INTEGRATION OF BLOCKCHAIN AND MACHINE LEARNING IN COMMUNICATION SYSTEMS

Sangeetha Chandramouli
EECS298 - Control, AI and ML
Department of Electrical Engineering and Computer Science
University of California, Irvine

***Abstract*—** This paper focuses on the integration of Blockchain and Machine Learning (ML) which are two evolving technologies that could be used in networks and communication systems. The potential of these technologies is used to provide secure and decentralized sharing of data and models as well as intelligent network operation and management.

Index terms: Blockchain, Machine learning (ML), Mobile Edge Computing (MEC), Distributed Denial-of-Service (DDoS), Natural Language Processing (NLP), Software Defined Networks (SDN).

I. INTRODUCTION

Blockchain is a decentralized ledger controlled by peer-peer networks that enables trust and eliminates the need for a central controller or intermediary. It can help to improve the efficacy and precision of blockchain-based applications. It has found applications in various domains, including ad hoc networks, IoT, healthcare systems, and security services. On the other hand, ML, which is a sub-domain of AI, offers powerful data processing capabilities and can optimize and enhance complex systems and networks. ML techniques such as supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning have been applied in image recognition, virtual reality, and UAV networks, among others.

Integration of blockchain with ML could transform various sectors including networks and communication systems.

By applying blockchain to ML systems, ML can leverage the information covered and manipulated by blockchain to solve the problems efficiently. It also enables distributed processing of data and enhances security and reliability of communication systems. In conventional centralized ML solutions, a central server gathers and processes data from different data sets to train the model. But, getting a lot of data for preparing machine learning models is many times inaccessible in current communication systems. Furthermore, collecting data from heterogeneous networks presents difficulties due to bandwidth constraints and communication costs. The inefficient management of data sharing among various parties is identified as a challenge for machine learning technique development. Blockchain technology can be used to resolve this issue by providing a secure and decentralized way to share data without relying on a third party. By making use of technologies such as timestamping, cryptography, and peer-to-peer networks, blockchains ensure decentralized and tamper-proof timestamping of data.

On the other hand, ML can bring advantages to blockchain in terms of resource efficiency, security and privacy, prevention of malicious attacks such as edge computing, and intelligent smart contracts. ML can contribute to several aspects of blockchain security, such as anomaly detection, intrusion detection, and threat analysis. ML based algorithms that use training data can handle multiple tasks in a smart way. They can quickly predict data, process it, offering miners as a viable means to offer one or more significant transactions. Integrating ML techniques into blockchain has the energy to revolutionize the energy sector making it more efficient and intelligent. ML techniques could enable blockchain applications to support predictive analysis, ensuring accurate fulfillment of energy and resource requirements and enhancing blockchain efficiency. To handle scalability challenges, ML methodologies can optimize data maintenance and storage by providing more efficient solutions like data pruning. ML techniques can also facilitate more efficient off-chain solutions or dynamically adjust the block-size to enhance the scalability of blockchain-based systems. Applying ML algorithms in blockchain based networks and communications systems can detect malicious behaviour on blockchain by deploying trained models and assisting in detecting and preventing fraud and hacking in transactions. ML techniques provide a practical approach in creating and executing complex smart contracts, making them more efficient. Natural Language Processing (NLP) techniques with the assistance of ML can provide smart contract negotiation and construction. Self-writing smart contracts empowered by NLP can securely and cost-effectively facilitate exchanges involving money, property, shares, or any valuable assets. ML could be used for creating smart contracts and enhancing the verification process.

II. PROBLEM DOMAIN

Blockchain can present a secure, transparent, tamper-proof transactions and decentralized platform for sharing data and machine learning models. Research has shown that integration of blockchain and machine learning is innovative, groundbreaking and is set to transform a large number of technical areas including communications and networking systems. Incorporation of machine learning models can assist with sustainability of agreements which were agreed before. blockchain-based machine learning solutions can sort data in a distributed manner without being confined to a single entity data set. By applying blockchain into ML frameworks, machine learning can utilize the data gathered and stored by blockchains to figure out issues more rapidly than any ever before. It can ease sample data and machine learning model sharing, decentralized intelligence, security, privacy, trust and allows for transparent data sharing within a network decision-making.

