето истражување го нагласуваат присуството на информации специфични за уредот генерирани од сензори во IoT уредите на возилата, како што се GPS, камерите и радарот. Споделувањето на овие информации е неопходно за кооперативно возење, но истото може да претставува безбедносен ризик. За да се одговори на овој предизвик, авторите предложија искористување на FL за подобрување на безбедноста и перформансите на системот. Нивниот пристап вклучува интегрирање на масивни IoT мрежи кои опфаќаат различни уреди, со цел да се постигне подобрена безбедност и перформанси во системот.

Од оваа релевантна литература, може да се забележи дека дел од истражувањата во областа на FL првенствено се фокусираат на ограничен број напади, како што е DDoS, што значително ја ограничува нивната применливост на сценарија од реалниот свет. Ова е многу штетно за самиот процес на класификација, бидејќи IDS бараат разновидни и ажурирани податоци со цел да се добие висока точност и робусност. Исто така, голем број од истражувањата не се фокусираат на делот со зачувување на приватноста на податоците на корисниците, туку повеќе се насочени кон подобрување на точноста на самиот IDS систем.

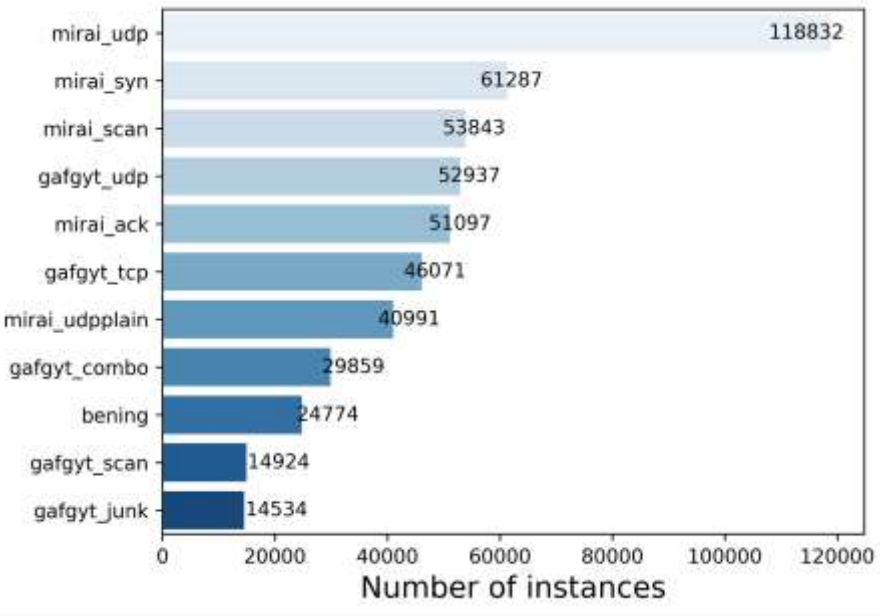
Глава 6 Развој на пристап за детектирање на аномалии во мрежен сообраќај во мрежи од Интернет на нешта

6.1 Избор на податочни множества

Врз основа на анализата на податочните множества прикажани во Глава 4 (Табела 4.1.) направен е избор на две податочни множества - IoTID20 и N-BaIoT како основа за развој на пристапот за детектирање на аномалии во IoT мрежен сообраќај. Изборот е направен врз основа на неколку карактеристики: актуелноста на самите множества, бројот на класи на напади, како и релевентноста на множествата врз развојот на овој пристап. За двете множества може да се каже дека се релевантни за проблемот кој се разгледува, имајќи во предвид дека се работи за означен IoT мрежен сообраќај. Исто така, овие податочни множества се избрани бидејќи содржат актуелни и разновидни типови на напади кои можат да се случат во ваквите мрежи, за разлика од другите множества. Битно е да се напомене дека за експериментите со ML и DL се искористени и двете податочни множества, додека пак експериментите со FL се извршени само врз податочното множество IoTID20, бидејќи по резултатите од експериментите со ML и DL констатирано е дека ова IoT множество содржи повеќе типови на различни напади за разлика од второто множество, кое содржи поспецифичен тип напади кои припаѓаат на семејството на ботнет напади. На сликите 6.1. и 6.2. е дадена дистрибуцијата по класи на напади на двете податочни множества. Првото множество содржи 10 класи на напади, додека второто содржи 11.



Слика 6.1. Дистрибуција на класи на напади во IoTID20. Податочното множество се состои од 9 типови на напади - Mirai UDP Flooding, Mirai Hostbruteforce, DoS SYN Flooding, Mirai HTTP Flooding, Mirai ACK Flooding, Scan Port OS, MITM ARP Spoofing, Scan Hostport & Bot и Normal категорија.



Слика 6.2. Дистрибуција на класи на напади во N-BaIoT. Типовите на напади се поделени во 10 категории: gafgyt_combo, gafgyt_junk, gafgyt_scan, gafgyt_tcp, gafgyt_udp, mirai_ack, mirai_raudp, mirai_rai_ како и benign категорија, која го содржи нормалниот мрежен сообраќај на уредите.

6.2 Експерименти со централизирано машинско учење

За развојот на пристапот за детектирање на аномалии во мрежен сообраќај во мрежи од IoT направени се повеќе типови на експерименти. Во овој дел се прикажани ML и DL алгоритмите со нивните параметри, како и резултатите кои се добиени од процесот на обука и евалуација на истите. Притоа, за секој експеримент прикажана е матрица на конфузија (confusion matrix), која што претставува табела која ги сумира перформансите на моделот за класификација со прикажување на резултатите на вистински позитивни (TP), вистински негативни (TN), лажно позитивни (FP) и лажно негативни (FN) предвидувања/одлуки.

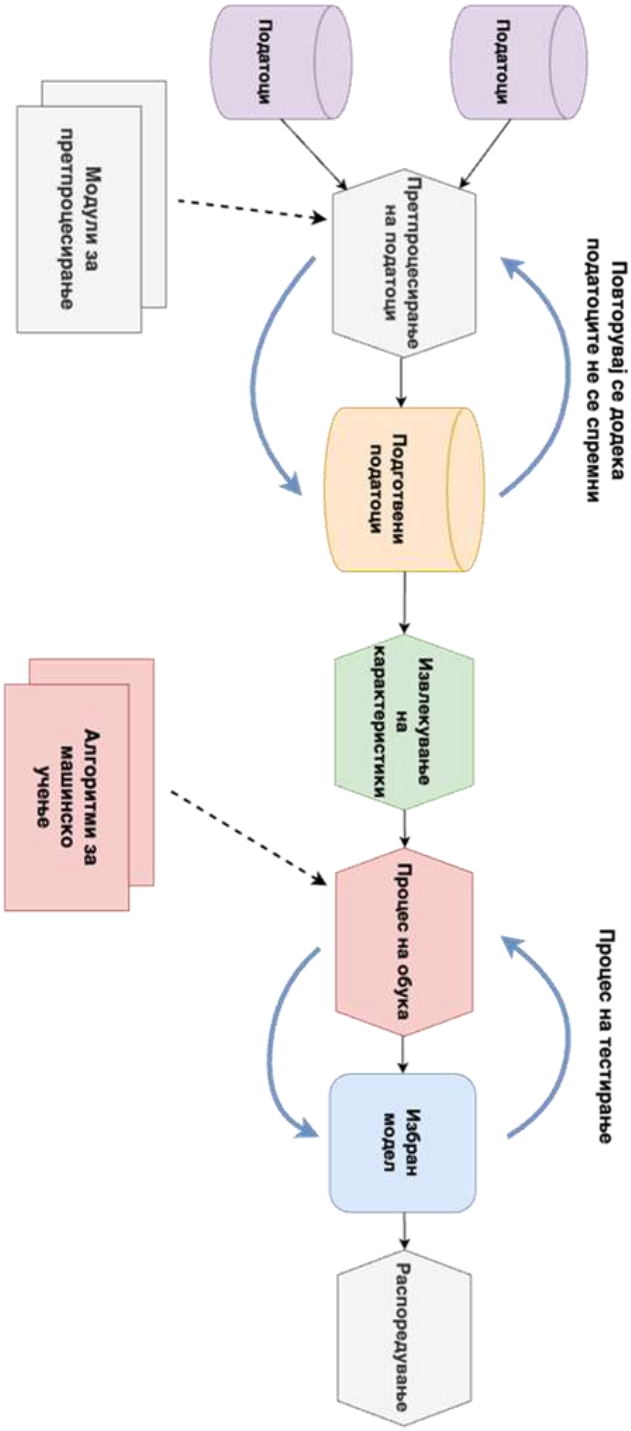
6.2.1. ML експерименти и резултати

Во експериментите со машинско учење искористени се четири познати и често користени алгоритми врз двете податочни множества и тоа: XGBoost, K-nearest neighbours, Naive Bayes и Random Forest. Пред да се започне со процесот на обука на моделите, двете податочни множества се соодветно претпроцесирани. Понатаму, множествата се поделени во однос 80/20, што значи дека 80% од податочните множества се искористени за процесот на обука на моделите, додека 20% од податочните множества се искористени за процесот на евалуација на обучените модели. Кај првото множество класата 0 го претставува нормалниот мрежен сообраќај; класите 1, 4, 5 и 6 ги претставуваат различните Мираи напади; класа 2 - DoS; класа 3 - ScanPortOS, класа 7 - Hostport Scan; класа 8 - MitM; класа 9 – Bot, додека кај второто множество класата 0 го претставува нормалниот мрежен сообраќај; класите 1, 2, 3, 4 и 5 ги претставуваат различните Mirai напади (Scan, UDP, ACK, SYN и UDP plain); класите 6, 7, 8, 9 и 10 ги претставуваат различните типови на Bashlite напади (Combo,TCP Flooding, UDP Flooding, Junk, Scan).

Во првиот дел од експериментите извршена е бинарна класификација на нападите т.е. бројот на класи е 2, со тоа што како класа 1 е означен нормалниот сообраќај, додека класа 0 ги групира сите типови на напади присутни во податочните множества. Во вториот дел од експериментите задржана е повеќекласната класификација која е присутна од самиот почеток во податочните множества (10 и 11 класи).

На Слика 6.3 прикажан е целосниот ML процес употребен во ML експериментите. Тој опфаќа неколку клучни фази, почнувајќи со претходна обработка на податоците, каде што необработените податоци се чистат, трансформираат и форматираат за да бидат соодветни за понатамошна анализа. По претходна обработка на податоците, фазата на подготовка на податоците вклучува задачи како што се поделба на податочното множество на множества за обука и тестирање, справување со вредностите што недостасуваат и решавање на нерамнотежата на класите, доколку е применливо. Следниот чекор е екстракција на карактеристики, каде што релевантните карактеристики се извлекуваат од претходно обработените податоци за да се претстават основните шаблони и карактеристики. Оваа фаза ја подобрува ефикасноста и ефективноста на ML алгоритмите. Понатаму, процесот вклучува примена на различни ML алгоритми на подготвеното податочно множество. Овие алгоритми опфаќаат низа техники. Секој алгоритам се применува на податоците, учи од податочното множество за обука и генерира предвидувања или класификации за множеството за тестирање. На крај, процесот се фокусира на идентификување на моделот со најдобри перформанси. Овој дел обично вклучува евалуација на перформансите на секој ML алгоритам врз

основа на метрика (како што е точноста). Моделот што покажува највисоки перформанси на метриката за евалуација е избран како најдобар модел за дадената задача.



Слика 6.3. Искористениот ML процес во ML експериментите. Се состои од претходна обработка на податоци, подготовка на податоци, екстракција на карактеристики, примена на различни ML алгоритми и наоѓање на најдобриот целокупен ML модел за оваа задача.

ML резултати од бинарна класификација

Анализирајќи ги резултатите добиени од бинарната класификација кај првото податочно множество (IoTID20), може со сигурност да се каже дека алгоритмот XGBoost има најголема точност (од 99%) за разлика од сите останати извршени алгоритми (види Додаток – Слика 6.4.). Истото важи и за експериментите извршени на второто податочно множество N-BaIoT (види Додаток – Слика 6.5.), каде XGBoost постигнува точност од 100%. Како втор најдобар алгоритам може да се спомне K - Nearest Neighbours со точност од околу 98%, односно 97,5% за второто множество. Исто така, може да се нагласи дека иако овој алгоритам има помала точност, неговото време на тренирање е пократко отколку она на XGBoost, што можеби е важно кога станува збор за IDS системи кои работат во реално време. Најлош алгоритам и во двата случаи е Naive Bayes. Резултатите од овие експерименти се сумаризирани во Табела 6.1.

Табела 6.1. Резултати од ML бинарни експерименти. Може да се види дека и за двете податочни множества, алгоритмот XGBoost обезбедува најдобра вкупна точност. Најниска точност може да се открие при експериментирање со Naive Bayes.

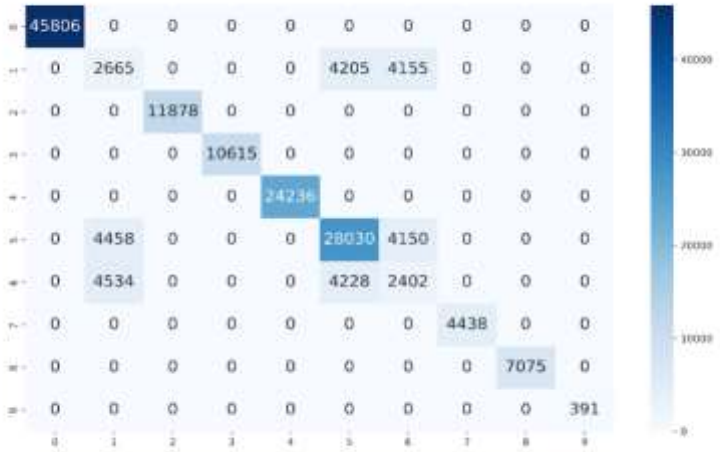
Алгоритми	Податочни множества							
	IoTID20				N-BaIoT			
	TP	TN	FP	FN	TP	TN	FP	FN
XGBoost	**117459**	**45806**	**1**	**0**	**96875**	**4955**	**0**	**0**
K- Nearest Neighbours	117401	45710	59	96	96863	4954	12	1
Naive Bayes	116651	31261	809	14545	96865	4937	10	18
Random Forest	117034	37778	426	8028	96871	4918	4	37

ML резултати од повеќекласна класификација

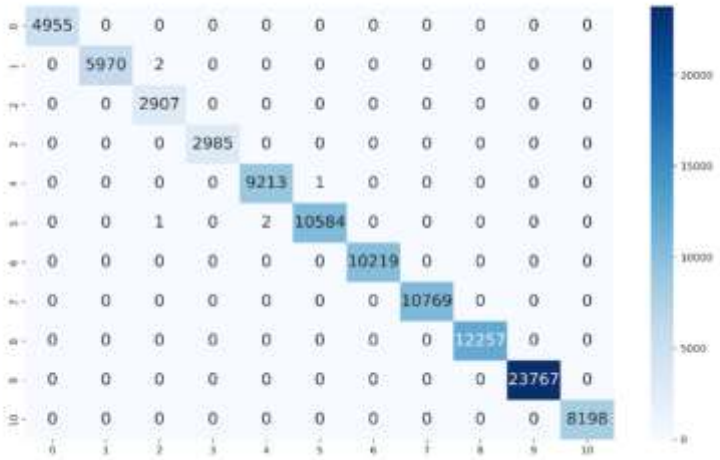
По класифицирањето на податочните множества во 10 и 11 класи соодветно може да се забележат резултатите добиени од спроведените експерименти. Од матриците за конфузија (види Додаток - Слика 6.7. и Слика 6.8.) може да се забележи дека алгоритмите ги мешаат класите 5 и 6 (што одговараат на истата фамилија на напади - Mirai). Исто така, од сликите може да се забележи дека алгоритмот Random Forest има најлоши перформанси, додека алгоритмот K- Nearest Neighbors е втор најдобар (речиси 100% точен).

Матриците на конфузија на Слика 6.6. покажуваат дека алгоритмот XGBoost има најдобри перформанси кога е обучен и тестиран на двете податочни множества. Меѓутоа, на Слика 6.6. а) може да се забележи истата конфузија помеѓу класите 5 и 6 во првото податочно множество. Ова не е случај со Слика 6.6. б), каде што XGBoost ги предвидува класите со точност од 100%.

а) XGBoost IoTID20



б) XGBoost N-BaIoT

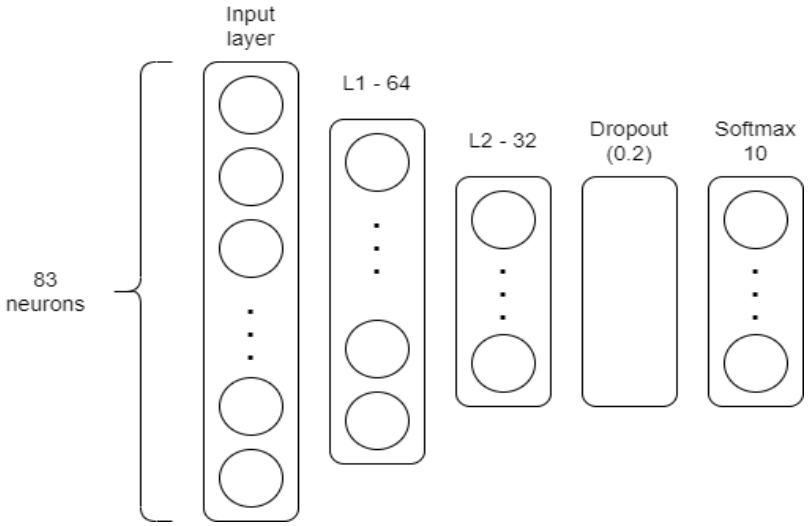


Слика 6.6. Резултати од ML повеќекласната класификација со IoTID20 и N-BaIoT - XGBoost. Двете матрици за конфузија го прикажуваат најдобриот севкупен алгоритам од експериментите кога се изведуваат на двете податочни множества.

Со внимателна анализа на податочните множества, може да се заклучи дека некои од нападите се многу поврзани според нивниот тип и својствени семејни карактеристики. На пример, постојат неколку различни Mirai напади кои покажуваат многу слично однесување на упад во мрежата. Како примарна цел на IDS во мрежата е точно и навремено откривање напади, така што може да биде многу корисно ако системот може да ги групира нападите и да ги подобри своите способности за детекција. Бидејќи групирањето ќе се врши на исто семејство на напади, системот сепак ќе може да го идентификува типот на нападот, но неговата грануларност ќе биде погруба. Затоа, кај експериментите со длабоко учење е направена повеќекласна класификација со групирање на нападите од тип Mirai, кои се присутни во двете податочни множества.

6.2.2. DL експерименти и резултати

Во експериментите со длабоко учење искористен е модел на длабока невронска мрежа со два целосно поврзани слоја со 64 и 32 неврони, соодветно (Слика 6.9.). Слоевите ја користат функцијата за активирање ReLU. Двата слоја се проследени со слој на напуштање (dropout) со стапка од 0,2. Излезниот слој е softmax слој кој се состои од 10 и 11 неврони за секое од податочните множества соодветно, кои ги претставуваат класите на напади во податочното множество. Покрај бинарна класификација каде што во крајниот слој има само еден неврон (т.е. се предвидува нормален или абнормален сообраќај), во експериментите со групирање на напади, бројот на неврони во излезниот слој е намален на 7. Во експериментите со групирање напади во првото податочно множество, типовите на напади 1, 4, 5 и 6 се групирани во класа 1 (Mirai), додека во второто податочно множество, типовите напади 1, 2, 3, 4 и 5 се групирани во класа 1 (Mirai).



Слика 6.9. Структура на DNN со 10 неврони. Искористената невронска мрежа се состои од два целосно поврзани слоја со 64 и 32 неврони. Двата слоја се проследени со слој на напуштање од 0,2. Излезниот слој е softmax слој кој се состои од 10/11 излезни неврони.

DL резултати од бинарна класификација

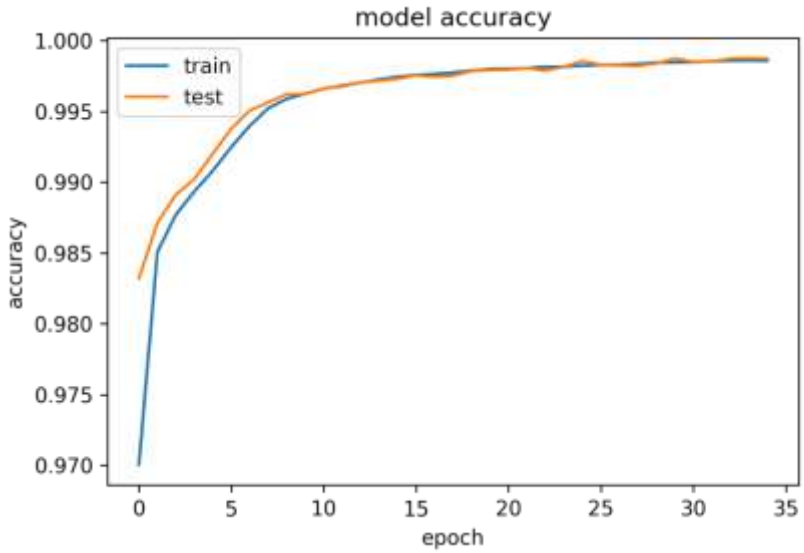
Во експериментите со DL, искористена е дополнителна метрика - точност по епоха како график за евалуација на моделите. Овој тип на график е визуелизација што го прикажува перформансот на DL моделот при повеќе епохи на обука. Во DL, епохата се однесува на целосно поминување низ целото податочно множество за обука за време на процесот на обука на моделот. Метриката за точност го претставува процентот на правилно предвидени ознаки или класификации направени од моделот на дадено податочно множество.

Анализирајќи ги резултатите од графиците за точност добиени од бинарната класификација (Слика 6.10.) со користење на DL врз двете податочни множества, може да се каже дека истата невронска мрежа дава слични резултати. Во првото податочно множество, точноста на предвидувањето е околу 99,8%, додека во втората 99,98%. Иако оваа разлика од 0,1% можеби е премногу мала за да се истакне, може да придонесе за конкретниот проблем што се разгледува. По приближно 10 рунди, може да се забележи

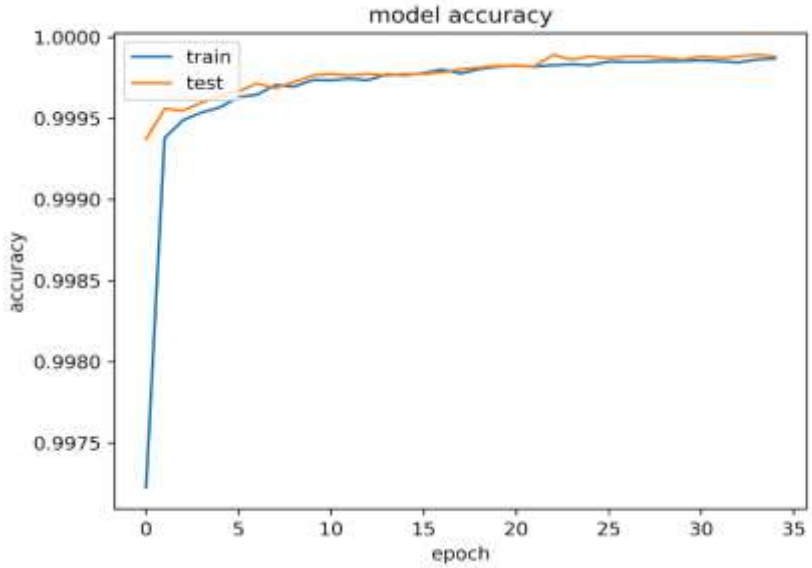
дека прецизноста на двата модели се искачува до 99%, што значи дека процесот на учење на двата модели е брз и можеби не се потребни повеќе од 35 рунди за обука.

Матриците за конфузија (види Додаток - Слика 6.11.) исто така го потврдуваат трендот прикажан на графиците за точност - истата невронска мрежа дава слични резултати за двете податочни множества (прецизност од речиси 100%).

а) График на точност - IoTID20



б) График на точност - N-BaIoT



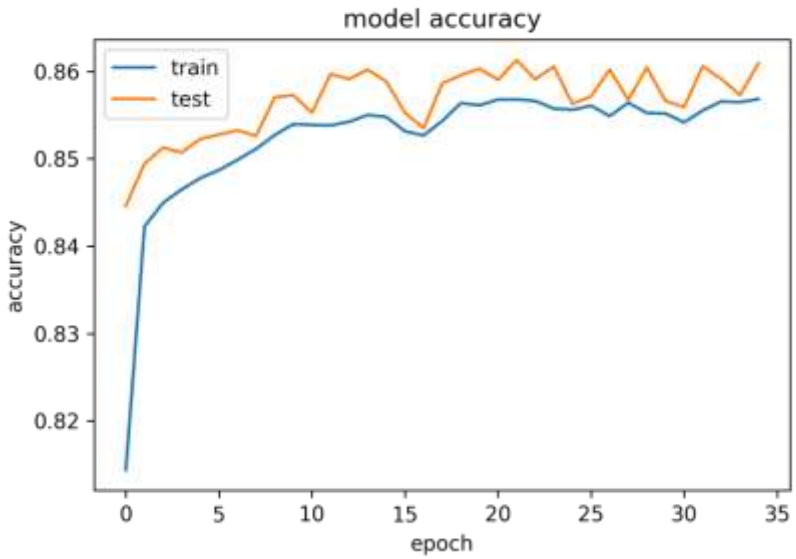
Слика 6.10. Резултати од DL бинарна класификација со IoTID20 и N-BaIoT – графици на точност. Може да се забележи дека по 10 рунди и двата модели постигнуваат прецизност од над 99%.

DL резултати од повеќекласна класификација

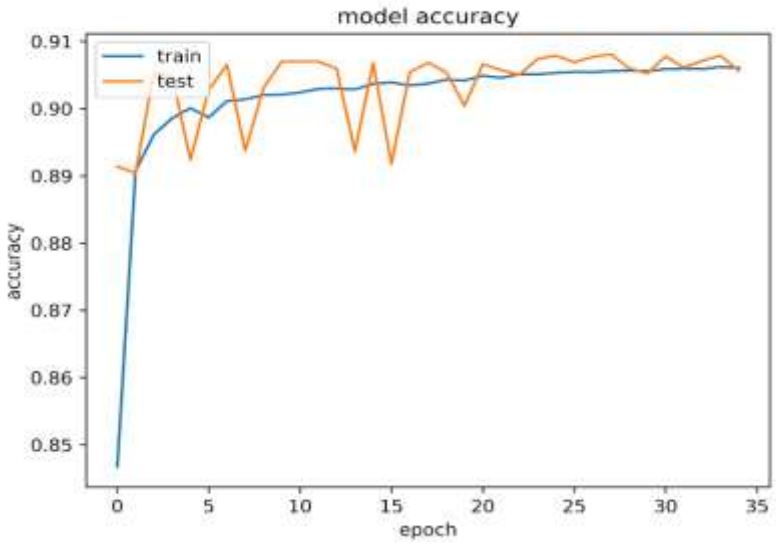
Графиците на точност прикажани на Слика 6.12. одговараат на експериментите за класификација со повеќе класи користејќи DL. Притоа, за првото податочно множество може да се забележи дека по класифицирањето на нападите во 10 различни класи, точноста на предвидувањето на првиот модел паѓа на 86%, што се очекува поради поделбата на самите податоци. Од графикот за точност може да се забележи дека кривата сè уште расте, што асоцира на можно продолжување на процесот на обука (моделот треба да се тренира уште неколку епохи). Во второто податочно множество, по класифицирањето на нападите во 11 различни класи, прецизноста на предвидувањето на вториот модел паѓа на 90,5%. Причината е иста како и во првиот модел - доаѓа од поделбата во различни класи. Конечно, може да се забележат скокови во графиците на точност за двата модели. Овие скокови се резултат на нестабилноста на моделот во справувањето со новите податоци, што се случува иако точноста за податочното множество за обука расте непречено.

Од матриците за конфузија (види Додаток - Слика 6.13.), може да се забележи дека погрешните класификации се случуваат кај двата модели: првиот модел погрешно ги класифицира класите 5 и 6, кои одговараат на истата фамилија на напади (Mirai), додека вториот модел погрешно ги класифицира класите 4 и 5, кои исто така припаѓаат на истата фамилија на напади - Mirai.

а) График на точност - IoTID20



б) График на точност - N-BaIoT

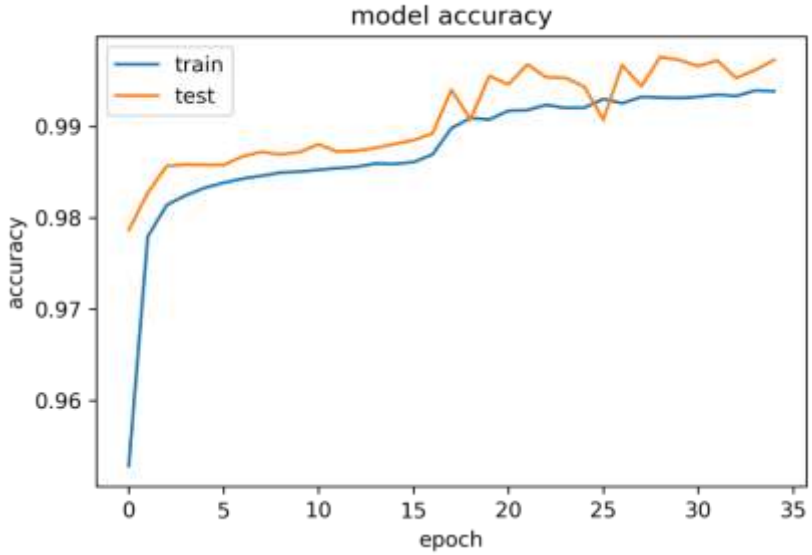


Слика 6.12. Резултати од DL со повеќекласна класификација со графици за точност на IoTID20 и N-BaIoT. По класифицирањето на нападите во 10 различни класи, прецизноста на предвидувањето на првиот модел паѓа на 86%, додека кај второто податочно множество точноста на предвидување на вториот модел паѓа на 90,5%.

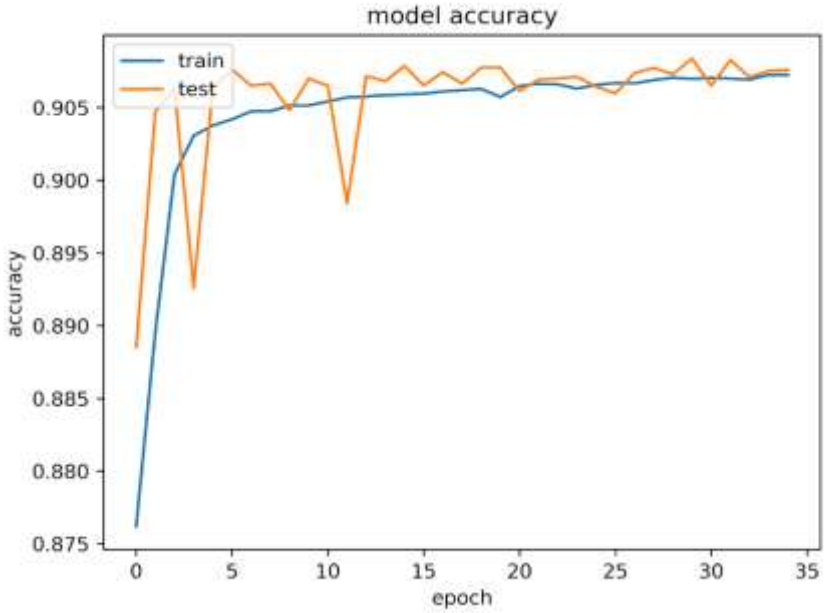
DL резултати со групирање на Mirai

Графиците на точност на Слика 6.14. ги прикажуваат резултатите по групирањето на Mirai нападите во двете податочни множества и намалувањето на класите од 10 на 7 и од 11 на 7 соодветно. Со првиот модел може да се забележи зголемување на точноста на предвидувањето - 99,5%. Исто така, може да се види дека кривата на учење сè уште расте, што укажува на можно продолжување на процесот на обука. Во случајот со вториот модел, може да се забележи дека точноста останува речиси иста (околу 91%). Исто така, во матриците за конфузија (види Додаток - Слика 6.15) може да се воочи погрешната класификација на класите 3 и 4, кои всушност се дел од истата фамилија на напади на ботнет (Bashlite). Ова води до заклучокот дека можеби овие напади можеби ќе треба дополнително да се групираат.

а) График на точност - IoTID20



б) График на точност - N-BaIoT

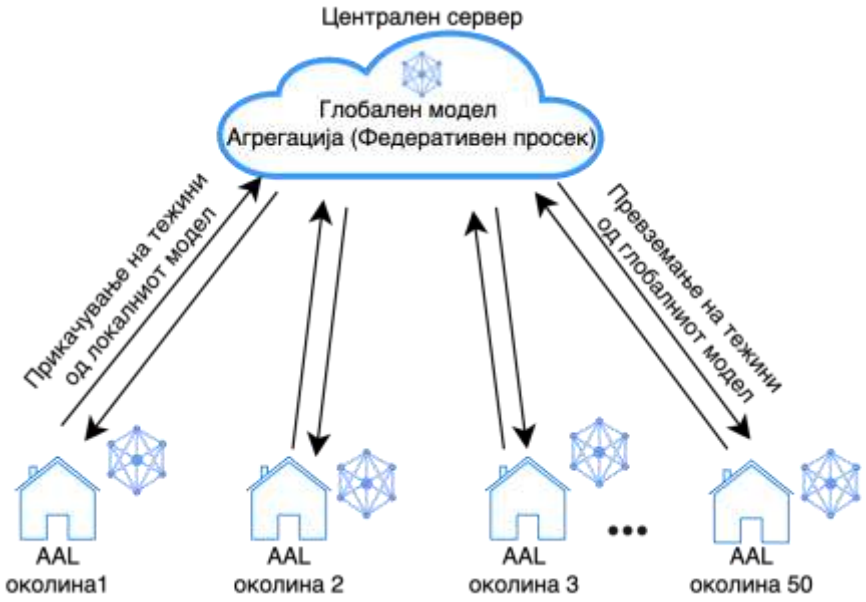


Слика 6.14. Резултати од DL класификацијата во повеќе класи со групирање на Mirai со графици на точност на IoTID20 и N-BaIoT. Може да се забележи зголемување на точноста на предвидувањето за првиот модел (99,5%).

6.3 Експерименти со федеративно (дистрибуирано) учење

Предложената архитектура на FL системот е дадена на Слика 6.16. Таа се состои од две главни компоненти: FL клиенти (околини) и централен сервер. Клиентите во овој случај претставуваат AAL (Ambient Assisted Living) околини/паметни домови (поради природата на податочното множество што се користи). Тие го обучуваат локалниот модел (користејќи ја истата длабока невронска мрежа како во експериментите со длабоко учење) со помош на нивните локални податоци. По одреден број локални епохи, FL клиентите ги испраќаат само тежините на обучените модели до централниот сервер. Централниот сервер ги агрегира моделите и го ажурира глобалниот модел со помош на FedAvg алгоритмот. Ажурираниот модел потоа се испраќа назад до клиентите, завршувајќи една FL рунда. Процесот се повторува до постигнување на потребните перформанси или конвергенција на моделот.

Од теоретска перспектива ниту еден FL модел не може да постигне поголема точност од централизираниот DL модел, кога FL ја користи истата невронска мрежа. Причината е поврзана со манипулацијата со податочното множество. Поточно, DL моделот се обучува на целото податочно множество, додека FL ги обучува локалните модели на делови од податочното множество, а потоа ги агрегира во еден глобален модел, при што губи дел од информациите поради партиционирање на множеството и барање просек (FedAvg).



Слика 6.16. Архитектура на FL системот. Клиентите во ова сценарио претставуваат AAL (Ambient Assisted Living) средини или паметни домови. Секој клиент обучува локален модел, користејќи ја истата DNN која е искористена во експериментите за длабоко учење, со нивните соодветни локални податоци. Откако ќе се завршат однапред одреден број на локални епохи, FL клиентите ги пренесуваат само тежините на моделот до централниот сервер. Централниот сервер потоа ги собира овие модели и го ажурира глобалниот модел користејќи го FedAvg алгоритмот.

За обука на DL моделот се искорстени 35 епохи, додека за FL моделите 35 рунди. Затоа, целосните податоци се поделени на 50 клиенти (во нашиот случај, секој клиент се однесува на AAL околина), каде што секој клиент има различен дел од податочното множество за тестирање и евалуација. Подмножеството за обука на секој клиент се користи за обука на локалните модели. Глобалниот модел се оценува (во секоја рунда) користејќи комбинирани тест подмножества од сите клиенти. Важно е да се напомене дека сите 50 клиенти вклучени во процесот на федеративното учење ја имаат истата невронска мрежа. Во секоја FL рунда се избира подгрупа од случајни клиенти за локална обука, контролирана од параметарот за фракции (fraction_fit). Секој од клиентите користи само 5 епохи за обука на локалните модели. Притоа, алгоритмот FedAvg се користи за агрегација на локалните модели во глобалниот FL модел. По секоја рунда, збирните тежини на глобалниот модел се дистрибуираат до клиентите и се користат како почетна точка за обука за локалниот модел во следната рунда. Евалуацијата на овие модели се фокусира на точноста на FL во зависност од бројот на FL рунди, како и параметарот за фракции (т.е. процент на FL клиенти што се користат во секоја рунда).

Важно е да се напомене дека иако и двете податочни множества беа користени за ML и DL експериментите, FL експериментите беа извршени само на податочното множество IoTID20, бидејќи по резултатите од ML и DL експериментите беше откриено дека ова податочно множество содржи повеќе видови на различни напади кога ќе се спореди со второто податочно множество, кое содржи специфични типови на напади кои припаѓаат на семејството на напади на ботнет. Истата DNN се користи во FL експериментите. Нејзината структура за откривање аномалии во првото податочно множество (каде што бројот на излезни неврони е 10, бидејќи имаме работа со 10 класи) може да се види на Слика 6.9. За второто податочно множество и при групирање на нападите, структурата на мрежата останува иста, со мала промена во последниот слој што ги разликува 11-те класи што треба да се детектираат.