There are significant impacts of ML on blockchain development, including energy and resource efficiency, scalability, security, privacy, and smart contracts. However, there are some challenges in the integration of blockchain and ML. It includes resource management, big data processing, and security and privacy. There could be other vulnerabilities such as Mobile Edge Computing (MEC) to edge attacks and fake service record attacks as MEC takes advantage of the computation resources of edge devices, for example access points (APs) to lessen service latency, safeguard security of clients and preserve the fundamental advantages of a high-performance cloud. MEC devices are in many cases deployed in physically accessible areas, like base stations or edge servers. This availability increases the risk of physical tampering or unapproved access to the devices. Attackers can gain physical access to MEC devices, exploit their integrity. install malware to steal confidential information or take control of the server. Mobile edge computing (MEC) includes handling and storing of data at the edge of the network nearer to the end-clients. This closeness raises privacy concerns particularly while managing confidential or personal data. If legitimate protection measures are not set up, then there is a chance of unauthorized access and violation of privacy guidelines. Since MEC devices depend upon network connectivity, in particular wireless communication, they are open to various DoS attacks and snooping. Hackers can intercept sensitive information transmitted between MEC devices and other network devices and can disrupt the network connection.

Another common issue that occurs in the integration of blockchain with machine learning is data security and privacy. ML algorithms are often trained on data. Large datasets give more chances to streamline and fine-tuning model hyperparameters. It can be utilized for pretraining models on related assignments, permitting them to learn valuable portrayals. This means it requires access to large amounts of data to achieve optimal performance. The more information they are prepared on, the better they can perform. This is because machine learning algorithms figure out how to recognize designs in the datasets. The more information they have the more patterns they can identify, the better they can make predictions and the more accurate it will be. The amount of data required for ideal performance relies upon the complexity of the machine learning algorithm and the task it is being trained to perform which means more perplexing algorithms and assignments require more data. Large datasets empower models to learn diverse and rich portrayals of the underlying patterns and features in the data. By presenting the model to a large amount of information, it can valuably generalize and catch complex relationships. Thus, by training on a large and diverse dataset, deep learning models can become more robust, and present summed up portrayals, lessening the possibilities of overfitting to explicit examples present in smaller datasets. However, in the blockchain environment, data privacy is a significant concern as all transactions are made on public ledger. This implies that anybody can see who is sending and receiving money, as well as the amount of money being transferred. This can be an issue for people who want to keep their monetary transactions hidden. In light of these exchanges, it may be utilized to build a profile of an individual's spending habits. This data could then be utilized to target the person with advertising or to make decisions about their creditworthiness. Also, if the data is not properly secured, blockchain can be utilized to store an individual’s personal data and this could be accessed by unauthorized individuals.

DDoS attacks pose a great threat to network security causing damage to network operators and internet service providers. There are different types of DDoS attacks among which DNS amplification attacks is the most devastating. In the history of DNS attacks, the biggest recorded DNS attack occurred on October 2, 2016[1], targeting Dyn's servers which resulted in unavailability of internet services such as Amazon, Twitter, etc.[2]. Those attacks not only harm the ISPs, but also lead to significant financial losses for enterprises.

III. DESCRIPTION OF RESULTS

A. Principle of Blockchain-based MEC

In the paper we read that the authors examined blockchain and machine learning based approaches in terms of security issues, for example, verification confidentiality, privacy and access control, data and asset provenance, and integrity assurance. The use of blockchain technology using ML brings a lot of advantages in these situations, such as increased security and efficiency, enhanced privacy, and low costs. Although blockchain and ML are promising technologies applied in communications and networking systems there are still a ton of open issues and challenges that we described in the above section. In order to overcome the challenges of Mobile Edge Computing (MEC) the paper outlines a solution which describes that a blockchain based MEC mechanism is proposed which leverages the reputation of edge nodes recorded on blockchain. Mobile edge computing (MEC) is a novel technology that carries cloud computing abilities to the edge of the mobile network. MEC allows mobile devices to offload computationally escalated undertakings to local MEC servers.

Reinforcement Learning (RL) and Deep Learning (DL) based algorithms are used to optimize the memory allocation in MEC. This mechanism improves the computation of edge devices, reduces response latency, and enhances the performance of edge and mobile devices against security attacks.

B. Zero-knowledge proof to boost data privacy

In blockchain technology each full node that processes transactions and constructs the blockchain essentially has access to the blockchain transaction information itself. This implies that the blockchain is freely accessible and each transaction can be traced back to the primary beginning block. This creates fundamental privacy problems as the data is publicly available to anyone on the internet. One way to mitigate this issue is through the use of privacy-pressing techniques such as zero-knowledge proofs. It refers to a cryptographic idea and protocol where one party, known as the prover, can demonstrate to another party, known as the verifier, that a specific statement is valid without uncovering any additional information apart from the legitimacy of the actual statement itself. In other words, the prover can show knowledge of a confidential or private piece of information without unveiling any details regarding that data. The zero-knowledge protocol is intended to resolve the issue of interactive proof systems, where one party needs to convince another party of the truth of a statement. This guarantees that the prover can convince the verifier without uncovering any extra data beyond the statement's truthfulness. The protocols accomplish this by permitting the prover to produce a proof that is verified by the verifier. The verifier can be guaranteed that the prover has the information or data necessary to validate the claim without really learning any specifics about it. They give a method to establish trust and verify information without the requirement for full disclosure, in this way protecting the privacy and confidentiality of the users. In simple terms for example, if Joe sends M to David, who checks its validity using the algorithm.