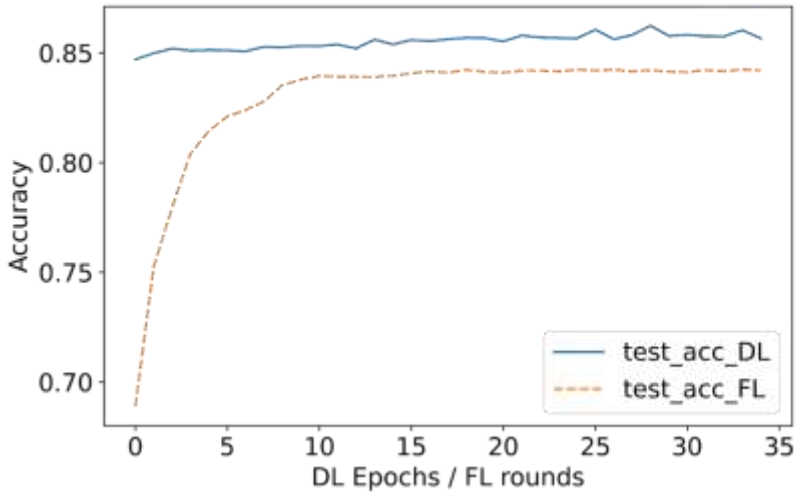
Во продолжение се дадени резултатите од експериментите извршени со FL врз податочното множество IoTID20. Во првиот дел се извршени експерименти со повеќекласно класифицирање (10 класи), слично на експериментите со DL. Вториот дел од експериментите се состои од повеќекласно класифицирање со групирање на Mirai нападите, кои ги испитуваат придобивките на групирањето со цел да се подобрат перформансите на FL. Третиот дел се состои од експерименти со карактеризација на параметрите на FL (fraction_fit), додека во четвртиот дел се анализираат перформансите на различни алгоритми на FL врз истото податочно множество.

6.3.1 Резултати од федеративно учење со повеќекласна класификација

Првиот експеримент се состои од споредба на перформансите на точноста и конвергенцијата на FL моделот со DL моделот со повеќекласна класификација. Во овој експеримент, fraction_fit параметарот е фиксиран на 1.0, што значи дека сите клиенти учествуваат во секоја рунда од процесот на обука на FL моделот. Слика 6.17. (види Додаток) ја прикажува матрицата на конфузија за FL моделот. Резултатите покажуваат дека повеќето погрешни класификации во двата модели (FL и DL) се случуваат помеѓу класите 1, 4, 5 и 6, кои одговараат на слични типови на напади, т.е. напади од семејството Mirai. Интуитивно, ова сугерира дека групирањето на Mirai нападите во една класа ќе ги подобри перформансите на класификацијата на аномалиите.

Од Слика 6.18. може да се забележи дека FL моделот постигнува малку помала точност (~84%) во споредба со DL моделот (~86%). Понатаму, по 20-тата рунда, се чини дека FL моделот ја достигнува својата конвергенција. Од друга страна, DL моделот сè уште има тенденција да ја подобри својата точност како што се зголемува бројот на епохи, но наидува на бавна конвергенција и варијабилност при повисоки перформанси (нестабилност на моделот). Овој резултат јасно ги демонстрира придобивките од користењето на FL за откривање на аномалии. По цена на мало намалување на перформансите на класификацијата, приватноста на корисникот може да се зачува во овие сценарија бидејќи моделот FL не ги споделува или изложува локалните податоци, туку само тежините на моделот. Понатаму, FL обезбедува подобра стабилност (најмногу поради тежините на NN во агрегација) и побрза конвергенција.

График на точност на DL и FL со сите 10 класи



Слика 6.18. Споредба на резултатите од DL и FL по повеќекласната класификација. Моделот FL постигнува малку помала точност (~84%) во споредба со моделот DL (~86%). Понатаму, по 20-та рунда, моделот FL се чини дека ја достигнува својата конвергенција.

6.3.2 Резултати од федеративно учење со групирање на класи

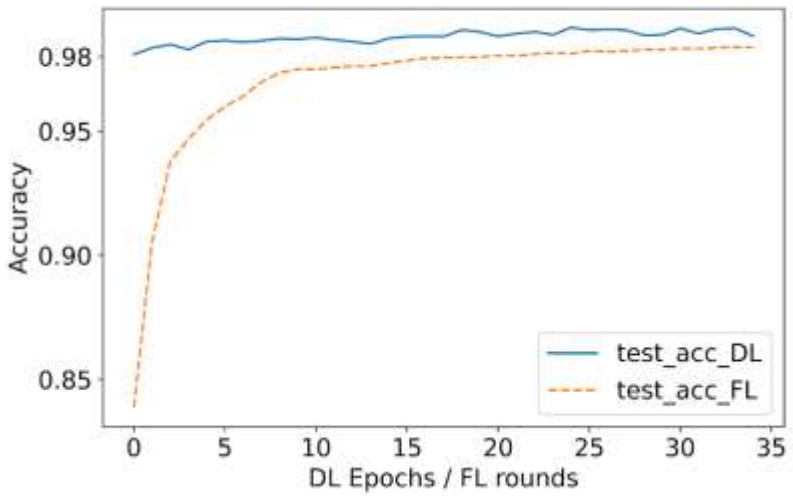
Вториот експеримент ги групира четирите Mirai класи во една, оставајќи го податочното множество со седум различни класи (како кај експериментите со ML и DL). Ова исто така го намалува бројот на јазли во завршниот слој на NN. Исто како и во првиот експеримент, fraction_fit параметарот е поставен на 1.0.

Слика 6.19. ја прикажува споредбата на моделите DL и FL во однос на точноста и конвергенција при групирање на Mirai нападите. Резултатите јасно покажуваат дека групирањето значително ја подобрува точноста на двата модели, а забележливо е и дека групирањето има повеќе придобивки за FL моделот. Поточно, разликата во перформансите помеѓу моделите е помала во споредба со случајот кога не постои

групирање на нападите. Понатаму, FL моделот има тенденција за подобрување на своите перформанси дури и по 20-тата рунда.

Причината зад ова однесување може да се најде со анализа на матриците на конфузија. Слика 6.20. (види Додаток) ја прикажува матрицата на конфузија за FL моделот по групирањето. Резултатите го потврдуваат претходното тврдење: групирањето на Mirai типовите на напади во една класа ги подобрува перформансите на класификацијата на аномалиите. Поради сличноста помеѓу нападите на Mirai и значително помалите податочни множества (1/50), локалните модели не успеваат да ги научат разликите помеѓу овие типови на напади. Затоа, групирањето на повеќе Mirai класи во една резултира со позначајна добивка во перформансите за FL моделот.

График на точност на DL и FL со групирање на напади



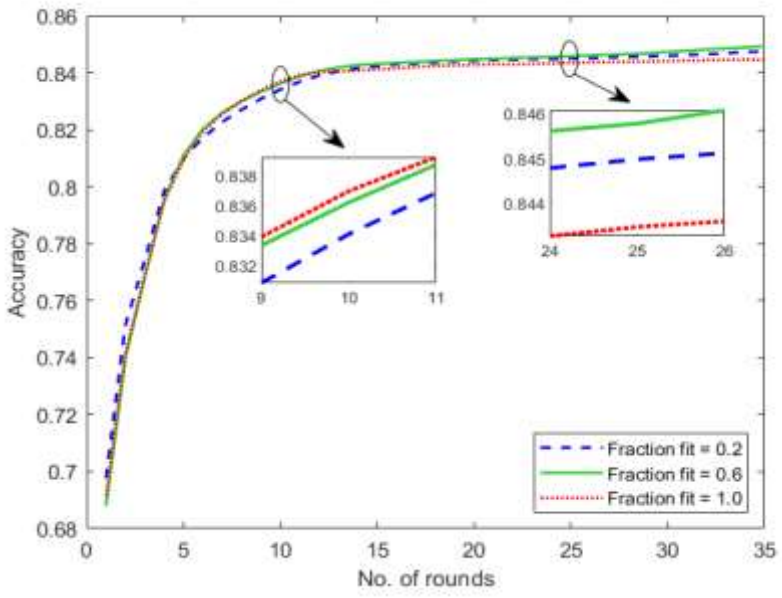
Слика 6.19. Споредба на DL и FL моделите при групирање на Mirai нападите. Резултатите јасно покажуваат дека групирањето значително ја подобрува точноста на двата модели, а забележливо е дека групирањето има повеќе придобивки за FL моделот.

6.3.3 Резултати од федеративно учење со карактеризација на параметри

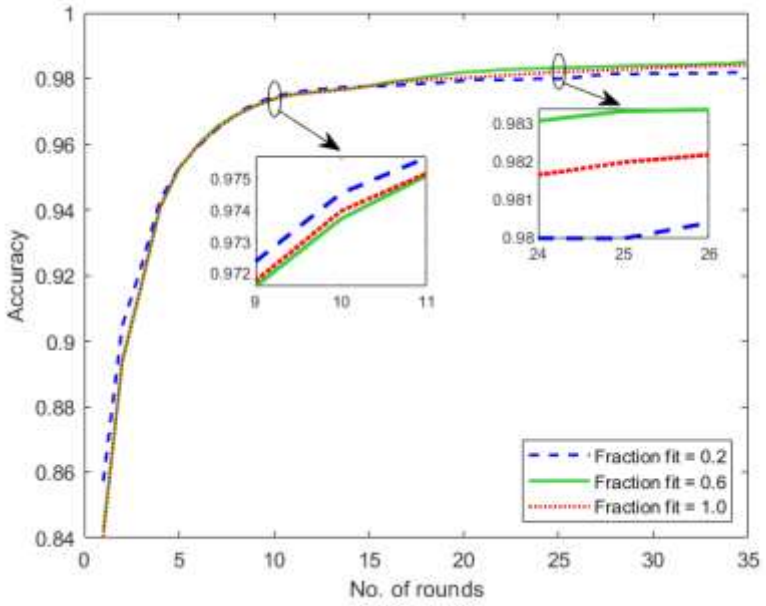
Третиот тип FL експеримент се фокусира на карактеризацијата на параметри на двата FL модели, т.е. FL модел кој ги користи сите 10 класи и FL модел кој користи групирање на Mirai нападите. Покрај бројот на рунди, овој експеримент исто така го истражува fraction_fit параметарот и неговиот придонес во точноста и конвергенцијата на моделите. Fraction fit е важен параметар во FL, бидејќи го контролира изборот на клиенти, како и изборот и стохастичноста на процесот на обука. Случајниот избор на подмножество на клиенти во секоја рунда за обука на моделот ги намалува пресметувањата ја олеснува комуникацијата, со што може да го намали преоптоварувањето во добиениот глобален FL модел. Резултатите од овој експеримент се добиени со менување на параметарот fraction_fit на три вредности: 0.2, 0.6 и 1.0. Поточно, fraction_fit еднаков на 0.2 значи дека во секоја рунда од федеративното учење, само 20% случајно избрани клиенти учествуваат во обуката (т.е. 10 по случаен избор избрани клиенти од 50 во нашиот случај).

На Слика 6.21. а) и б) е прикажана конвергенцијата и точноста за двата FL модели со менување на параметарот fraction_fit. Притоа, може да се воочи слично однесување кај двата FL модели, што значи дека параметарот не влијае значително на точноста и на конвергенцијата за овој тип проблем (класификација на аномалии). Помеѓу различниот избор на вредности за параметарот fraction_fit може да се каже дека постојат минорни разлики (<0,3%). Сепак, тука можат да се посочат неколку работи. Имено, поставувањето на помала вредност во параметарот fraction_fit обезбедува малку подобра стапка на конвергенција за помал број на FL рунди (<10). Вредноста од 0.6 обезбедува најдобра точност кога бројот на FL рунди е над 15, на пр., FL моделот со групирање на Mirai постигнува точност од 98,3% во 25-тата рунда. Исто така, може да се забележи дека поставување на параметарот на вредност 1.0 прави мал overfitting на моделот. Како заклучок, резултатите јасно покажуваат дека постои оптимален fraction_fit кој ги балансира точноста и конвергенцијата на FL моделите.

а) FL модел со сите класи



б) FL модел со групирање на напади

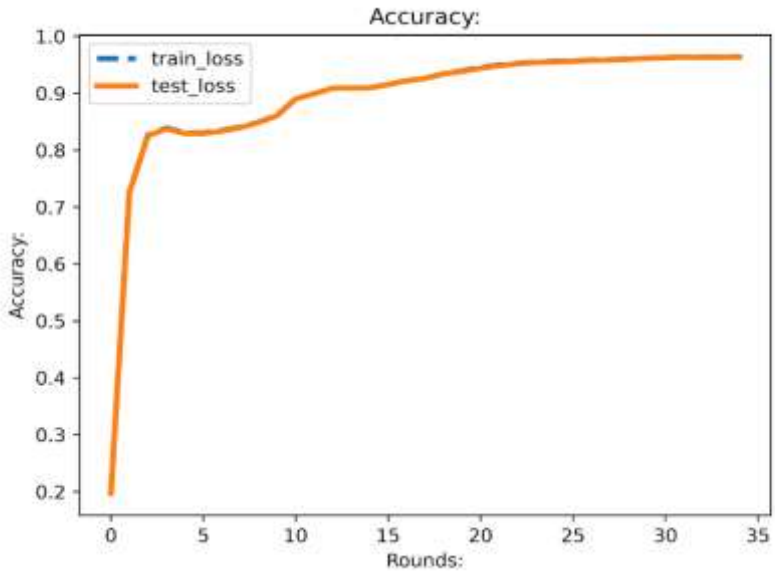


Слика 6.21. Резултати од FL со карактеризација на параметри. Слично однесување може да се забележи и кај двата FL модели, што значи дека параметарот не влијае значително на точноста и конвергенцијата за овој тип на проблем. Се покажува дека вредноста од 0,6 обезбедува најдобра точност кога бројот на FL рунди е над 15.

6.3.4 Резултати од федеративно учење со различни алгоритми

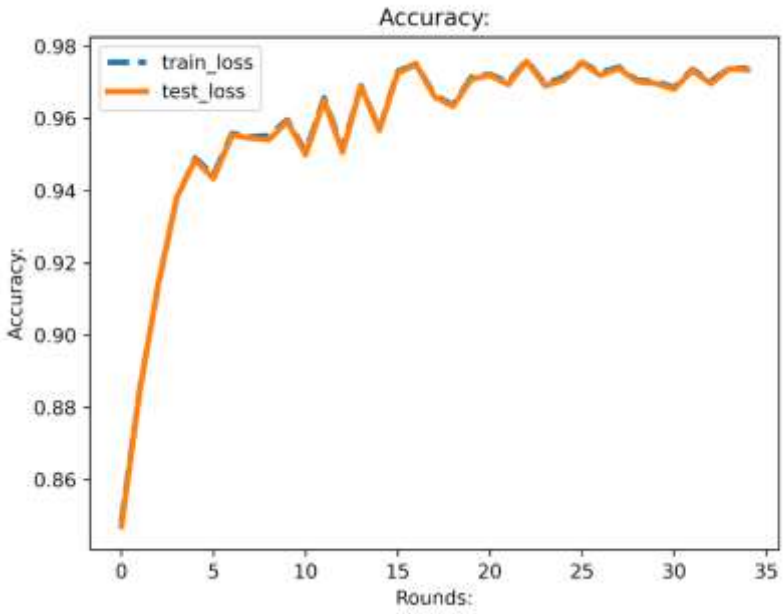
Четвртиот и последен експеримент се состои од имплементирање на различни алгоритми за FL на избраното податочно множество. Тука, експериментите се направени врз податочно множество со групирани Mirai напади, како и fraction_fit параметар поставен на 0.6. Идејата е да се искористат сознанијата од претходните експерименти со цел да се добие оптимален пристап за детекција на аномалии во IoT мрежен сообраќај. Избрани се 6 различни алгоритми: FedProx, FedMedian, FedYogi, FedAdagrad, FTFedAvg и DPFedAvg. Шестиот алгоритам е всушност имплементација на диференцијална приватност во FedAvg алгоритам.

Од анализата на алгоритмите спроведена во Глава 5, оптималниот FL алгоритам за овој тип на проблем (ако имаме избор) би бил FedYogi (Слика 6.22). Неговата точност е околу 100%, што е најдобра од сите тестирани алгоритми. Од резултатите прикажани на Слика 6.24. и Слика 6.25. (види Додаток) може да се каже дека основниот алгоритам FedAvg ги исполнува очекувањата за откривање на аномалии во IoT мрежен сообраќај. FedProx, FedMedian, како и FTFedAvg може да се каже дека постигнуваат слична точност како основниот алгоритам (околу 98%), што сепак не е занемарливо за IDS системите. За алгоритмот FedAdagrad може да се каже дека графикот на точност заситува на 97%. Исто така, на овој алгоритам во споредба со другите му требаат 10 рунди за да се стабилизира, што значи дека тука потенцијално решение би било промена на параметрите на процесот на обука на моделот.



Слика 6.22. Резултати од FL со FedYogi. Во споредба со другите FL алгоритми, може да се забележи дека користењето на FedYogi најдобро одговара за пристапот.

На Слика 6.23. е прикажан шестиот модел искористен во овој експеримент, а тоа е модел креиран со додавање на диференцијална приватност на основниот модел на FL (DPFedAvg). Од сликата може да се констатира дека моделот има добри резултати и покрај додавањето на шум. Тоа значи дека е овозможена дополнителна заштита на моделот, како и самите податоци на клиентите во системот т.е. имплементацијата на диференцијалната приватност не влијае на квалитетот на перформансите на моделот.



Слика 6.23. FedAvg резултати со диференцијална приватност. Може да се забележи дека моделот има добри резултати и покрај додавањето на бучава.

6.4 Дискусија на резултатите - резиме

Во овој дел беа направени неколку експерименти со цел да се поддржи развојот на пристап за откривање аномалии во мрежниот сообраќај на IoT мрежите.

По извршувањето на ML експериментите (бинарни и повеќекласни), се забележува дека најдобриот севкупен ML алгоритам е XGBoost проследен со KNN. Иако вториот алгоритам има помала прецизност од XGboost, тој има побрзо време на обука, што е од најголема важност за IDS системите. Со внимателна анализа на податочните множества, беше откриено дека Mirai нападите работат слично, така што во DL експериментите беше извршена дополнителна класификација со повеќе класи со групирање на Mirai нападите (присутни во двете податочни множества). Предноста на групирањето може да се види на графиците за точност и во матриците за конфузија. Тие покажуваат дека погрешните класификации се случуваат кога има повеќе класи, но кога се врши групирањето на Mirai нападите, точноста значително се зголемува. Од овие три типа на експерименти може да се заклучи дека ML и DL се одлична опција за IDS системите и треба дополнително да се истражат. Но, бидејќи овие методи се централизирани, тие имаат потешкотии да се справат со новите закани, така што дополнително беше истражен и FL како дистрибуиран пристап кон истиот проблем.

FL експериментите беа извршени само врз податочното множество IoTID20, бидејќи по резултатите од ML и DL експериментите беше откриено дека ова податочно множество содржи повеќе видови на различни напади во споредба со второто податочно множство, кое содржи специфични типови на напади кои припаѓаат на семејството на ботнет напади. При извршување на мултикласификација со FL, може да се забележи дека FL моделот постигнува малку помала точност во споредба со DL моделот. Од друга страна, DL моделот сè уште има тенденција да ја подобри својата точност како што се зголемува бројот на епохи, но наидува на бавна конвергенција и варијабилност при повисоки перформанси (нестабилност на моделот). Овој резултат јасно ги покажува придобивките

од користењето на FL за откривање аномалии. По цена на мало намалување на перформансите на класификацијата, приватноста на корисникот може да се зачува во овие сценарија бидејќи FL моделот не ги споделува или изложува локалните податоци, туку само тежините на моделот. Понатаму, FL обезбедува подобра стабилност (најмногу поради NN тежините во агрегација) и побрза конвергенција. По групирањето на Mirai нападите, резултатите јасно покажуваат дека групирањето значително ја подобрува точноста на двата модели, а забележливо е дека групирањето има повеќе придобивки за FL моделот. Резултатите од карактеризацијата на параметрите покажуваат дека постои оптимален fraction_fit (0,6) кој ја балансира точноста и конвергенцијата на FL моделите, додека fraction_fit од 1,0 прави overfitting на моделот. Што се однесува до резултатите добиени со имплементирање на различни типови на алгоритми, може да се каже дека најдобриот избор за IDS систем ќе биде FedYogi (со користење на оптималната fraction_fit од претходниот експеримент). При користење на диференцијална приватност (DPFedAvg), може да се забележи дека моделот има добри резултати и покрај додавањето на шум. Тоа значи дека DP има мало влијание кога станува збор за перформансот на моделот.

Како заклучок, оваа докторска теза ги истражи предизвиците и можностите во IDS за IoT мрежите. Преку ригорозни истражувања и експерименти, утврдено е дека идеалното решение лежи во усвојувањето на FedYogi алгоритмот, притоа користејќи fraction_fit од 0,6 проценти, како и групирање на Mirai нападите во една класа. Овој алгоритам не само што ги зема во предвид уникатните карактеристики на IoT мрежите, како што се ограничувањата на ресурсите и дистрибуираните податоци, туку исто така обезбедува робустен пристап кон IDS, во исто време зачувувајќи ја приватноста. Со искористување на FL техниките и инкорпорирање на Yogi алгоритмот, предложеното решение нуди ефикасно и ефективно средство за обезбедување на IoT мрежите од новите закани. Наодите од ова истражување го отвораат патот за подобрена безбедност во IoT околини, што на крајот придонесува за унапредување на побезбеден и доверлив Интернет на нештата.

Глава 7 Заклучок

IoT мрежите со себе носат уникатни безбедносни предизвици поради нивната дистрибуирана природа, големиот број поврзани уреди и разновидноста на апликациите и услугите што ги поддржуваат. Дополнително, разновидноста, како и големиот број на напади кои ги таргетираат овие мрежи ги можат да ги загрозат чувствителните податоци, да доведат до финансиска загуба и/или да влијаат на јавната безбедност. Многу IoT уреди имаат лоши безбедносни механизми, како што се стандардните лозинки кои лесно се погодуваат или нешифрирани комуникациски канали, што ги прави ранливи. Друг предизвик во обезбедувањето на IoT мрежите е приватноста на податоците. IoT уредите често собираат и пренесуваат чувствителни податоци, како што се лични и финансиски информации, што ги прави посакувана цел за напаѓачите. Неовластен пристап и контрола на IoT уредите е уште една значајна закана за безбедноста на IoT мрежите. Напаѓачите можат да добијат неовластен пристап до уредите и да ги користат за да извршат DDoS напади или други напади. Исто така, IoT уредите често се распоредени на оддалечени локации и имаат ограничени компјутерски ресурси, што го отежнува инсталирањето и стартувањето на традиционалниот безбедносен софтвер. Потребно е да се спроведат безбедносни мерки како што се откривање на аномалии и упади во реално време. Обезбедувањето на IoT мрежите е сложена задача која бара холистички пристап за решавање на уникатните безбедносни предизвици на уредите и мрежите на IoT. Токму обезбедувањето на ваквите мрежи беше главна мотивација за ова истражување.

На почетокот на дисертацијата беа анализирани карактеристиките на IoT, нивната архитектура и протоколниот пласт. Потоа, се продолжи кон анализа на потребите кои овие уреди ги имаат, како и ризиците со кои се соочуваат. Анализирани беа и соодветните постоечки методологии за заштита на ваквите системи и колку истите носат безбедност во нив. Потоа се пристапи кон анализа и класификација на постоечките системи за детекција на аномалии во мрежен сообраќај во IoT, со посебен осврт на оние кои користат машинско учење како метод на детекција. Тие се класифицирани и споредени според нивните предности и недостатоци, ресурси потребни за извршување и алгоритми и методи кои ги користат. Оваа класификација беше направена со цел за избор на методологија (алгоритам) од постоечките, синтеза на променет или нов сопствен алгоритам за детекција на аномалии во мрежен сообраќај.

Исто така, анализирани беа постоечки податочни множества со цел да се избере релевантно множество кое понатаму би се искористило во самите експерименти. Врз база на добиените резултати од анализите, синтетизиран е нов пристап за IDS во мрежен сообраќај во IoT мрежи. Над податочното множество се извршени експерименти со методи на ML, DL и FL. Како заклучок, може да се истакне дека пристапот со FL ги задоволува сите четири претходно поставени хипотези - обезбедува детекција на аномалии со голема точност и приватност на податоците, како и соработка помеѓу ентитетите. По споредба со други релевантни модели, може да се каже дека овој пристап има прифатливи резултати споредено со централизираните модели на ML и DL.

7.1 Придонеси во областа на истражување

Спроведеното истражување го дава следниот придонес кон литературата во областа на IDS во IoT мрежен сообраќај:

- Направен е темелен преглед на постоечката литература од областа на IoT. Притоа, дискутирани се различните нивоа во архитектурата на IoT, типичните протоколи што се користат при секојдневното функционирање на уредите и нивниот начин на работа.

- Направен е темелен преглед на безбедносните барања и предизвици во мрежите од Интернет на нешта, со посебен осврт кон нападите на секое ниво од архитектурата.

- Направена е теоретска анализа на техниките и методите кои се користат за детекција на аномалии во мрежен сообраќај кај IoT мрежи.

- Направена е компаративна анализа на различни IDS во мрежен сообраќај кај IoT мрежи, нивни предности, недостатоци и применливост. Посебен акцент е ставен на примената на машинското, длабокото и федеративното машинско учење во ваквите системи.

- Направена е компаративна анализа на различни податочни множества кои се користат за развој на IDS во мрежен сообраќај, како и нивни предности, недостатоци и применливост.

- Предложена е архитектура на дистрибуиран систем за детекција на аномалии во IoT мрежи, кој овозможува зачувување на приватноста на ентитетите и истовремено поттикнува нивна соработка.

- Развиен е пристап за дистрибуиран систем за детекција на аномалии во IoT мрежен сообраќај преку модели за бинарна класификација, повеќекласна класификација, длабоко и федеративно машинско учење.

- По развојот на пристапот за дистрибуиран систем за детекција на аномалии, анализирани се перформансите на новосоздадениот систем и направена е споредба со веќе постоечки системи, како и анализа на применливост.

Врската меѓу горенаведените придонеси и трудовите коишто се публикувани во текот на истражувањето е следна:

Во **[Труд 2]** е даден преглед на архитектурата на IoT, како и (1) безбедносните прашања, (2) предизвиците и (3) потребите во овие мрежи. Следно, проучени се домените на IoT, како и можните типови на напади што можат да се случат во секој од нив. Прикажани се и мерки кои можат да се превземат за да се заштити самата мрежа.

Во **[Труд 8]** е предложен ML базиран пристап за истражување на разликите помеѓу нормален и сообраќај со аномалии, како и влијанието на дистрибуцијата на податоците за обука во квалитетот на перформансите на моделите на машинско учење. Резултатите

покажуваат дека XGBoost покажува најдобри перформанси пред и по подсемплирање на податоците за обука.

Во **[Труд 3]** направена е анализа на алгоритми за класификација од машинско учење за детектирање на аномалии со помош на податочното множество NSL-KDD. Добиените резултати покажуваат дека алгоритмот RF ги дава најдобрите резултати при негова имплементација врз ова податочно множество, со точност при детекција од 99%. Анализа на алгоритми за класификација од машинско учење за детектирање аномалии е направена и во **[Труд 6]**, **[Труд 9]** и **[Труд 10]**. Резултатите ја потврдуваат успешноста на алгоритмите од машинско учење врз овој тип проблеми. Овие трудови се поврзани со првата хипотеза поставена во оваа докторска дисертација.

Во [**Труд 4]** е предложена архитектурата на дистрибуираниот систем за откривање на аномалии во IoT мрежи, кој овозможува безбедно користење на мрежата, како и безбедна дистрибуција на податоците меѓу ентитетите кои учествуваат во процесот. Имплементацијата на системот, како и последователната анализа е претставена **[Труд 5]** и **[Труд 1]**. Резултатите ја потврдуваат успешноста на федеративното машинско учење во зачувување на приватноста на податоците на ентитетите вклучени во системот, како и поттикнување на соработка меѓу истите. Резултатите од овие трудови се поврзани со првата, втората, третата и четвртата хипотеза поставени во оваа докторска дисертација.

7.2 Примена на резултатите од истражувањето

Пристапот и моделите кои произлегуваат од оваа докторска дисертација можат да се применат за интелегентна и ефикасна детекција на аномалии во мрежен сообраќај во IoT мрежи. Притоа, ова решение може да се имплементира во разни здравствени установи, како и паметни домови, соби за интензивна нега итн. Системот е предвиден да открива аномалии и истовремено да гарантира безбедност и приватност на корисниците и нивните лични (сензорски) податоци. Предложеното решение на различните ентитети ќе им овозможи размена на податоци на контролиран начин, притоа креирајќи партнерство помеѓу истите. Поради тоа што се работи за флексибилен систем, по потреба на истиот може да му се дадат дополнителни информации за нови напади кои би се појавиле во иднина, со цел да се задржи квалитетот на детекцијата на аномалии во мрежен сообраќај, како и точноста и прецизноста. Извршените анализи, класификации, како и евалуација на ефикасноста на постојните IDS во мрежен сообраќај од IoT мрежи може да им послужат на други истражувачи како основа за натамошни истражувања и подобрувања.

7.3 Насоки за идни истражувања

Постојат неколку насоки за дополнителни истражувања. Меѓу нив, вреди да се истакнат следните:

- IoT мрежите обично се состојат од хетерогени уреди со различни способности и ограничувања на ресурсите. Според тоа, идните истражувања треба да бидат насочени кон развој и имплементација на алгоритми за FL кои можат да се справат со варијации во компјутерската моќ, како и пропусниот опсег на мрежата.

- Процесот на агрегација на модели во FL е критична компонента на IDS, бидејќи истиот одредува како локалните модели се комбинираат со цел да се произведе глобалниот модел. Понатамошни истражуваања во оваа насока би биле оптимизација на процесот на агрегација на моделите, како и развој на нови алгоритми за агрегација кои можат да се справат со сложените структури на податоци и ограничувањата на приватноста на IoT мрежите.

- Предложениот пристап на FL за детекција на аномалии во IoT мрежен сообраќај во оваа дисертација ја зема во предвид само диференцијалната приватност како метод за зачувување на приватноста. Во иднина, треба да се направи проширување на предложениот пристап, така што ќе бидат испробани уште неколку различни методи за зачувување на приватноста и ќе биде направена соодветна анализа, како и споредба помеѓу истите.

- Анализите може да се збогатат со нови FL оптимизатори, кои ќе се тестираат на нови или веќе постоечки податочни множества. Исто така, симулациите во дисертацијата се направени користејќи податочни множества од јавно достапни бази. Во идните истражувања овој пристап дополнително може да се евалуира на ново реално податочно множество.

- Во IoT мрежите, податоците може да се менуваат со текот на времето поради промени во околината или конфигурацијата на системот. Ова може да поттикне феномен познат како „концептуален дрифт“, каде што перформансите на IDS се деградираат со текот на времето. Како идни истражување во оваа насока, треба да се размислува кон развивање техники кои можат да се прилагодат на промените во дистрибуцијата на податоците со текот на времето.

- Квалитетот на податоците може да се разликува помеѓу уредите во IoT мрежата, што може значително да влијае на точноста на предвидување и детектирање на IDS. Идните истражувања во оваа насока ќе вклучуваат како да се земе во предвид квалитетот на податоците, како и да се продолжи кон развој на методи за подобрување или оценување на нивниот квалитет при локални итерации на моделот.

- На крај, предложениот пристап на FL за детекција на аномалии во IoT мрежен сообраќај во иднина може да се дополни и примени во други научни истражувања кои се занимаваат со детекција на аномалии во мрежен сообраќај во мрежи од следна генерација.

Користена литература

[1] C. V. Nisha Angeline, S. Muthuramlingam, E. Rahul Ganesh, S. Siva Pratheep, V. Nishanthan, “Medical iot - automatic medical dispensing machine”, Machine Learning, Deep Learning and Computational Intelligence for Wireless Communication, *Springer Singapore*, pp. 323–333, 2021.

[2] O. Ribeiro, L. Gomes, Z. Vale, “IoT-Based Human Fall Detection System,” *Electronics*, vol. 11, no. 4, pp. 592, 2022, doi: 10.3390/electronics11040592.

[3] V. C. Jadala, S. K. Pasupuletti, S. H. Raju, S. Kavitha, C. M. H. Sai Bhaba, B. Sreedhar, "Need of Internet of Things, Industrial IoT, Industry 4.0 and Integration of Cloud for Industrial Revolution," 2021 Innovations in Power and Advanced Computing Technologies (i-PACT), Kuala Lumpur, Malaysia, pp. 1-5, 2021, doi: 10.1109/i-PACT52855.2021.9696696.

[4] MT. Zhou, TF. Ren, ZM. Dai, XY. Feng, “Real-Time Task Scheduling in Smart Factories Employing Fog Computing” Industrial IoT Technologies and Applications, Industrial IoT 2020, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 365. Springer, Cham., 2021, doi: 10.1007/978-3-030-71061-3_2.

[5] M. Dhanaraju, P. Chenniappan, K. Ramalingam, S. Pazhanivelan, R. Kaliaperumal, “Smart Farming: Internet of Things (IoT)-Based Sustainable Agriculture”, *Agriculture*, vol. 12, no. 10, pp. 1745, 2022, doi: 10.3390/agriculture12101745.

[6] S. Mishra, S.K. Sharma, “Advanced contribution of IoT in agricultural production for the development of smart livestock environments”, *Internet of Things*, vol. 22, 2023, doi: 10.1016/j.iot.2023.100724.

[7] R. V. Ravi, S. B. Goyal, B. C. Neagu, M. S. Raboaca, C. Verma, “A Low-Cost Industrial Automation System Using IoT and Cloud Computing”, 2022 International Conference and Exposition on Electrical And Power Engineering (EPE), Romania, 2022, pp. 649-653, doi: 10.1109/EPE56121.2022.9959772.

[8] F. Adelantado, M. Ammouriova, E. Herrera, A. A. Juan, S. S. Shinde, D. Tarchi, “Internet of Vehicles and Real-Time Optimization Algorithms: Concepts for Vehicle Networking in Smart Cities,” Vehicles, vol. 4, no. 4, pp. 1223–1245, Nov. 2022, doi: 10.3390/vehicles4040065.

[9] I. N. Stăncel, C. M. Dumitrescu, "IoT Technology and Supply Chain Management," 13th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Pitesti, Romania, 2021, pp. 1-6, doi: 10.1109/ECAI52376.2021.9515128.

[10] H. Marcos, R. Gernowo, I. Rosyida, O. D. Nurhayati, "Intelligent Traffic Management System using Internet of Things: A Systematic Literature Review," 2022 IEEE 8th International Conference on Computing, Engineering and Design

(ICCED), Sukabumi, Indonesia, 2022, pp. 1-6, doi: 10.1109/ICCED56140.2022.10010602.

[11] S. C. Rai, S. P. Nayak, B. Acharya, V. C. Gerogiannis, A. Kanavos, T. Panagiotakopoulos, “ITSS: An Intelligent Traffic Signalling System Based on an IoT Infrastructure,” Electronics, vol. 12, no. 5, p. 1177, Feb. 2023, doi: 10.3390/electronics12051177.

[12] Q. A. Tran, Q. H. Dang, T. Le, H. T. Nguyen, T. D. Le, "Air Quality Monitoring and Forecasting System using IoT and Machine Learning Techniques," 2022 6th International Conference on Green Technology and Sustainable Development (GTSD), Nha Trang City, Vietnam, 2022, pp. 786-792, doi: 10.1109/GTSD54989.2022.9988756.

[13] V. A. Orfanos, S. D. Kaminaris, P. Papageorgas, D. Piromalis, D. Kandris, “A Comprehensive Review of IoT Networking Technologies for Smart Home Automation Applications,” Journal of Sensor and Actuator Networks, vol. 12, no. 2, p. 30, Apr. 2023, doi: 10.3390/jsan12020030.

[14] C. Stolojescu-Crisan, C. Crisan, B.-P. Butunoi, “An IoT-Based Smart Home Automation System,” Sensors, vol. 21, no. 11, p. 3784, May 2021, doi: 10.3390/s21113784.

[15] A. P. Nirmala, V. Asha, P. Chandra, H. Priya, S. Raj, "IoT based Secure Smart Home Automation System," 2022 IEEE Delhi Section Conference (DELCON), New Delhi, India, 2022, pp. 1-7, doi: 10.1109/DELCON54057.2022.9753086.

[16] B. Mustafa, M. W. Iqbal, M. Saeed, A. R. Shafqat, H. Sajjad, M. R. Naqvi, "IOT Based Low-Cost Smart Home Automation System," 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), Ankara, Turkey, 2021, pp. 1-6, doi: 10.1109/HORA52670.2021.9461276.

[17] S. Shekhawat, "Use of AI and IoT to make Retail Smarter," 2022 3rd International Informatics and Software Engineering Conference (IISEC), Ankara, Turkey, 2022, pp. 1-5, doi: 10.1109/IISEC56263.2022.9998202.

[18] A. Khanna, R. Tomar, "IoT based interactive shopping ecosystem," 2016 2nd International Conference on Next Generation Computing Technologies (NGCT), Dehradun, India, 2016, pp. 40-45, doi: 10.1109/NGCT.2016.7877387.

[19] V. Perera, L. Ekanayake, A. Bandara, D. Shakya, U. S. Oruthota, "IOT Based Smart Shopping System," 2021 10th International Conference on Information and Automation for Sustainability (ICIAfS), Negambo, Sri Lanka, 2021, pp. 225-229, doi: 10.1109/ICIAfS52090.2021.9606124.