Assuming if the algorithm says the proof is legitimate, David is convinced of the truth of M while never knowing S. This allows for a more private, less data sharing and secure blockchain environment.

C. Detection and Mitigation of DDoS attack

1) ML based DDOS detection Module

As networks complexity increases, Software Defined Networking (SDN) has emerged as a trustworthy technology that enables dynamic network management and provides novel approach to address DDoS attack. By decoupling the control plane from the data plane, SDN provides enhanced control over the network, making it an effective tool against attacks like DNS application.

In this paper[3], we present a scalable and efficient SDN-based scheme designed to protect blockchain from DDoS attacks, specifically focusing on DNS amplification attacks, called BrainChain. It leverages SDN's benefits by employing flow statistics collection schemes (FS) using sFlow[4] protocol, reducing exchanges between the data plane and control plane. Additionally, an Entropy Scheme (ES) is introduced to measure the randomness of data, while a real-time detection scheme based on the

ML algorithm of Bayesian Network Filtering Scheme (BF) automatically detects network anomalies. The proposed DNS Mitigation (DM) module effectively mitigates illegitimate flows such as DNS requests.

The security and availability of permissioned blockchain nodes, given their limited number of peers, are critical concerns. While previous works have shown effectiveness against Distributed Reflection Denial-of-Service (DRDoS) attacks, this paper focuses on DNS amplification attacks and utilizes ML algorithms to decrease the FPR while maintaining a high detection rate.

Existing solutions face various limitations, including false positives, performed degradation, and communication overhead. In contrast, BrainChain overcomes these weaknesses by utilizing the scalability of sFlow and ML algorithms, achieving accurate detection and mitigation. The REST API is employed to facilitate the management of any SDN controller for efficient detection and mitigation based on the DM scheme.

2) BrainChain Scheme

The objective of BrainChain is to ensure real-time detection, protecting the limited Ternary Content Addressable Memory (TCAM) of data plane devices, and achieving scalability. The architecture of BrainChain comprises of four schemes namely: Flow Statistics collection (FS), Entropy calculation Scheme (ES), real time detection (BF) and DNS mitigation (DM). The system consists of two phases: an ML-based DDoS detection module that detects illegitimate flows using FS, ES, and BF, and DM that effectively mitigates illegitimate flows for network recovery.

i. FS

Flow statistics are information about the flow of data through the network, such as number of bytes transferred, number of packets sent and the source and destination of traffic. This can be collected at every node in the network and helps in the efficiency of the network and hence provides network security. By collecting flow statistics, nodes can identify malicious activity, such as DDoS attacks or spam attacks. This information can then be used to block malicious traffic and prevent attacks from succeeding. The nodes can also learn about the patterns of traffic and prevent attacks from succeeding. FS can be used to provide transparency into the network which helps in building trust between the users and network. By collecting FS, the nodes can identify the areas where the network is unutilized. By knowing this, the network can be optimized which helps in reducing the cost of implementing.

ii. ES

The purpose of ES is to extract network features from collected information and calculate entropy values. Entropy is a concept of information theory that is used to measure the randomness change of incoming flows within a given time window. During DDoS attacks, the number of flows for a specific flow increases significantly, leading to a more concentrated distribution of IP source addresses. ES calculates entropy values to identify changes in traffic patterns, such as IP source addresses, DNS request types, and UDP source ports, during each monitoring interval.

If IP_{src} is a random variable that denoted number of flows during time interval ΔT, then the entropy is denoted by:

$$H(IP_{src}) = -\sum_{i=1}^{N} p_{i,j}(IPsi, S) log_2 p_{i,j}(IP_{src}S_{\mathrm{j}})$$

iii. BF

BF is a binary classifier that utilizes stateful features (network feature vectors) to classify flows as legitimate or illegitimate. It receives the network X_k at the kth time period ΔT and calculates the probability of illegitimacy using the Bayesian filter. If the probability of illegitimacy exceeds a predefined threshold, BF notifies the SDN controller to deploy mitigation measures against the detected illegitimate flow. BF focuses on detecting abrupt behaviors of attackers who overload the network with DNS traffic.