[20] J. Xu et al., "Design of Smart Unstaffed Retail Shop Based on IoT and Artificial Intelligence," in IEEE Access, vol. 8, pp. 147728-147737, 2020, doi: 10.1109/ACCESS.2020.3014047.

[21] A. Zahra, M. Asif, A. A. Nagra, M. Azeem, S. A. Gilani, "Vulnerabilities and

Security Threats for IoT in Transportation and Fleet Management," 2021 4th International Conference on Computing & Information Sciences (ICCIS), Karachi, Pakistan, 2021, pp. 1-5, doi: 10.1109/ICCIS54243.2021.9676388.

[22] R. Sivapriyan, S. V. Sushmitha, K. Pooja, N. Sakshi, "Analysis of Security Challenges and Issues in IoT Enabled Smart Homes," 2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), Bangalore, India, 2021, pp. 1-6, doi: 10.1109/CSITSS54238.2021.9683324.

[23] Z. Davoody-Beni, N. Sheini-Shahvand, H. Shahinzadeh, M. Moazzami, M. Shaneh, G. B. Gharehpetian, "Application of IoT in Smart Grid: Challenges and Solutions," 2019 5th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), Shahrood, Iran, 2019, pp. 1-8, doi: 10.1109/ICSPIS48872.2019.9066025.

[24] F. Kamalov, B. Pourghebleh, M. Gheisari, Y. Liu, S. Moussa, “Internet of Medical Things Privacy and Security: Challenges, Solutions, and Future Trends from a New Perspective,” Sustainability, vol. 15, no. 4, p. 3317, Feb. 2023, doi: 10.3390/su15043317.

[25] Z. Wang, X. Yu, P. Xue, Y. Qu, L. Ju, “Research on Medical Security System Based on Zero Trust,” Sensors, vol. 23, no. 7, p. 3774, Apr. 2023, doi: 10.3390/s23073774.

[26] K. Gopalakrishnan, A. Balakrishnan, K. Govardhanan, S. Selvarasu, “Propositional Inference for IoT Based Dosage Calibration System Using Private Patient-Specific Prescription against Fatal Dosages,” Sensors, vol. 23, no. 1, p. 336, Dec. 2022, doi: 10.3390/s23010336.

[27] W. Iqbal, H. Abbas, M. Daneshmand, B. Rauf , Y. A. Bangash, "An In-Depth Analysis of IoT Security Requirements, Challenges, and Their Countermeasures via Software-Defined Security," in IEEE Internet of Things Journal, vol. 7, no. 10, pp. 10250-10276, Oct. 2020, doi: 10.1109/JIOT.2020.2997651.

[28] A. A. Cook, G. Mısırlı, Z. Fan, "Anomaly Detection for IoT Time-Series Data: A Survey," in IEEE Internet of Things Journal, vol. 7, no. 7, pp. 6481-6494, July 2020, doi: 10.1109/JIOT.2019.2958185.

[29] A. Remesh, D. Muralidharan, N. Raj, J. Gopika, P. K. Binu, "Intrusion Detection System for IoT Devices," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 826-830, doi: 10.1109/ICESC48915.2020.9155999.

[30] R. Hattarki, S. Houji, M. Dhage, "Real Time Intrusion Detection System For IoT Networks," 2021 6th International Conference for Convergence in Technology (I2CT), Maharashtra, India, 2021, pp. 1-5, doi: 10.1109/I2CT51068.2021.9417815.

[31] ITU-T, "Overview of the Internet of Things. Y.2060", June, 2012.

[32] D. Navani, S. Jain, M. S. Nehra, "The Internet of Things (IoT): A Study of Architectural Elements," 2017 13th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), Jaipur, India, 2017, pp. 473-478, doi: 10.1109/SITIS.2017.83.

[33] M. Lombardi, F. Pascale, D. Santaniello, "Internet of Things: A General Overview between Architectures, Protocols and Applications," Information, vol. 12, no. 2, p. 87, February, 2021, doi: 10.3390/info12020087.

[34] K. Kaur, "A Survey on Internet of Things – Architecture, Applications, and Future Trends," 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), Jalandhar, India, 2018, pp. 581-583, doi: 10.1109/ICSCCC.2018.8703341.

[35] D. Aksu, M. A. Aydin, "A Survey of IoT Architectural Reference Models," 2019 16th International Multi-Conference on Systems, Signals & Devices (SSD), Istanbul, Turkey, 2019, pp. 413-417, doi: 10.1109/SSD.2019.8893170.

[36] R. Khan, S. U. Khan, R. Zaheer, S. Khan, "Future Internet: The Internet of Things Architecture, Possible Applications and Key Challenges," 2012 10th International Conference on Frontiers of Information Technology, Islamabad, Pakistan, 2012, pp. 257-260, doi: 10.1109/FIT.2012.53.

[37] Y. Fan et al., "SNPL: One Scheme of Securing Nodes in IoT Perception Layer," Sensors, vol. 20, no. 4, p. 1090, Feb. 2020, doi: 10.3390/s20041090.

[38] A. Goulart, A. Chennamaneni, D. Torre, B. Hur, F. Y. Al-Aboosi, "On Wide-Area IoT Networks, Lightweight Security and Their Applications—A Practical Review," Electronics, vol. 11, no. 11, p. 1762, Jun. 2022, doi: 10.3390/electronics11111762.

[39] R. Duan, X. Chen, T. Xing, "A QoS Architecture for IOT," 2011 International Conference on Internet of Things and 4th International Conference on Cyber, Physical and Social Computing, Dalian, China, 2011, pp. 717-720, doi: 10.1109/iThings/CPSCom.2011.125.

[40] S. Kalyani, D. Vydeki, "Measurement and Analysis of QoS Parameters in RPL Network," 2018 Tenth International Conference on Advanced Computing (ICoAC), Chennai, India, 2018, pp. 307-312, doi: 10.1109/ICoAC44903.2018.8939052.

[41] C. Wheelus, X. Zhu, "IoT Network Security: Threats, Risks, and a Data-Driven Defense Framework," IoT, vol. 1, no. 2, pp. 259–285, Oct. 2020, doi: 10.3390/iot1020016.

[42] A. M. AlAali, A. AlAteeq, W. Elmedany, "Cybersecurity Threats and Solutions of IoT Network Layer," 2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), Sakheer, Bahrain, 2022, pp. 250-257, doi: 10.1109/3ICT56508.2022.9990734.

[43] M. Z. Alam, F. Reegu, A. A. Dar, W. A. Bhat, "Recent Privacy and Security Issues

in Internet of Things Network Layer: A Systematic Review," 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, 2022, pp. 1025-1031, doi: 10.1109/ICSCDS53736.2022.9760927.

[44] L. Zhao, B. Cheng, J. Chen, "Service-Oriented IoT Resources Access and Provisioning Framework for IoT Context-Aware Environment," 2020 IEEE World Congress on Services (SERVICES), Beijing, China, 2020, pp. 245-251, doi: 10.1109/SERVICES48979.2020.00056.

[45] M. Balamurugan, G. Prabhakar, G. Amsaveni, M. Karthikumar, J. J. Shifa, E. Sharmila, "IoT-based Intelligent Mobile Application for Shopping," 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS), Pudukkottai, India, 2022, pp. 1-4, doi: 10.1109/ICACRS55517.2022.10029137.

[46] A. D. I. A. Kadir, et al., "Mobile IoT Cloud-based Health Monitoring Dashboard Application for The Elderly," 2022 4th International Conference on Smart Sensors and Application (ICSSA), Kuala Lumpur, Malaysia, 2022, pp. 161-166, doi: 10.1109/ICSSA54161.2022.9870913.

[47] S. M. Azarudeen, B. Mohan, S. Pujita, P. Swapna, "Automatic Smart Receptacle Using IoT and Mobile Application," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2021, pp. 389-392, doi: 10.1109/ICICCS51141.2021.9432137.

[48] ZigBee Alliance. (2022, Mar.). ZigBee specification: ZigBee PRO (2022). [Online].Available:https://zigbeealliance.org/wp-content/uploads/2022/03/docs-15-11-00-Zigbee-PRO-2022.pdf

[49] IEEE Standard for Local and metropolitan area networks--Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs), IEEE Std 802.15.4-2015 (Revision of IEEE Std 802.15.4-2011). [Online].
Available: https://ieeexplore.ieee.org/document/7270287

[50] LoRa Alliance. (2020, Nov.). LoRaWAN® 1.0.4 Specification. [Online]. Available: https://lora-alliance.org/resource-hub/lorawantm-specification-v104

[51] Bluetooth SIG. (2019, Dec.). Bluetooth Core Specification Version 5.2. [Online]. Available:https://www.bluetooth.com/specifications/bluetooth-core-specification/

[52] Z-Wave Alliance. (2021, Mar.). Z-Wave Application Programmers Guide, Z-Wave Plus V2, SD3503. [Online].
Available:https://z-wavealliance.org/wp-content/uploads/2021/03/Z-Wave-Application-Programmers-Guide-SD3503.pdf

[53] 3GPP. (2016, Jun.). Technical Specification Group Radio Access Network; Study on 3GPP Support for Next Generation Narrow Band IoT (NB-IoT) (Release 13). [Online].
Available:https://www.etsi.org/deliver/etsi_ts/136100_136199/136101/13.01.00_

60/ts_136101v130100p.pdf

[54] Montenegro, G., Kushalnagar, N., Eds. (2011, Sep.). Transmission of IPv6 Packets over IEEE 802.15.4 Networks. RFC 6282. [Online]. Available: https://tools.ietf.org/html/rfc6282

[55] Winter, T., Thubert, P., Brandt, A., Hui, J., Kelsey, R., Levis, P., Pister, K., Struik, R., Vasseur, J.P. (2012, Mar.). RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks. RFC 6550. [Online]. Available: https://tools.ietf.org/html/rfc6550

[56] OASIS Standard. (2014, Mar.). MQTT Version 3.1.1. [Online]. Available: https://docs.oasis-open.org/mqtt/mqtt/v3.1.1/mqtt-v3.1.1.html

[57] Shelby, Z., Hartke, K., Bormann, C. (2014, Dec.). The Constrained Application Protocol (CoAP). RFC 7252. [Online]. Available: https://tools.ietf.org/html/rfc7252

[58] OASIS Standard. (2012, Oct.). Advanced Message Queuing Protocol (AMQP) Version 1.0. [Online].
Available: https://www.oasis-open.org/standards#amqpv1.0

[59] Object Management Group. (2021, Jan.). Data Distribution Service for Real-Time Systems Specification, v1.4. [Online]. Available: https://www.omg.org/spec/DDS/1.4/

[60] P. S. Bangare, K. P. Patil, "Security Issues and Challenges in Internet of Things (IOT) System," 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering, Greater Noida, India, 2022, pp. 91-94, doi: 10.1109/ICACITE53722.2022.9823709.

[61] S. Pal, M. Hitchens, T. Rabehaja, S. Mukhopadhyay, “Security Requirements for the Internet of Things: A Systematic Approach,” Sensors, vol. 20, no. 20, p. 5897, Oct. 2020, doi: 10.3390/s20205897.

[62] K. Gai et al., "Attacking the Edge-of-Things: A Physical Attack Perspective," in IEEE Internet of Things Journal, vol. 9, no. 7, pp. 5240-5253, 1 April, 2022, doi: 10.1109/JIOT.2021.3109917.

[63] J. Dofe, A. Nguyen, A. Nguyen, "Unified Countermeasures against Physical Attacks in Internet of Things - A survey," 2021 IEEE International Symposium on Smart Electronic Systems (iSES), Jaipur, India, 2021, pp. 194-199, doi: 10.1109/iSES52644.2021.00053.

[64] H. Mohammed, T. A. Odetola, S. R. Hasan, S. Stissi, I. Garlin, F. Awwad, "(HIADIoT): Hardware Intrinsic Attack Detection in Internet of Things; Leveraging Power Profiling," 2019 IEEE 62nd International Midwest Symposium on Circuits and Systems (MWSCAS), Dallas, TX, USA, 2019, pp. 852-855, doi: 10.1109/MWSCAS.2019.8885183.

[65] F. Khan et al., “Development of a Model for Spoofing Attacks in Internet of Things,” Mathematics, vol. 10, no. 19, p. 3686, Oct. 2022, doi:

10.3390/math10193686.

[66] B. Kepçeoğlu, A. Murzaeva, S. Demirci, "Performing energy consuming attacks on IoT devices," 2019 27th Telecommunications Forum (TELFOR), Belgrade, Serbia, 2019, pp. 1-4, doi: 10.1109/TELFOR48224.2019.8971102.

[67] C. Hang, K. K. Venkatasubramanian, "Detecting data manipulation attacks on physiological sensor measurements in wearable medical systems," EURASIP Journal on Information Security, 2018, no. 1,pp.1-21, doi:10.1186/s13635-018-0082-y

[68] A. Munshi, N. A. Alqarni, N. Abdullah Almalki, "DDOS Attack on IOT Devices," 2020 3rd International Conference on Computer Applications & Information Security (ICCAIS), Riyadh, Saudi Arabia, 2020, pp. 1-5, doi: 10.1109/ICCAIS48893.2020.9096818.

[69] A. A. Olazabal, J. Kaur, A. Yeboah-Ofori, "Deploying Man-In-the-Middle Attack on IoT Devices Connected to Long Range Wide Area Networks (LoRaWAN)," 2022 IEEE International Smart Cities Conference (ISC2), Pafos, Cyprus, 2022, pp. 1-7, doi: 10.1109/ISC255366.2022.9922377.

[70] A. Alazab, A. Khraisat, S. Singh, S. Bevinakoppa, O. A. Mahdi, “Routing Attacks Detection in 6LoWPAN-Based Internet of Things,” Electronics, vol. 12, no. 6, p. 1320, Mar. 2023, doi: 10.3390/electronics12061320.

[71] M. S. Wara, Q. Yu, "New Replay Attacks on ZigBee Devices for Internet-of-Things (IoT) Applications," 2020 IEEE International Conference on Embedded Software and Systems (ICESS), Shanghai, China, 2020, pp. 1-6, doi: 10.1109/ICESS49830.2020.9301593.

[72] A. S. Ali et al., "Performance Analysis and Evaluation of RF Jamming in IoT Networks," GLOBECOM 2022 - 2022 IEEE Global Communications Conference, Rio de Janeiro, Brazil, 2022, pp. 2745-2751, doi: 10.1109/GLOBECOM48099.2022.10001334.

[73] G. M. Menon, N. V. Nivedya, N. S. Nair, "Ensuring Security in IoT Applications by Detecting Sybil Attack," in Ubiquitous Intelligent Systems, P. Karuppusamy, I. Perikos, and F. P. García Márquez, Eds. Singapore: Springer, 2022, vol. 243, pp. 325-335. doi: 10.1007/978-981-16-3675-2_23.

[74] R. Raman, "Detection of Malware Attacks in an IoT based Networks," 2022 Sixth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Dharan, Nepal, 2022, pp. 430-433, doi: 10.1109/I-SMAC55078.2022.9987253.

[75] S. R. Zahra, M. Ahsan Chishti, "RansomWare and Internet of Things: A New Security Nightmare," 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2019, pp. 551-555, doi: 10.1109/CONFLUENCE.2019.8776926.

[76] A. Alaali, W. Elmedany, "Hardware Trojan and Countermeasures for Internet of

Things," 2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), Sakheer, Bahrain, 2022, pp. 233-238, doi: 10.1109/3ICT56508.2022.9990806.

[77] R. Kumar Yadav, K. Karamveer, "A Survey on IOT Botnets and their Detection Approaches," 2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India, 2022, pp. 1901-1906, doi: 10.1109/ICAC3N56670.2022.10074482.

[78] G. Kambourakis, C. Kolias, A. Stavrou, "The Mirai botnet and the IoT Zombie Armies," MILCOM 2017 - 2017 IEEE Military Communications Conference (MILCOM), Baltimore, MD, USA, 2017, pp. 267-272, doi: 10.1109/MILCOM.2017.8170867.

[79] B. Nazzal, A. A. Zaid, M. H. Alalfi, A. Valani, "Vulnerability Classification of Consumer-based IoT Software," 2022 IEEE/ACM 4th International Workshop on Software Engineering Research and Practices for the IoT (SERP4IoT), Pittsburgh, PA, USA, 2022, pp. 17-24, doi: 10.1145/3528227.3528566.

[80] G. Mullen, L. Meany, "Assessment of Buffer Overflow Based Attacks On an IoT Operating System," 2019 Global IoT Summit (GIoTS), Aarhus, Denmark, 2019, pp. 1-6, doi: 10.1109/GIOTS.2019.8766434.

[81] S. Altayaran, W. Elmedany, "Security threats of application programming interface (API's) in internet of things (IoT) communications," 4th Smart Cities Symposium (SCS 2021), Online Conference, Bahrain, 2021, pp. 552-557, doi: 10.1049/icp.2022.0399.

[82] M. El-Hajj, A. Fadlallah, M. Chamoun, A. Serhrouchni, "A Survey of Internet of Things (IoT) Authentication Schemes," in Sensors, vol. 19, no. 5, pp. 1141, Mar. 2019. doi: 10.3390/s19051141.

[83] S. Duddu, A. Rishita sai, C. L. S. Sowjanya, G. R. Rao, K. Siddabattula, "Secure Socket Layer Stripping Attack Using Address Resolution Protocol Spoofing," 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2020, pp. 973-978, doi: 10.1109/ICICCS48265.2020.9120993.

[84] O. P. Dwyer, A. K. Marnerides, V. Giotsas, T. Mursch, "Profiling IoT-Based Botnet Traffic Using DNS," 2019 IEEE Global Communications Conference (GLOBECOM), Waikoloa, HI, USA, 2019, pp. 1-6, doi: 10.1109/GLOBECOM38437.2019.9014300.

[85] R. Petrović, D. Simić, S. Stanković, M. Perić, "Man-In-The-Middle Attack Based on ARP Spoofing in IoT Educational Platform," 2021 15th International Conference on Advanced Technologies, Systems and Services in Telecommunications (TELSIKS), Nis, Serbia, 2021, pp. 307-310, doi: 10.1109/TELSIKS52058.2021.9606392.

[86] W. Ahmad, A. Rasool, A. R. Javed, T. Baker, Z. Jalil, “Cyber Security in IoT-

Based Cloud Computing: A Comprehensive Survey," Electronics, vol. 11, no. 1, p. 16, Dec. 2021, doi: 10.3390/electronics11010016.

[87] H. Lee, S. Kwon, J.-H. Lee, "Experimental Analysis of Security Attacks for Docker Container Communications," Electronics, vol. 12, no. 4, p. 940, Feb. 2023, doi: 10.3390/electronics12040940.

[88] M. Lalit, S. K. Chawla, A. K. Rana, K. Nisar, T. R. Soomro, M. A. Khan, "IoT Networks: Security Vulnerabilities of Application Layer Protocols," 2022 14th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS), Karachi, Pakistan, 2022, pp. 1-5, doi: 10.1109/MACS56771.2022.10022971.

[89] R. Tahri, Y. Balouki, A. Jarrar, et al., "Intrusion Detection System Using machine learning Algorithms," ITM Web Conference, 2022, doi: 10.1051/itmconf/20224602003.

[90] M. Barhoush, A. Jaradat, R. Easa, "Network Intrusion Detection System: Machine Learning Approach," Indonesian Journal of Electrical Engineering and Computer Science, vol. 25, no. 2, pp. 1151-1158, 2022, doi: 10.11591/ijeecs.v25.i2.pp1151-1158

[91] H. Jmila, M. Kheder, "Adversarial machine learning for network intrusion detection: A comparative study," Computer Networks, vol. 214, p. 109073, 2022, doi: 10.1016/j.comnet.2022.109073.

[92] R. Aditya, H. H. Nuha, S. Prabowo, "Intrusion Detection using Support Vector Machine on Internet of Things Dataset," 2022 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT), Solo, Indonesia, 2022, pp. 62-66, doi: 10.1109/COMNETSAT56033.2022.9994392.

[93] T. Le, E. O. Yustus, K. Howon, "XGBoost for Imbalanced Multiclass Classification-Based Industrial Internet of Things Intrusion Detection Systems," Sustainability, vol. 14, no. 14, p. 8707, 2022, doi: 10.3390/su14148707.

[94] Y. K. Saheed, A. I. Abiodun, S. Misra, M. K. Holone, R. Colomo-Palacios, "A machine learning-based intrusion detection for detecting internet of things network attacks," Alexandria Engineering Journal, vol. 61, no. 12, pp. 9395-9409, 2022. doi: 10.1016/j.aej.2022.07.039.

[95] V. Golmah, "An Efficient Hybrid Intrusion Detection System based on C5.0 and SVM," International Journal of Database Theory and Application, vol. 7, no. 2, pp. 59-70, 2014.

[96] S.S. Tanpure, G.D. Patel, Z. Raja, J. Jagtap, A. Pathan, "Intrusion Detection System in Data Mining using Hybrid Approach," International Journal of Computer Applications, 2016.

[97] P. Bhatt, A. Morais, "HADS: Hybrid anomaly detection system for IoT environments," in 2018 International Conference on Internet of Things, Embedded Systems and Communications (IINTEC), 2018, pp. 191-196. DOI:

10.1109/IINTEC.2018.8639001.

[98] K. S. Kiran, R. K. Devisetty, N. P. Kalyan, K. Mukundini, R. Karthi, "Building a intrusion detection system for IoT environment using machine learning techniques," Procedia Computer Science, vol. 171, pp. 2372-2379, 2020.

[99] T. Gaber, A. El-Ghamry, A. E. Hassanien, "Injection attack detection using machine learning for smart IoT applications," Physical Communication, vol. 52, p. 101685, 2022.

[100] M. Nobakht, "IoT-NetSec: Policy-Based IoT Network Security Using OpenFlow," in IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), 2019.

[101] H. R. Zeidanloo, F. Hosseinpour, P. Najafi, "Botnet detection based on common network behaviors by utilizing Artificial Immune System (AIS)," 2010 Second International Conference on Software Technology and Engineering, Puerto Rico, 2010, pp. 249-253, doi: 10.1109/ICSTE.2010.5607683.

[102] S. Gao, G. Thamilarasu, "Machine-Learning Classifiers for Security in Connected Medical Devices," 2017 26th International Conference on Computer Communication and Networks (ICCCN), Vancouver, BC, 2017, pp. 1-6, doi: 10.1109/ICCCN.2017.8038522.

[103] L. Fernandez Maimo, M. Huerta-Ruiz, J. J. Martinez-Diaz, P. H. Melin, O. Castillo, "Intelligent and Dynamic Ransomware Spread Detection and Mitigation in Integrated Clinical Environments," Sensors, vol. 19, no. 5, pp. 1121, Mar. 2019. DOI: 10.3390/s19051121.

[104] A. Verner, D. Butvinik, "A Machine Learning Approach to Detecting Sensor Data Modification Intrusions in WBANs," 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), 2017, pp. 161-169, doi: 10.1109/ICMLA.2017.0-163.

[105] M. Shafiq, Z. Tian, A. K. Bashir, X. Du, M. Guizani, "CorrAUC: A Malicious Bot-IoT Traffic Detection Method in IoT Network Using Machine-Learning Techniques," in IEEE Internet of Things Journal, vol. 8, no. 5, pp. 3242-3254, May 2021, doi: 10.1109/JIOT.2020.3041024.

[106] A. Sarwar, S. Hasan, W. U. Khan, S. Ahmed, S. N. K. Marwat, "Design of an Advance Intrusion Detection System for IoT Networks," in 2022 2nd International Conference on Artificial Intelligence (ICAI), March 2022, pp. 46-51, IEEE.

[107] I. Ullah, Q. H. Mahmoud, "A Scheme for Generating a Dataset for Anomalous Activity Detection in IoT Networks," in Goutte C., Zhu X. (eds) Advances in Artificial Intelligence. Canadian AI 2020. Lecture Notes in Computer Science, vol. 12109. Springer, Cham, 2020, pp. 645-651, https://doi.org/10.1007/978-3-030-47358-7_52.

[108] M. Bagaa, T. Taleb, J. B. Bernabe, A. Skarmeta, "A machine learning security

framework for IoT systems," IEEE Access, vol. 8, pp. 114066-114077, 2020, doi: 10.1109/ACCESS.2020.3000930.

[109] S. M. Kasongo, "An advanced intrusion detection system for IIoT based on GA and tree-based algorithms," IEEE Access, vol. 9, pp. 113199-113212, 2021, doi: 10.1109/ACCESS.2021.3108406.

[110] N. Moustafa, J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," 2015 Military Communications and Information Systems Conference (MilCIS), Canberra, ACT, Australia, 2015, pp. 1-6, doi: 10.1109/MilCIS.2015.7348942.

[111] A. Abdullah Alfrhan, R. Hamad Alhusain, R. Ulah Khan, "SMOTE: Class Imbalance Problem In Intrusion Detection System," 2020 International Conference on Computing and Information Technology (ICCIT-1441), Tabuk, Saudi Arabia, 2020, pp. 1-5, doi: 10.1109/ICCIT-144147971.2020.9213728.

[112] I. Sharafaldin, A. Habibi Lashkari, A. A. Ghorbani, "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization," in Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP), Portugal, Jan. 2018, pp. 108-116. doi: 10.5220/0006636301080116.

[113] Chawla, N.V., Bowyer, K.W., Hall, L.O., et al. "SMOTE synthetic minority over-sampling technique." Journal of Artificial Intelligence Research, 2002, vol. 16, pp. 321–357. DOI: 10.1613/jair.953.

[114] S. Kumar, "Stop using SMOTE to handle all your Imbalanced Data," [Online]. Available:https://towardsdatascience.com/stop-using-smote-to-handle-all-your-imbalanced-data-34403399d3be

[115] G. Thamilarasu, S. Chawla, "Towards deep-learning-driven intrusion detection for the Internet of Things," Sensors, vol. 19, no. 9, pp. 1977, 2019. doi: 10.3390/s19091977.

[116] Y. Xiao, C. Xing, T. Zhang, et al., "An intrusion detection model based on feature reduction and convolutional neural networks," IEEE Access, vol. 7, pp. 42210-42219, 2019.

[117] A. Awajan, "A Novel Deep Learning-Based Intrusion Detection System for IoT Networks," Computers, vol. 12, no. 12, pp. 34, 2023. doi: 10.3390/computers12020034.

[118] S. Ikhwan, A. Wibowo, B. Warsito, "Intrusion Detection using Deep Neural Network Algorithm on the Internet of Things," 2022 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT), Solo, Indonesia, 2022, pp. 84-87, doi: 10.1109/COMNETSAT56033.2022.9994499.

[119] K. Wu, Z. Chen, W. Li, "A novel intrusion detection model for a massive network using convolutional neural networks," IEEE Access, vol. 6, pp. 50850-50859, 2018. DOI: 10.1109/ACCESS.2018.2875945.

[120] Y. Otoum, D. Liu, A. Nayak, "DL-IDS: a deep learning-based intrusion detection framework for securing IoT", Transactions on Emerging Telecommunications Technologies, vol. 33, no. 3, Mar. 2022.

[121] N. D. Kathamuthu, P. Murugesan, M. J. R. Peter, S. Nair, P. L. Balakrishnan, "Deep Q-Learning-Based Neural Network with Privacy Preservation Method for Secure Data Transmission in Internet of Things (IoT) Healthcare Application," Electronics, vol. 11, no. 1, Jan. 2022, doi: 10.3390/electronics11010051.

[122] P. Mohamed Shakeel, Baskar .S, V.R.Sarma Dhulipala, Sukumar Mishra, Mustafa Jaber, "Maintaining Security and Privacy in Health Care System Using Learning Based Deep-Q-Networks," Journal of Medical Systems, vol. 42, Aug. 2018, doi: 10.1007/s10916-018-1045-z.

[123] E. Tsogbaatar, M. H. Bhuyan, Y. Taenaka, D. Fall, K. Gonchigsumlaa, E. Elmroth, Y. Kadobayashi, "DeL-IoT: A deep ensemble learning approach to uncover anomalies in IoT," Internet of Things, vol. 14, p. 100391, 2021.

[124] L. Aversano, M.L. Bernardi, M. Cimitile, R. Pecori, L. Veltri, "Effective anomaly detection using deep learning in IoT systems," Wireless Communications and Mobile Computing, vol. 2021, 2021.

[125] A. Abusitta, M. Bellaiche, M. Dagenais, T. Halabi, "A deep learning approach for proactive multi-cloud cooperative intrusion detection system," Future Gener. Comput. Syst., vol. 98, pp. 308-318, 2019.

[126] A. Abusitta, O.A. Wahab, T. Halabi, "Deep learning for proactive cooperative malware detection system," in Edge Intelligence Workshop, vol. 711, 2020, p. 7.

[127] N. Moustafa, B. Turnbull, K.-K. R. Choo, "An Ensemble Intrusion Detection Technique Based on Proposed Statistical Flow Features for Protecting Network Traffic of Internet of Things," IEEE Internet Things J., vol. 6, no. 3, pp. 4815-4830, June 2018.

[128] R. Kale, Z. Lu, K. W. Fok, V. L. Thing, "A Hybrid Deep Learning Anomaly Detection Framework for Intrusion Detection," in 2022 IEEE 8th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE, 2022, pp. 137-142.

[129] S. K. Sarma, "Optimally configured deep convolutional neural network for attack detection in internet of things: impact of algorithm of the innovative gunner," Wireless Personal Communications, vol. 118, no. 1, pp. 239-260, Jan 2021.

[130] S. Gurung, M. K. Ghose, A. Subedi, "Deep learning approach on network intrusion detection system using NSL-KDD dataset," International Journal of Computer Network and Information Security, vol. 11, no. 3, pp. 8-14, 2019.

[131] M. Tavallaee, E. Bagheri, W. Lu, A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, Ottawa, ON, Canada, 2009, pp. 1-6, doi: 10.1109/CISDA.2009.5356528.

[132] K. H. Le, M. H. Nguyen, T. D. Tran, N. D. Tran, "IMIDS: An Intelligent Intrusion

Detection System Against Cyber Threats in IoT," Electronics, vol. 11, no. 4, p. 524, 2022.

[133] A. Derhab, A. Aldweesh, A. Z. Emam, F. A. Khan, "Intrusion detection system for internet of things based on temporal convolution neural network and efficient feature engineering," Wireless Communications and Mobile Computing, 2020.

[134] N. Koroniotis, N. Moustafa, E. Sitnikova, B. Turnbull, "Towards the development of realistic botnet dataset in the internet of things for network forensic analytics: Bot-iot dataset," Future Generation Computer Systems, vol. 100, pp. 779-796, Nov. 2019, doi: 10.1016/j.future.2019.06.034.

[135] I. Ullah, Q. H. Mahmoud, "Design and Development of a Deep Learning-Based Model for Anomaly Detection in IoT Networks," in IEEE Access, vol. 9, pp. 103906-103926, 2021.

[136] S. Garcia, A. Parmisano, M. J. Erquiaga, "IoT-23: A labeled dataset with malicious and benign IoT network traffic (Version 1.0.0) [Data set]," Zenodo, 2020. [Online]. Available: http://doi.org/10.5281/zenodo.4743746.

[137] Hanan Hindy, Christos Tachtatzis, Robert Atkinson, Ethan Bayne, Xavier Bellekens, June 23, 2020, "MQTT-IoT-IDS2020: MQTT Internet of Things Intrusion Detection Dataset", IEEE Dataport, doi: https://dx.doi.org/10.21227/bhxy-ep04.

[138] M. Ge, X. Fu, N. Syed, Z. Baig, G. Teo, A. Robles-Kelly, "Deep learning-based intrusion detection for IoT networks," in 2019 IEEE 24th Pacific Rim International Symposium on Dependable Computing (PRDC), Dec. 2019, pp. 256-265.

[139] M. Lopez-Martin, B. Carro, A. Sanchez-Esguevillas, J. Lloret, "Network Traffic Classifier With Convolutional and Recurrent Neural Networks for Internet of Things," in IEEE Access, vol. 5, pp. 18042-18050, 2017, doi: 10.1109/ACCESS.2017.2724638.

[140] I. Ullah, Q. H. Mahmoud, "Design and Development of RNN Anomaly Detection Model for IoT Networks," in IEEE Access, vol. 10, pp. 62722-62750, 2022.

[141] B. Kolosnjaji, A. Zarras, G. Webster, C. Eckert, "Deep learning for classification of malware system call sequences," in Proceedings of the Australasian Joint Conference on Artificial Intelligence, Springer, 2016, pp. 137-149.

[142] J. Saxe and K. Berlin, "Deep neural network based malware detection using two dimensional binary program features," 2015 10th International Conference on Malicious and Unwanted Software (MALWARE), IEEE, pp. 11-20.

[143] Zhou X., Hu Y., Wu J., Liang W., Ma J., Jin Q.,"Distribution bias aware collaborative generative adversarial network for imbalanced deep learning in industrial IoT", IEEE Trans. Ind. Inform., vol. 18, no. 5, pp. 3406-3415, 2022.

[144] C. Thomas, V. Sharma, N. Balakrishnan, "Usefulness of DARPA dataset for intrusion detection system evaluation," in Proceedings of SPIE - The International Society for Optical Engineering, March 2008, doi: 10.1117/12.777341.

[145] N. Moustafa, "A new distributed architecture for evaluating AI-based security systems at the edge: Network TON_IoT datasets," Sustainable Cities and Society, vol. 102994, 2021.

[146] Y. Meidan et al., "N-BaIoT—Network-Based Detection of IoT Botnet Attacks Using Deep Autoencoders," in IEEE Pervasive Computing, vol. 17, no. 3, pp. 12-22, Jul.-Sep. 2018, doi: 10.1109/MPRV.2018.03367731.

[147] A. Rosay, E. Cheval, F. Carlier, P. Leroux, "Network Intrusion Detection: A Comprehensive Analysis of CIC-IDS2017," in Proceedings of the IEEE International Conference on Big Data, 2019, pp. 3331-3339.

[148] J. L. Leevy, T. M. Khoshgoftaar, "A survey and analysis of intrusion detection models based on CSE-CIC-IDS2018 Big Data," Journal of Big Data, vol. 7, article no. 104, 2020.

[149] U. Banerjee, A. Vashishtha, M. Saxena, "Evaluation of the Capabilities of WireShark as a tool for Intrusion Detection," IJCA Journal, vol. 6, no. 7, Sep. 2010.

[150] H. B. McMahan et al., "Federated learning of deep networks using model averaging," CoRR, vol. abs/1602.05629, 2016.

[151] K. Wei, J. Li, M. Ding, C. Ma, H. H. Yang, F. Farokhi, S. Jin, T. Q. S. Quek, H. V. Poor, "Federated Learning with Differential Privacy: Algorithms and Performance Analysis," IEEE Transactions on Information Forensics and Security, vol. 16, pp. 2886-2901, 2021.

[152] D. Yin, Y. Chen, K. Ramchandran, P. Bartlett, "Byzantine-Robust Distributed Learning: Towards Optimal Statistical Rates," in Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 2018, pp. 6862-6866.

[153] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, V. Smith, "Federated Optimization in Heterogeneous Networks," in Proceedings of the 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 2020, pp. 3272-3279.

[154] S. Reddi, Z. Charles, M. Zaheer, Z. Garrett, K. Rush, J. Konečný, S. Kumar, H. B. McMahan, "Adaptive Federated Optimization," in Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS), Vancouver, Canada, 2020.

[155] J. Á. Morell, E. Alba, "Dynamic and adaptive fault-tolerant asynchronous federated learning using volunteer edge devices," Future Generation Computer Systems, vol. 117, pp. 140-151, 2021.

[156] H. B. McMahan, D. Ramage, K. Talwar, L. Zhang, "Learning Differentially Private Recurrent Language Models," in Proceedings of the 2017 Conference on Neural Information Processing Systems (NeurIPS), Long Beach, CA, USA, 2017, pp. 3993-4003.

[157] T. T. Huong et al., "LocKedge: Low-complexity cyberattack detection in IoT edge computing," IEEE Access, vol. 9, pp. 18694-18704, 2021.

[158] S. A. Rahman et al., "Internet of Things intrusion detection: Centralized, on-device, or federated learning?" IEEE Network, vol. 34, no. 6, pp. 36-42, Nov. 2020.

[159] K. Li et al., "Distributed network intrusion detection system in satellite-terrestrial integrated networks using federated learning," IEEE Access, vol. 8, pp. 180254-180263, 2020.

[160] V. Mothukuri et al., "Federated learning-based anomaly detection for IoT security attacks," IEEE IoT Journal, vol. 9, no. 4, pp. 3931-3942, May 2021.

[161] U. M. Aïvodji, S. Gambs, A. Martin, "IOTFLA: A Secured and Privacy-Preserving Smart Home Architecture Implementing Federated Learning," in 2019 IEEE Security and Privacy Workshops (SPW), 2019, pp. 175-180, doi: 10.1109/SPW.2019.00041.

[162] Y. Mirsky, T. Golomb, Y. Elovici, "Lightweight collaborative anomaly detection for the IoT using blockchain," J. Parallel and Distrib. Comput., vol. 145, Jun. 2020, doi: 10.1016/j.jpdc.2020.06.008.

[163] D. Unal, M. Hammoudeh, M. A. Khan, A. Abuarqoub, G. Epiphaniou, and R. Hamila, "Integration of Federated Machine Learning and Blockchain for the Provision of Secure Big Data Analytics for Internet of Things," Comput. Secur., vol. 109, C, Oct. 2021, doi: 10.1016/j.cose.2021.102393.

[164] S. I. Popoola, R. Ande, B. Adebisi, G. Gui, M. Hammoudeh, and O. Jogunola, "Federated Deep Learning for Zero-Day Botnet Attack Detection in IoT-Edge Devices," IEEE Internet of Things Journal, vol. 9, no. 5, pp. 3930-3944, 2022, doi: 10.1109/JIOT.2021.3100755.

[165] R. A. Sater, A. B. Hamza, "A Federated Learning Approach to Anomaly Detection in Smart Buildings," ACM Trans. Internet Things, vol. 2, no. 4, Article 28, Aug. 2021, 23 pages, doi: 10.1145/3467981.