In BF, each sample is represented by a vector x = (x_1, x_2, x_3) where x_1, x_2, x_3 corresponds to the entropy values of IP source address H'(IP_{src}), H'(Sr c_{port}) and H'(ANY) respectively. These random variables indicate the entropy levels of specific flow properties. BF considers two classes of request namely: illegitimate (illeg) and legitimate (leg). The class of a flow, denoted by c, is determined based on the maximal probability of belonging to that class (p(c/x)). Bernoulli's distribution is a discrete probability distribution that takes the value 1 with probability p and 0 with probability q = 1-p. H'(IP_{src}), H'(Sr c_{port}) and H'(ANY) are conditionally independent given the flow category c. The conditional probabilities of a flow being illegitimate (pi(ill) and pi(leg) during ΔT are trained by BF to classify the network feature X_k as either legitimate or illegitimate.

If the class of flow x is denoted by c, the probability of being in this class i.e., p(c/x) is maximal:

$$c=\text{argmax}_{c\in\{\text{illeg,leg}\}}p(c/x)$$

Since, p(illeg/x)+p(leg/x)=1 implies that selection criteria can be defined as: x is illegitimate if

$$p(\text{illeg}/x)>5$$

Flow x is considered illegitimate if:

$$p(C=c/\mathrm{X}=\mathrm{x})=\frac{p(C=c).p(\mathrm{X}=\mathrm{x}/C=c)}{\Sigma_{c\epsilon\{illeg,leg\}}p(C=c).p(\mathrm{X}=\mathrm{x}/C=c)}>.5$$

In terms of Bernoulli classifier:

$$\frac{\prod_{i=1}^{n}p_i^{x_i}(ill)(1-p_i(ill))^{(1-x_i)}p(ill)}{\prod_{i=1}^{n}p_i^{x_i}(ill)(1-p_i(ill))^{(1-x_i)}p(ill)+\prod_{i=1}^{n}p_i^{x_i}(\mathrm{leg})(1-p_i(\mathrm{leg}))^{(1-x_i)}p(\mathrm{leg})}>.5$$

3) *DDOS MITIGATION SCHEME*

The goal of DM is to effectively reduce the impact of illegitimate data flows. To accomplish this, each data flow entry specifies a meter with a unique Meter_id which allows for monitoring the speed of illegitimate data flows. If the flow rates exceed the defined threshold (rate limiter), DM discards the packets that are suspected to be unlawful.

When BF identifies that X_k is not legitimate, a mitigation action is taken to safeguard the nodes of the blockchain (i.e., Target). To achieve this, new actions from OF actions are implemented using the SDN controller's API. These rules are given high priority to closely monitor the speed of the suspicious data stream.

4) *EXPERIMENTAL RESULTS*

To evaluate the BrainChain module, an experimental environment is used, followed by evaluation of performance of BF in terms of Detection Rate (DR) and False Positivity Rate (FPR) using Receiver Operating Characteristic (ROC) curve.

To assess the performance of BrainChain, a realistic experiment was carried out. The DDoS detection module was implemented as an application on top of the SDN controller. The module utilized the sFlow protocol to control the network traffic statistics and automatically detect illegitimate traffic. For emulating a real network environment, Mininet, a popular SDN emulation tool was used. Mininet employed Linux containers and virtual switches (Open Vswitch) to create a realistic test environment with hosts and OF switches.

The experimental setup involved installing Mininet on a VirtualBox- VM. It was connected to the internet through a Network Address Translation (NAT) for software installation and updates. A host-only adapter was configured on the VM to enable communication with the host system. Secure Shell (SSH) was used for accessing the VM and running different softwares simultaneously. The network monitor (sFlow-RT) and SDN controller (Floodlight) were installed on the host system. The experiments were performed on a PC with an Intel core i7-8750H CPU operating at 2.2GHz and 16GB RAM.

BF is a supervised ML algorithm that requires training. Hence Scapy's Python library was used to forge DNS requests, and the dnsd package in Node.js was employed to create a DNS test server. The DNS server was used to send amplified DNS responses to the victim nodes in the blockchain.

It was observed that with the control disabled, the attack traffic exceeded 2000 packets per second (illegitimate DNS). However, when BrainChain was enabled, the attack traffic was halted as BF classified the kth vector X_k as illegitimate and instructed the SDN controller to mitigate the traffic. BrainChain pipeline's detection and mitigation process took less than 13 seconds, enabling the network to recover quickly.

The performance of BF was evaluated using the ROC curve. Two experiments were conducted with different attack rates (100 Mbps and 500 Mbps), and BF's accuracy and False Positive Rate (FPR) were compared with previous results. The ROC curve illustrates the True Positive Rate (TPR), also known as sensitivity or Detection Rate (DR), against FPR which is also referred to as anti-specificity. In order to determine BF performance, DR and FPR were defined as follows[6]:

DR = TP / (TP + FN)
FPR = FP / (TN + FP)

Here, TP represents correctly classified illegitimate flows (illegitimate DNS requests), FN represents illegitimate flows mistakenly classified as legitimate, FP represents legitimate flows (legitimate DNS requests) incorrectly identified as illegitimate, and TN represents correctly classified legitimate flows.

IV. CRITICAL REVIEW

In this paper we analysed that the authors have introduced blockchain technology including the key prerequisites, consensus mechanisms, and existing blockchain platforms. They researched about decentralized consensus mechanisms in blockchain frameworks, such as Byzantine Fault Tolerance (BFT)-based consensus protocols, Nakamoto protocols, virtual mining protocols, hybrid protocols and a series of parallel consensus protocols. They give a far-reaching review on the network layer of permissionless blockchains. Considering network performance and requirements such as minimal expense of investment anonymousness, and topology hiding, they present a plan of the network layer of permissionless blockchains to work on the capability and security of frameworks.

Similarly, ML techniques can be used in blockchain systems by providing powerful intelligence for data processing and execution of complex tasks. There are myriad applications that can benefit from using ML in blockchain to protect their exchanges. These mechanisms can process different types of transactions and enhance the performance of smart systems. Table 1 provides a brief correlation of existing overviews on blockchain and machine learning.

We also read about Ocean protocol which is an example of a blockchain-based platform that can be utilized to share information for ML applications. It is based on three layers:

The Service layer - takes care of the service agreements, low-level access control, records, balances, and the block reward.

The Verification layer - introduces cryptographic difficulties to enhance the morals and security of the services.

The Curation layer - works as a discovery component as well as signalling and governance perspectives.

However, the Ocean protocol is still a work in progress, but it has the capability to reform the way that data is shared for machine learning applications.

On the other hand, like any detection mechanism, ES can suffer from false positives and false negatives. False positives occur when legitimate traffic is incorrectly classified as malicious, leading to unnecessary mitigation measures. In other words, it could happen if the normal traffic patterns exhibit characteristics that resemble those of a DDoS attack. False negatives occur when the detection system fails to identify an actual DDoS attack and classifies it as normal traffic. In the context of entropy calculations, false negatives can occur if the attack traffic patterns do not significantly deviate from the normal traffic patterns in terms of entropy. If the DDoS attack traffic exhibits similar entropy characteristics as legitimate traffic, the detection system may fail to detect the attack, leading to a false negative. Achieving an accurate balance between sensitivity and specificity in detecting DDoS attacks can be challenging. By combining ES and BF, real-time attack destinations with a low false positive rate and high detection rate can be achieved.

Another drawback, FS method might not be suitable for detecting high-rate DDoS attacks as it can exhaust the bandwidth and TCAM resources. To address this, FS utilizes flow sampling technologies with sFlow. By separating flow monitoring from forwarding logic, FS improves efficiency and scalability. sFlow performs flow aggregation, which is crucial during high-rate DDoS attacks.

Figure 1b shows that BrainChain achieves approximately 100% detection rate for the 100 Mbps case with a false positive ratio of only 23%. However, Chainsecure[5] achieves the same detection rates with false positive ratio of 31% and 40% respectively. According to **Figure 1c**, BrainChain achieves approximately 100% detection rate for the 500 Mbps case with false positive ratio of only 21%, while ChainSecure reaches similar detection rate with false positive ratios of 30% and 34% respectively.

V. CONCLUSION

Blockchain technology is an emerging opportunity for machine learning algorithms to be applied in communications and networking systems. It can be utilized to create a secure and tamper-proof record of exchanges, which can assist in preventing fraud and unauthorized access.

Blockchain is a worldwide data set in which all network nodes can hold and trade information in a way that is decentralized and verifiable. It can help data and model sharing, security and privacy, decentralized intelligence, and help in trustful decision-making. Having access to a large amount of data for training machine learning models may not be available in present communication systems. Moreover, aggregating data in heterogeneous systems for machine learning training is also an issue. In particular, the conventional centralized ML solutions, the bandwidth restrictions and the communication cost essentially restrict the aggregation of data in heterogeneous systems. Getting this information lies in gathering, organizing, and evaluating the data for precision. But, ML can be used to identify and characterize malicious traffic, as well as to identify and prevent intrusions.