[166] T. T. Huong, T. P. Bac, D. M. Long, T. D. Luong, N. M. Dan, L. A. Quang, L. T. Cong, B. D. Thang, K. P. Tran, "Detecting cyberattacks using anomaly detection in industrial control systems: A Federated Learning approach," Computers in Industry, vol. 132, p. 103509, 2021. doi:10.1016/j.compind.2021.103509.

[167] N. A. Al-Athba Al-Marri, B. S. Ciftler, and M. M. Abdallah, "Federated Mimic Learning for Privacy Preserving Intrusion Detection," in 2020 IEEE International Black Sea Conference on Communications and Networking, BlackSeaCom 2020, 2020.

[168] B. Cetin, A. Lazar, J. Kim, A. Sim, and K. Wu, "Federated Wireless Network Intrusion Detection," in Proceedings - 2019 IEEE International Conference on Big Data, Big Data 2019, pp. 6004-6006, 2019.

[169] C. Kolias, G. Kambourakis, A. Stavrou, and S. Gritzalis, "Intrusion detection in 802.11 networks: Empirical evaluation of threats and a public dataset," IEEE Communications Surveys Tutorials, vol. 18, no. 1, pp. 184-208, 2016.

[170] M. Driss, I. Almomani, Z. Huma, and J. Ahmad, "A federated learning framework for cyberattack detection in vehicular sensor networks," Complex Intelligent Systems, vol. 8, pp. 4221-4235, 2022. doi: doi.org/10.1007/s40747-022-00388-0.

[171] Z. Du, C. Wu, T. Yoshinaga, K.-L.A. Yau, Y. Ji, and J. Li, "Federated Learning for Vehicular Internet of Things: Recent Advances and Open Issues," IEEE Open Journal of the Computer Society, vol. 1, pp. 45-61, 2020. doi: doi.org/10.1109/OJCS.2020.2977543.

[172] ISO/IEC, "ISO/IEC 30141:2018 - Internet of Things (IoT) -- Reference Architecture," [Online]. Available: https://www.iso.org/standard/77572.html, 2018.

[173] ISO/IEC, "ISO/IEC 27400:2022 series - Information technology - Big data reference architecture," in Proceedings of IEEE, vol. 1, pp. 1-1, 2022. Available: https://www.iso.org/standard/78453.html.

[174] National Institute of Standards and Technology, "NIST Special Publication 800-160 Volume 2: Systems Security Engineering: Considerations for a Multidisciplinary Approach in the Engineering of Trustworthy Secure Systems," 2016. [Online]. Available: https://doi.org/10.6028/NIST.SP.800-160v2

[175] ENISA, "Baseline Security Recommendations for IoT," 2017. [Online]. Available:https://www.enisa.europa.eu/publications/baseline-security-recommendations-for-iot

[176] S. Raschka and V. Mirjalili, Python Machine Learning: Machine Learning and Deep Learning with Python, Scikit-Learn, and TensorFlow 2, 3rd Edition. Packt Publishing, 2019, ISBN: 978-1-78995-575-0.

[177] S. M. Almeghlef, A. A.-M. AL-Ghamdi, M. S. Ramzan, and M. Ragab, “Application Layer-Based Denial-of-Service Attacks Detection against IoT-CoAP,” Electronics, vol. 12, no. 12, p. 2563, Jun. 2023, doi: 10.3390/electronics12122563.

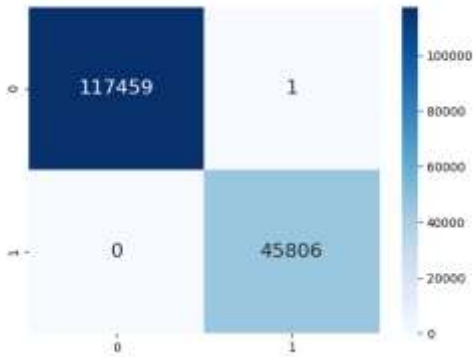
[178] J. C. Sapalo Sicato, P. K. Sharma, V. Loia, and J. H. Park, “VPNFilter Malware Analysis on Cyber Threat in Smart Home Network,” Applied Sciences, vol. 9, no. 13, p. 2763, Jul. 2019, doi: 10.3390/app9132763.

[179] E. C. Ogu, O. A. Ojesanmi, O. Awodele, and ‘Shade Kuyoro, “A Botnets Circumspection: The Current Threat Landscape, and What We Know So Far,” Information, vol. 10, no. 11, p. 337, Oct. 2019, doi: 10.3390/info10110337.

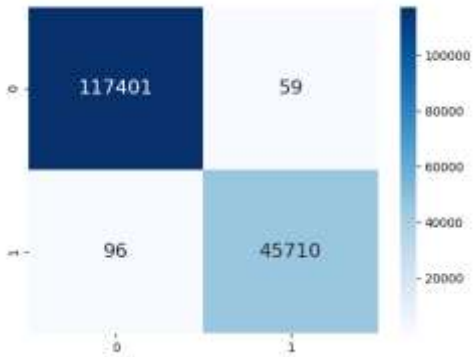
[180] A. Costin and J. Zaddach, "IoT malware: Comprehensive survey, analysis framework and case studies," in Proc. BlackHat, Aug. 2018, pp. 1–9.

Додаток

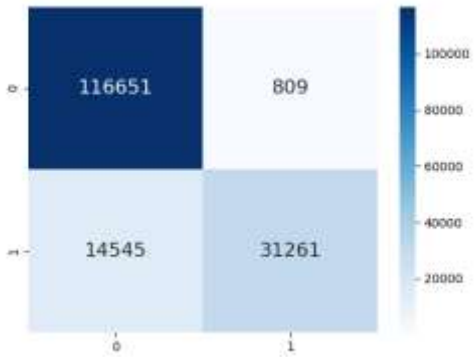
а) XGBoost



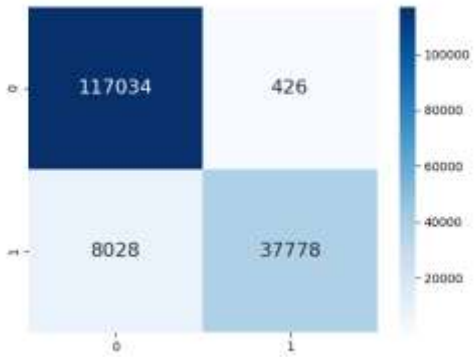
б) K - Nearest neighbours



в) Naive Bayes

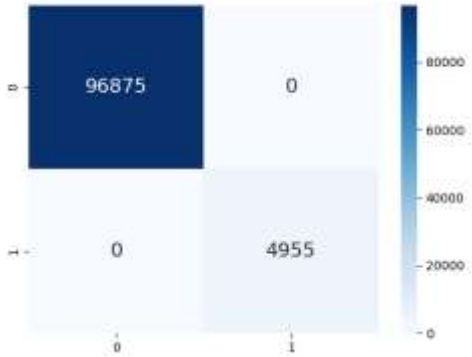


д) Random Forest

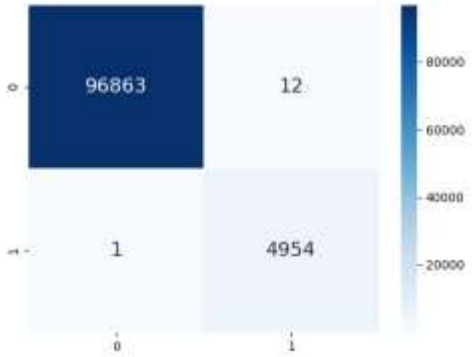


Слика 6.4. Резултати од ML бинарна класификација со IoTID20. Од матриците за конфузија, може да се забележи дека алгоритмот XGBoost е најдобриот алгоритам, а Naive Bayes најлош.

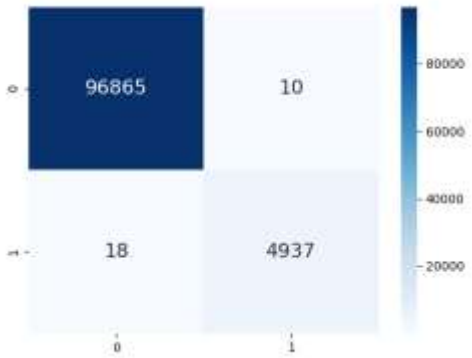
а) XGBoost



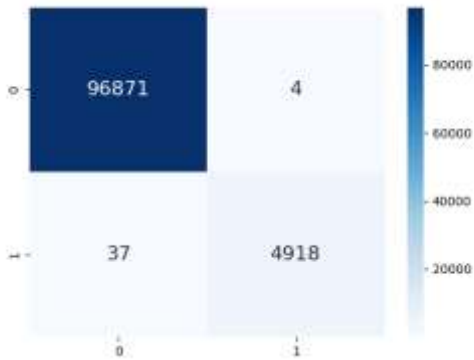
б) K - Nearest Neighbours



в) Naive Bayes

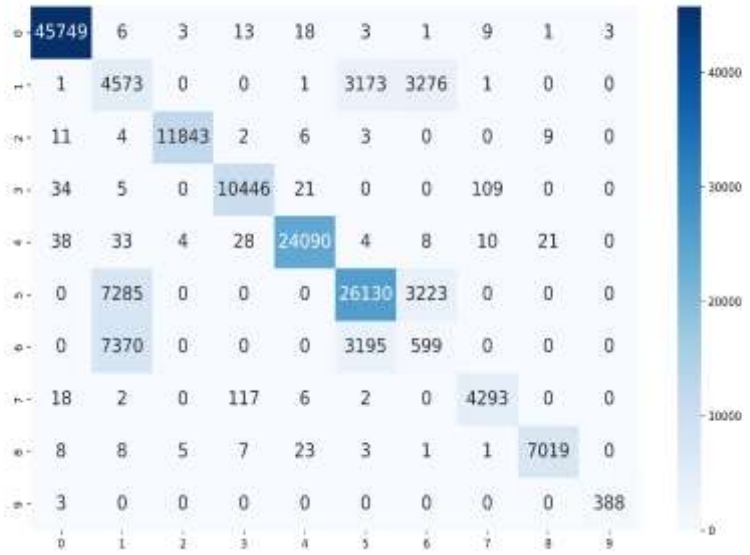


г) Random Forest

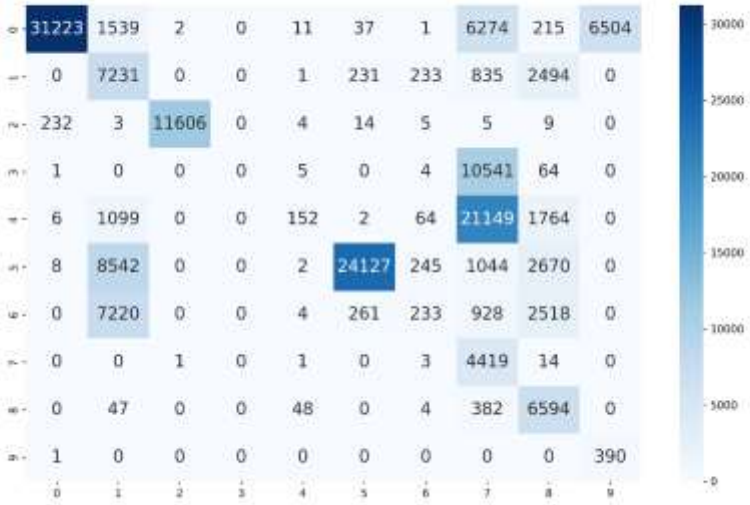


Слика 6.5. Резултати од ML бинарна класификација со N-BaIoT. Од матриците за конфузија, може да се забележи дека алгоритмот XGBoost е најдобриот алгоритам.

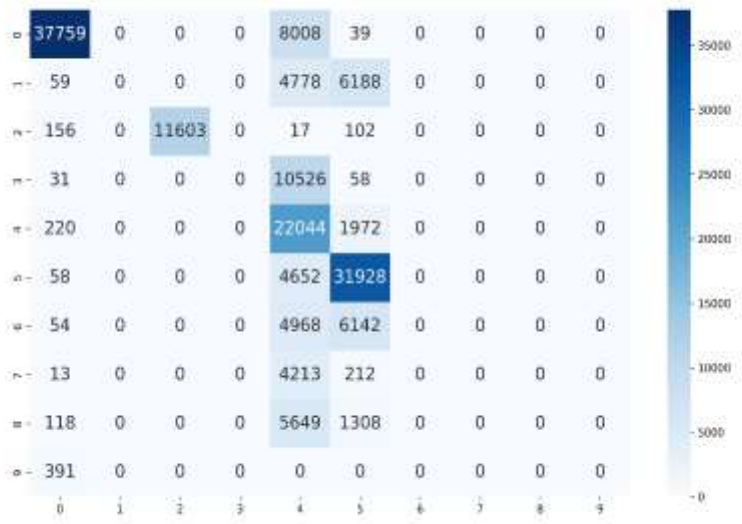
а) K - Nearest Neighbours



б) Naive Bayes

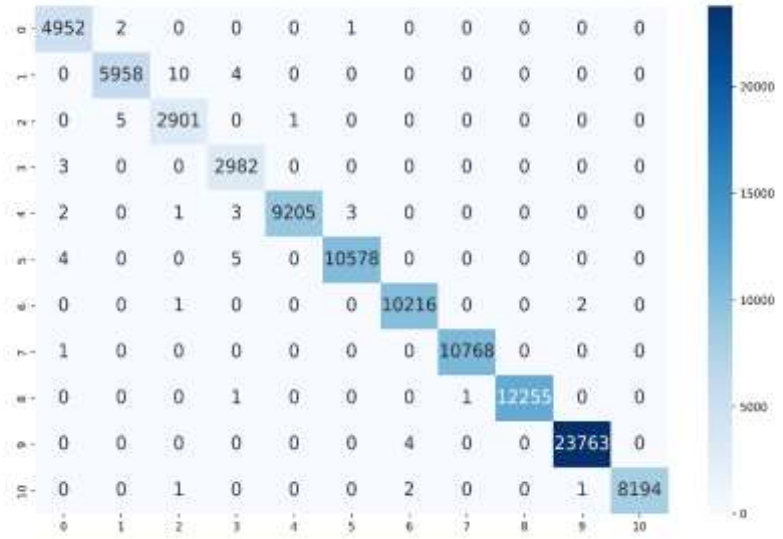


в) Random Forest

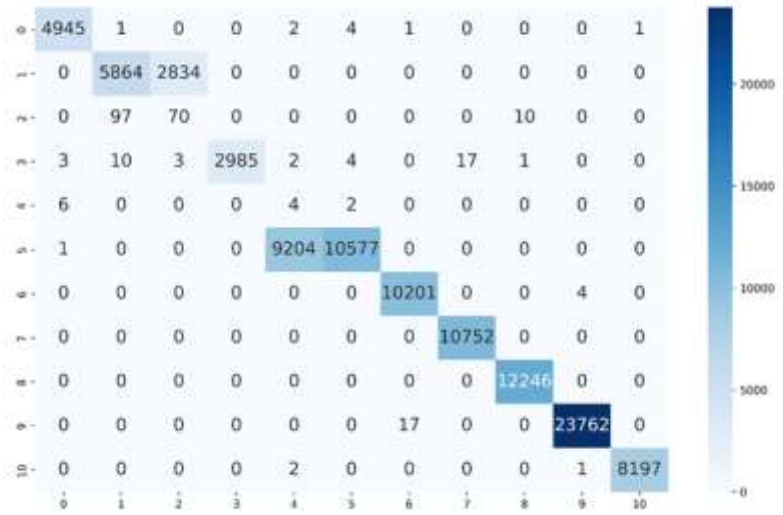


Слика 6.7. Резултати од ML повеќекласна класификација со IoTID20 - K Nearest Neighbours, Naïve Bayes и Random Forest. Може да се забележи дека моделите погрешно ги класифицираат класите 5 и 6.

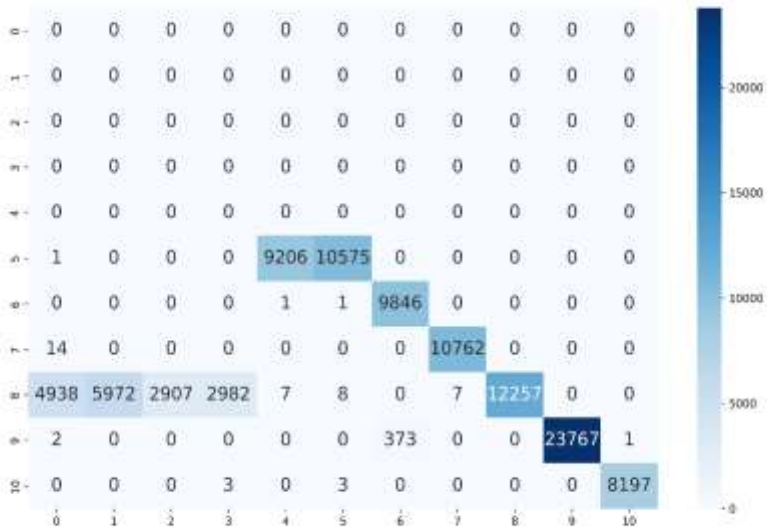
а) K - Nearest Neighbours



б) Naive Bayes

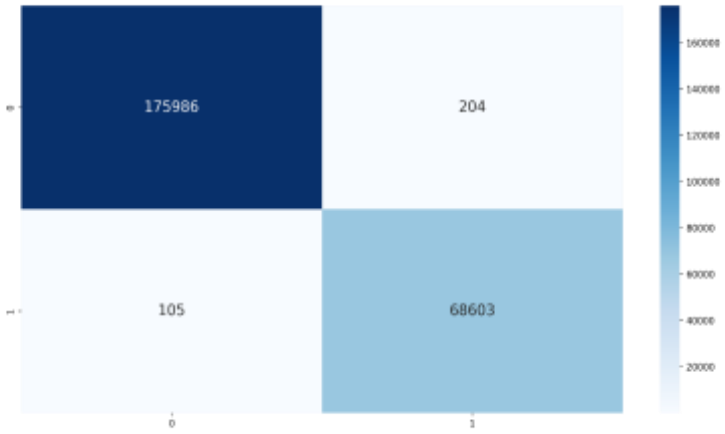


в) Random Forest



Слика 6.8. Резултати од ML повеќекласната класификација со N-BaIoT - K Nearest Neighbours, Naive Bayes и Random Forest. Овде, алгоритмот K - Nearest Neighbors најдобро функционира.