Permissioned blockchain has a limited number of peers, making them vulnerable to DDOS attacks which can prevent the working of blockchain operation. To prevent this attack, FS is used to significantly gather the feature flows. This is followed by design of real-time detection scheme using ES-BF combination to identify malicious flows. Finally, DNS mitigation (DM) is implemented to block illegitimate attacks. Evaluation is done using real-world data and found that it can effectively identify and prevent DNS amplification attacks.

VI. FUTURE DIRECTIONS

BrainChain is a new blockchain technology that makes use of machine learning algorithms such as Neural Networks, to improve the efficiency and security of networks. It is a scalable and efficient scheme to protect the blockchain from DNS attacks in SDN. But the implementation appears to be incomplete due to some of the common issues that remain unsolved. One of the biggest drawbacks is the Centralization. The BrainChain algorithm relies on central authority to train the neural network. That is, the central network has the control authority has the control over the security of the network. If the central authority is compromised, there is a probability that it would allow the attacker to manipulate the neural network and gain control over the network. The neural network could also make some incorrect decisions. An attacker could use adversarial attacks to create malicious traffic that looks like legitimate traffic to the neural network. This could allow the attacker to jam the network with malicious traffic and bring down the network. In future as the network size will increase, more processing of data will be required which could affect the network and hence making it more vulnerable to attack. Overall, the BrainChain mechanism could be a promising one but it's important to resolve the flaws discussed before implementing it on a large scale. Mobile Edge Computing (MEC) can be utilized to give a more secure environment to blockchain. For instance, MEC servers can be utilized to encrypt information before it is transmitted over the network. MEC servers can also be deployed to filter out malicious traffic. MEC is a promising new innovation that can possibly change the blockchain technology thus it should be further improved upon to reduce the network congestion, and improve security and provide a seamless user experience.

Blockchain technology has opened up unlimited doors and potential in different fields because of technological advancements in current computing technologies and storage systems. However, such models have shown constraints concerning high expenses and less resource efficiency. Presently, existing ML models require high-performing computing devices for the purpose of training and blockchain is a costly storage medium. More research is required to propose budget friendly, resource-friendly, quick, and high-performance-based blockchain assisted machine learning frameworks.

VII. REFERENCES

[1] B. Schneier, *Lessons from the Dyn DDoS Attack, Mai. 2019*.

[2] S. Sharwood, *GitHub Wobbles Under DDOS Attack, Mai. 2019*.

[3] Zakaria Abou El Houda, Abdelhakim Hafid, Lyes Khoukhi, *“BrainChain - A Machine learning Approach for protecting Blockchain applications using SDN”* ,IEEE International Conference on Communications (ICC), 2020.

[4] ‘sFlow’ Available: https://www.ietf.org/rfc/rfc3176.txt

[5] Z. A. El Houda, L. Khoukhi and A. Hafid, *“ChainSecure - A Scalable and Proactive Solution for Protecting Blockchain Applications Using SDN”* , 2018 IEEE Global Communications Conference (GLOBECOM), pp. 1-6, 2018.

[6] K Giotis, C Argyropoulos, G Androulidakis et al., “Combining OpenFlow and sFlow for an effective and scalable anomaly detection and mitigation mechanism on SDN environments” , Computer Networks, vol. 62, pp. 122-136, 2014.

[7] Yiming Liu , F. Richard Yu , Fellow, IEEE, Xi Li , Hong Ji , Senior Member, IEEE, and Victor C. M. Leung , Fellow, IEEE, “Blockchain and Machine Learning for Communications and Networking Systems”, IEEE COMMUNICATIONS SURVEYS & TUTORIALS, VOL. 22, NO. 2, SECOND QUARTER 2020

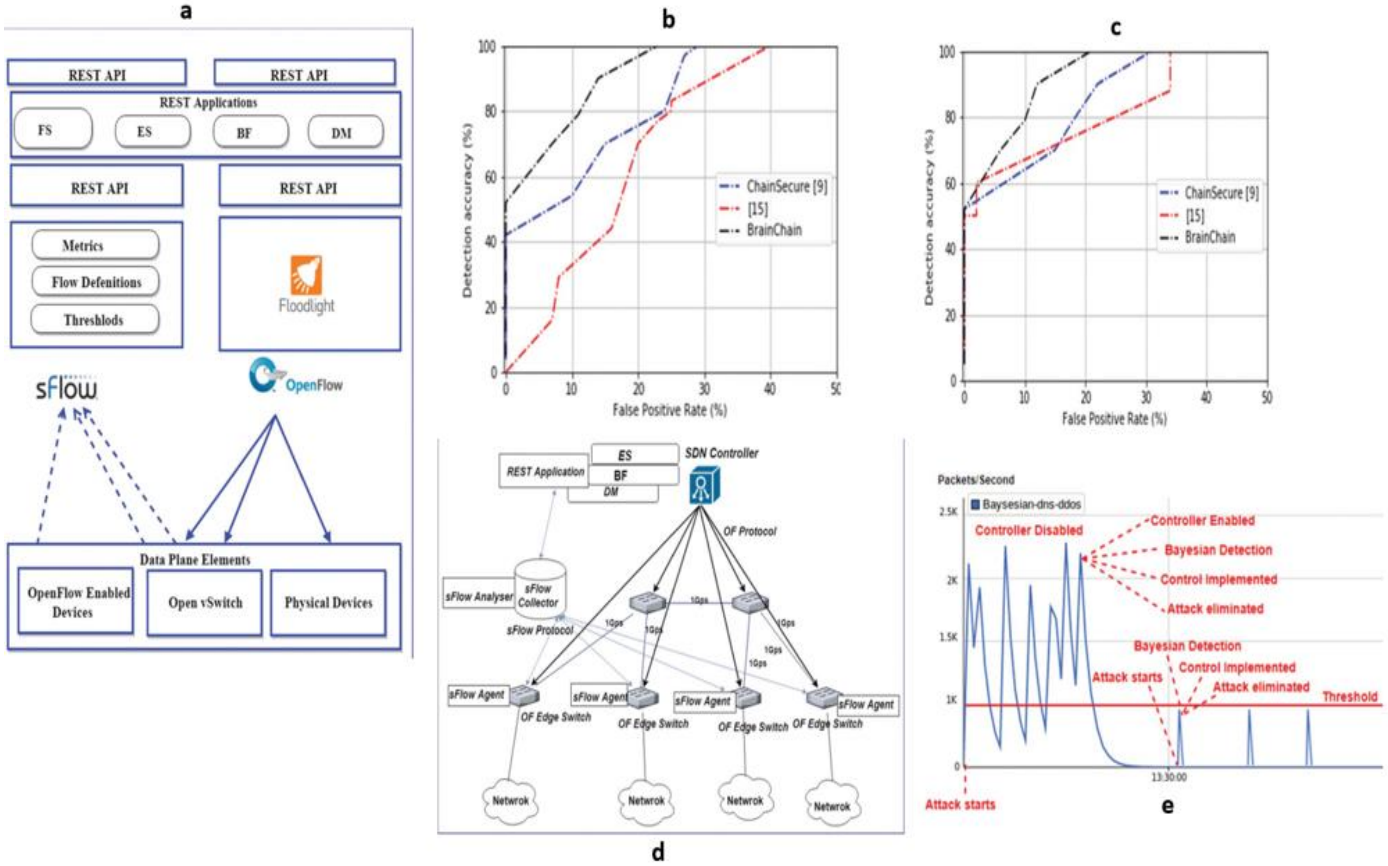
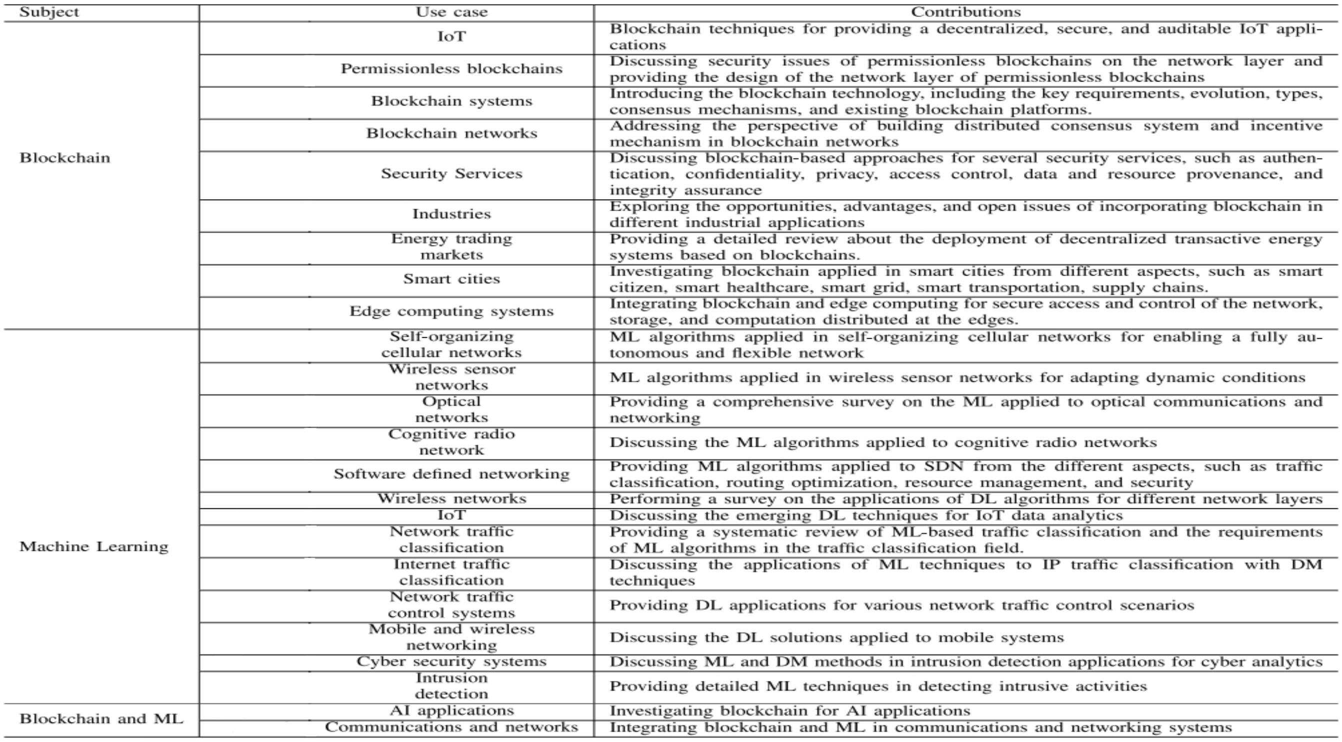