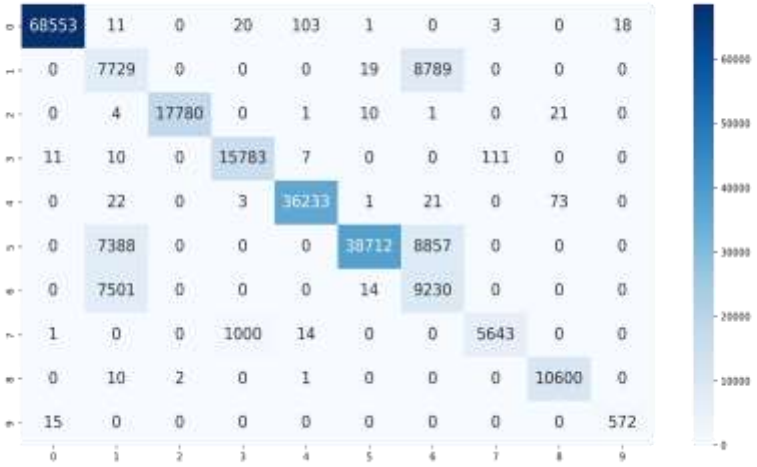
а) Матрица на конфузија - IoTID20



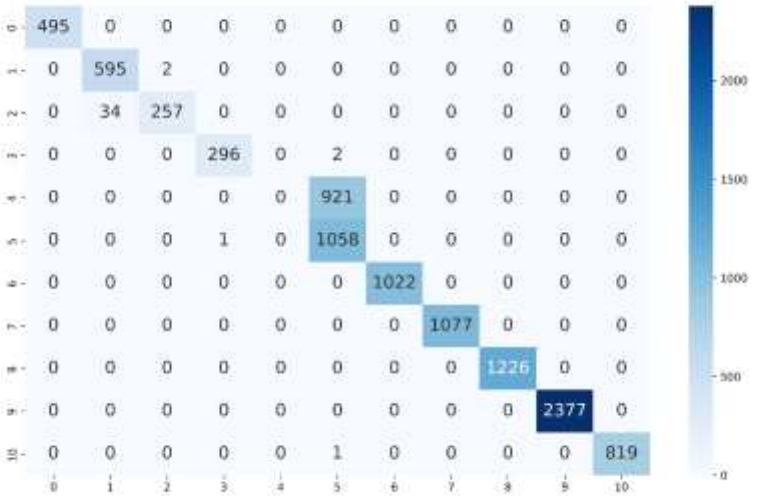
б) Матрица на конфузија - N- BaIoT

Слика 6.11. Резултати од DL бинарната класификација со матрици за конфузија за IoTID20 и N-BaIoT. Матриците за конфузија го потврдуваат трендот од графиците за точност, покажувајќи речиси 100% точност во предвидувањето на класите.

а) Матрица на конфузија - IoTID20

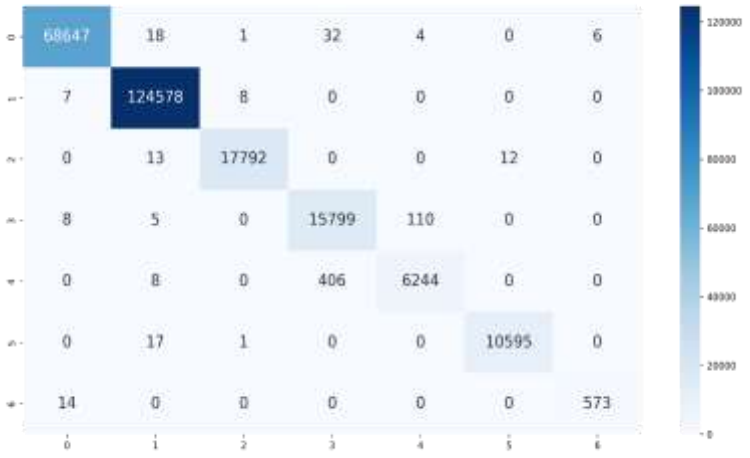


б) Матрица на конфузија - N- BaIoT

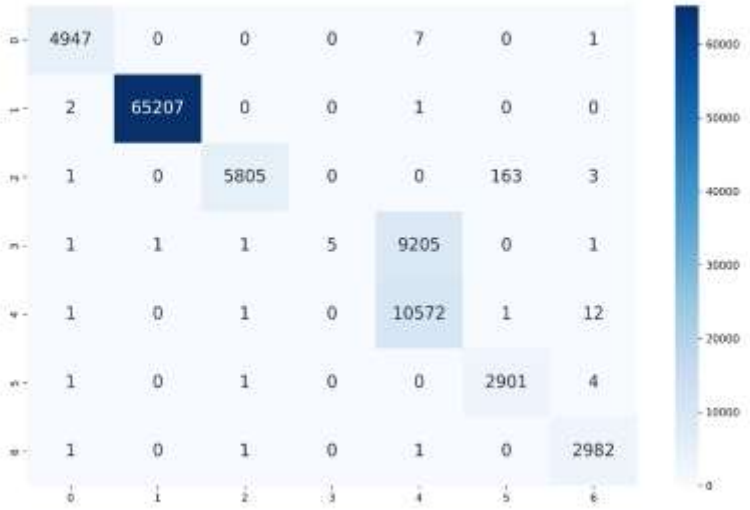


Слика 6.13. DL класификацијата со повеќе класи резултира со матрици за конфузија за IoTID20 и N - BaIoT. Првиот модел погрешно ги класифицира класите 5 и 6, кои одговараат на истата фамилија на напади (Mirai), додека вториот модел погрешно ги класифицира главно класите 4 и 5, кои исто така припаѓаат на истото семејство напади –Mirai.

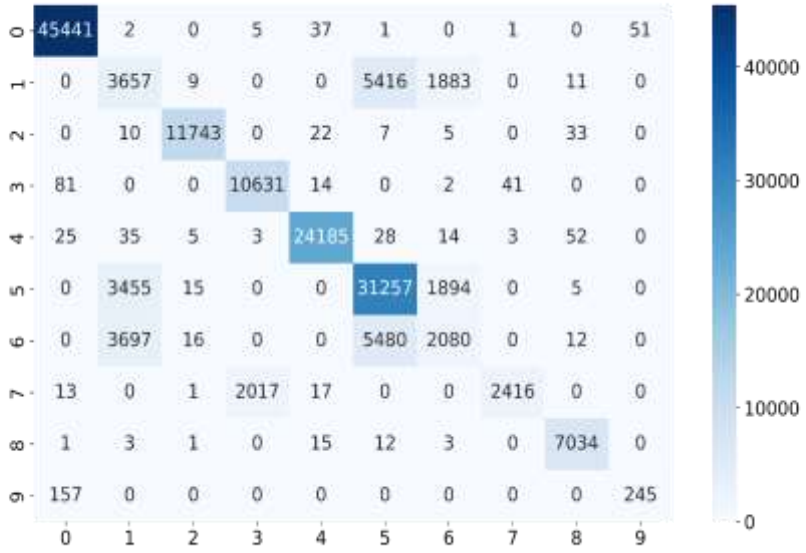
а) Матрица на конфузија - IoTID20



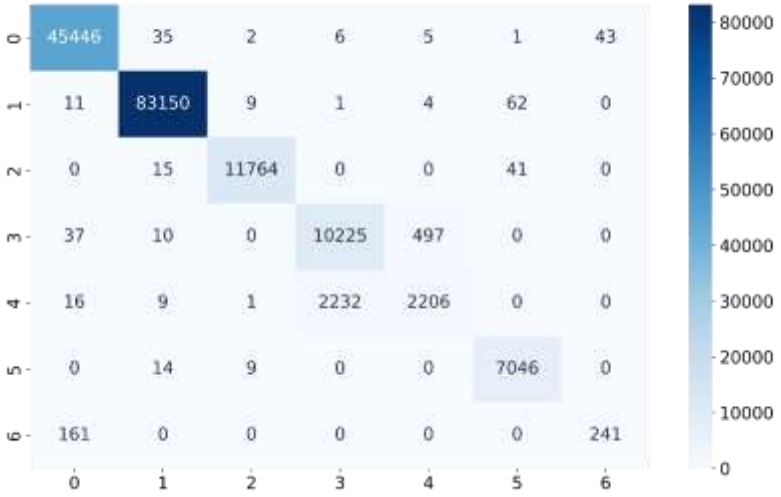
б) Матрица на конфузија - N - BaIoT



Слика 6.15. DL повеќекласна класификација со Mirai групирање - резултати со матрици за конфузија IoTID20 и N-BaIoT. Може да се забележи погрешната класификација на класите 3 и 4, кои всушност се дел од истото семејство на напади на ботнет (Bashlite).

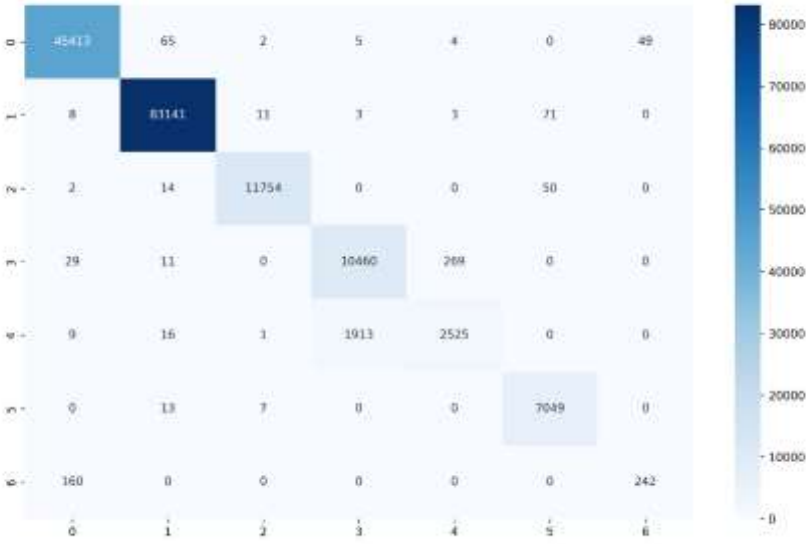


Слика 6.17. Резултати од FL експериментот со мултикласификација. Резултатите покажуваат дека повеќето погрешни класификации се случуваат помеѓу класите 1, 4, 5 и 6, кои одговараат на слични типови на напади, т.е. напади од семејството Mirai.



Слика 6.20. Матрица на конфузија за FL со групирање на Mirai. Групирањето на повеќе Mirai класи во една резултира со значајно зголемување на перформансите за FL моделот.

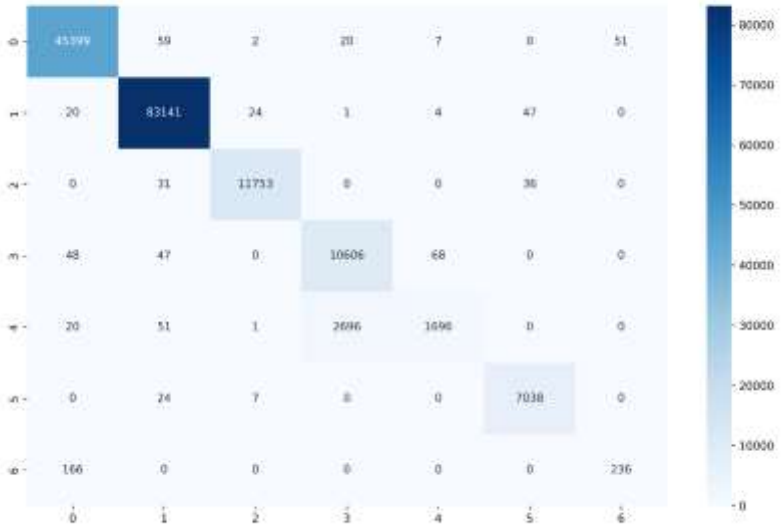
а) Матрица на конфузија за FL со FedProx



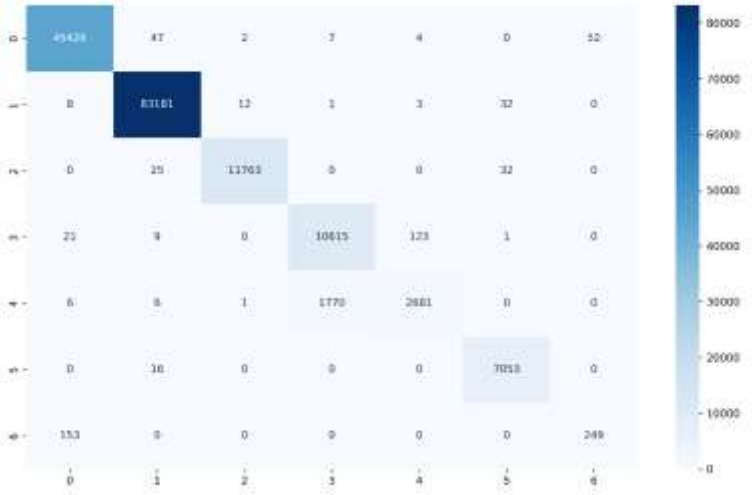
б) Матрица на конфузија за FL со FedMedian

в) Матрица на конфузија за FL со FedYogi

г) Матрица на конфузија за FL со FedAdagrad



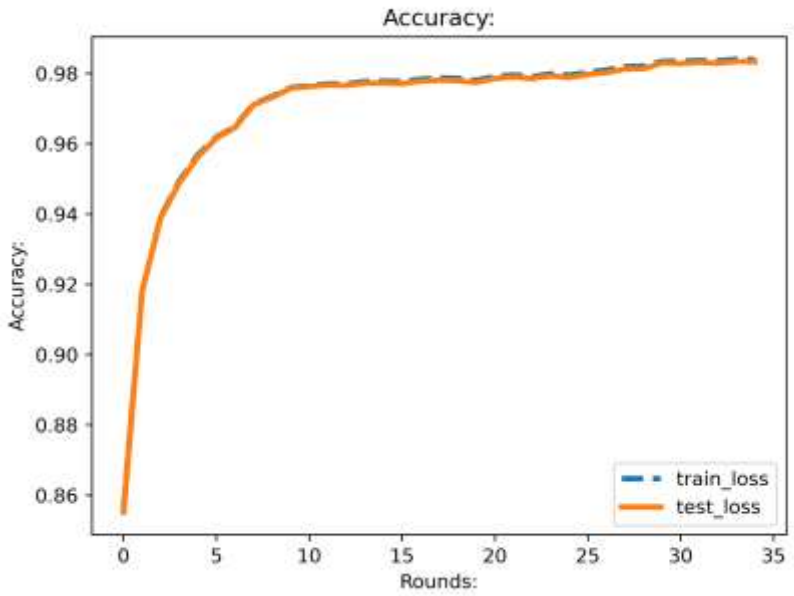
д) Матрица на конфузија за FL со FTFedAvg



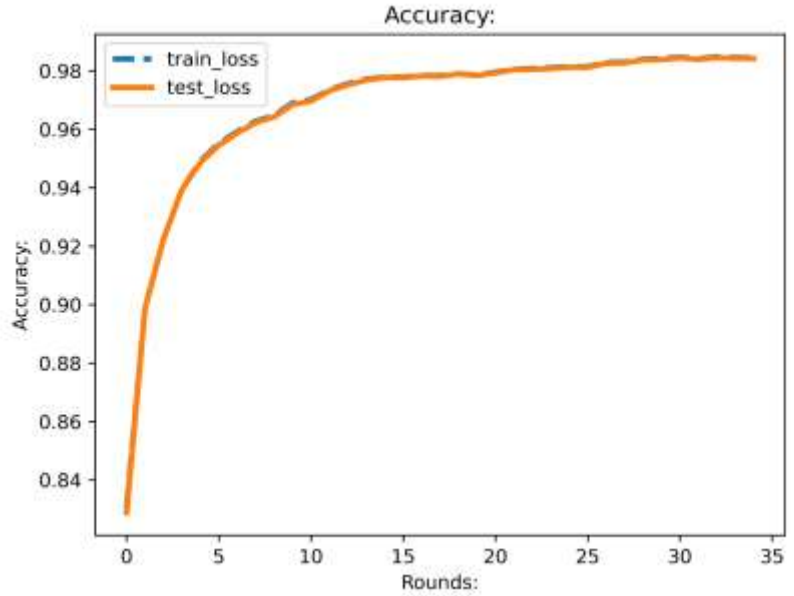
ѓ) Матрица на конфузија за FL со DPFedAvg

Слика 6.24. Матрици на конфузија за различни FL алгоритми. Анализата покажува дека алгоритмот FedYogi дава најдобри резултати.

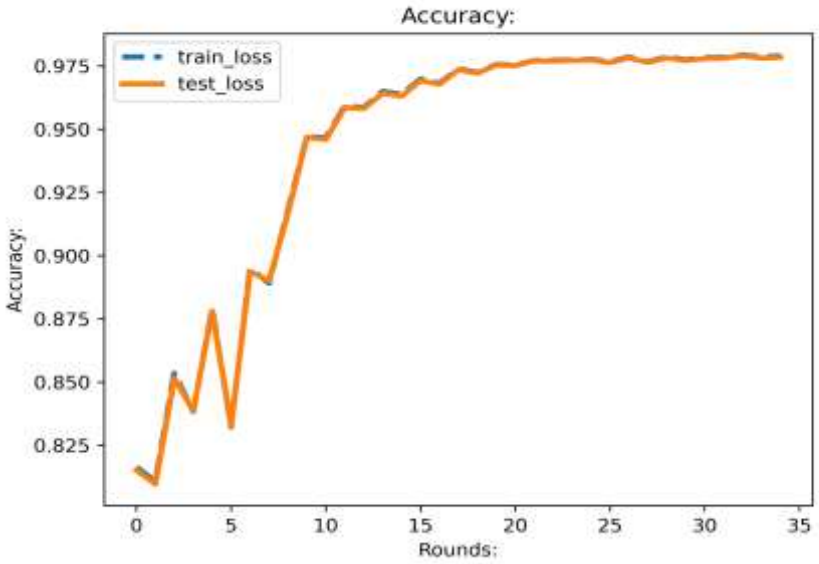
а) График на точност за FL со FedProx



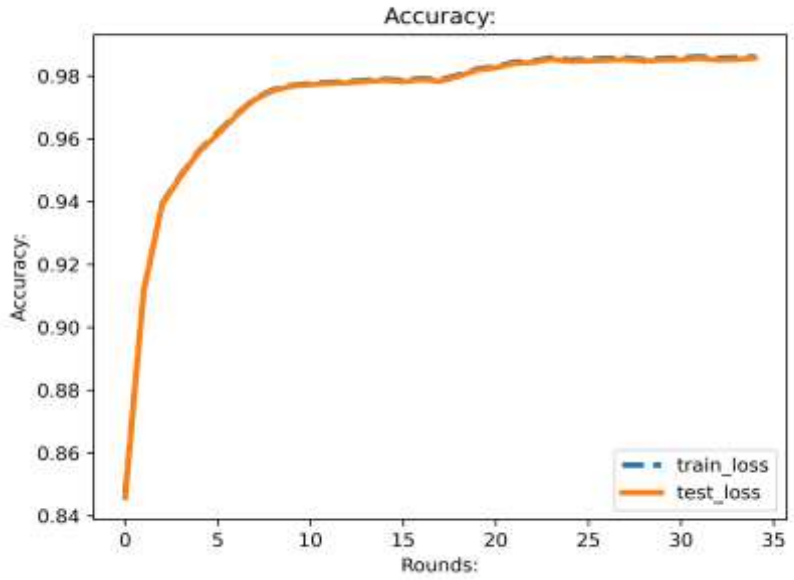
б) График на точност за FL со FedMedian



в) График на точност за FL со FedAdagrad



г) График на точност за FL со FTFedAvg



Слика 6.25. Графици на точност на FL со различни FL алгоритми. Повеќето од FL алгоритмите постигнуваат висока точност, но FedYogi е најдобар во целина.