Figure 1: (a) System Overview (b) ROC curve for 100 Mbps (c) ROC curve for 500 Mbps (d) Experimental Setup (e) Blockchain network traffic before and after BrainChain



Subject	Use case	Contributions
Blockchain	IoT	Blockchain techniques for providing a decentralized, secure, and auditable IoT applications
	Permissionless blockchains	Discussing security issues of permissionless blockchains on the network layer and providing the design of the network layer of permissionless blockchains
	Blockchain systems	Introducing the blockchain technology, including the key requirements, evolution, types, consensus mechanisms, and existing blockchain platforms.
	Blockchain networks	Addressing the perspective of building distributed consensus system and incentive mechanism in blockchain networks
	Security Services	Discussing blockchain-based approaches for several security services, such as authentication, confidentiality, privacy, access control, data and resource provenance, and integrity assurance
	Industries	Exploring the opportunities, advantages, and open issues of incorporating blockchain in different industrial applications
	Energy trading markets	Providing a detailed review about the deployment of decentralized transactive energy systems based on blockchains.
	Smart cities	Investigating blockchain applied in smart cities from different aspects, such as smart citizen, smart healthcare, smart grid, smart transportation, supply chains.
	Edge computing systems	Integrating blockchain and edge computing for secure access and control of the network, storage, and computation distributed at the edges.
Machine Learning	Self-organizing cellular networks	ML algorithms applied in self-organizing cellular networks for enabling a fully autonomous and flexible network
	Wireless sensor networks	ML algorithms applied in wireless sensor networks for adapting dynamic conditions
	Optical networks	Providing a comprehensive survey on the ML applied to optical communications and networking
	Cognitive radio network	Discussing the ML algorithms applied to cognitive radio networks
	Software defined networking	Providing ML algorithms applied to SDN from the different aspects, such as traffic classification, routing optimization, resource management, and security
	Wireless networks	Performing a survey on the applications of DL algorithms for different network layers
	IoT	Discussing the emerging DL techniques for IoT data analytics
	Network traffic classification	Providing a systematic review of ML-based traffic classification and the requirements of ML algorithms in the traffic classification field.
	Internet traffic classification	Discussing the applications of ML techniques to IP traffic classification with DM techniques
	Network traffic control systems	Providing DL applications for various network traffic control scenarios
	Mobile and wireless networking	Discussing the DL solutions applied to mobile systems
	Cyber security systems	Discussing ML and DM methods in intrusion detection applications for cyber analytics
	Intrusion detection	Providing detailed ML techniques in detecting intrusive activities
Blockchain and ML	AI applications	Investigating blockchain for AI applications
	Communications and networks	Integrating blockchain and ML in communications and networking systems

Table 1[7]: Correlation of existing overviews on blockchain and machine learning