International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

# Early Detection of Cyber-Attack by Using Machine Learning and Blockchain Technology

**Servesh Gupta[1], Pratik Tawde[2]**

[1]Information Technology Department, Vidyalankar Polytechnic, Wadala, India
[2]Electronics and Telecommunication Department, Vidyalankar Polytechnic, Wadala, India

**Abstract:**
Blockchain has moved beyond the hype to real-world implementation in a broad range of industries (eg. Finance, supply chain management and Internet of Things) and applications (eg. cybersecurity and digital forensics), partly evidenced by its evolution over the last decade or so. When writing this foreword, we are already in the fourth generation of blockchain (also known as Blockchain 4.0) and inching towards the fifth generation of blockchain. This study looks into the current, and potential uses of Blockchain technology in business, specifically in cybersecurity. The highlight of the documentation is the technology which is integrated not only to prevent fraud but also to predict it, using machine learning this technology is known as 'Federative Learning'.

**Keywords**: Cybersecurity, Federative Learning

## I. Introduction

Blockchain technology is defined as a decentralized system of distributed registers that are used to record data transactions on multiple computers. Blockchain technology has infiltrated all areas of our lives, from manufacturing to healthcare and beyond. Cybersecurity is an industry that has been significantly affected by this technology and maybe more so in the future.

Blockchain technology promises a new dimension of conducting business transactions among untrusted entities; its features that support verification, identification, authentication and integrity are guaranteed through cryptography, transparency and decentralized smart contracts and smart ledgers.

Blockchain technology comprises tamper-proof and tamper-evident digital ledgers executed without a central repository, as a distributed system, and often without a central authority, such as a government, a bank or a firm. It allows users in a community to store transactions in a shared ledger within that community.

There are various types of blockchain such as private, public/permissionless and federated/consortium blockchain. Private blockchain include Multichain, Monax, and Quorum. Public blockchain allows anyone to write or read the data stored in the blockchain network, without any permission from any authority and the operation is entirely decentralized and anomalous. Some examples are Monero, Ethereum, and Bitcoin.

Bitcoin is probably the most well-known example of a public blockchain and it achieves consensus through "bitcoin mining". Computers on the bitcoin network, or "miners," try to solve a complex cryptographic problem to create proof of work and thereby validate the transaction. Outside of public keys, there are few identity and access controls in this type of network.

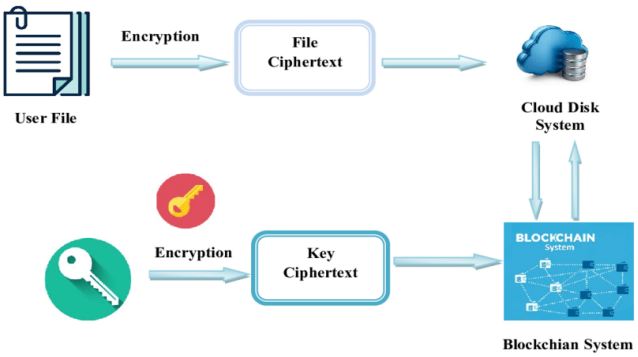


**Fig 1: Block Diagram**

## II. Cyber Attacks and Frauds

Blockchain networks are not immune to cyberattacks and fraud. Those with ill intent can manipulate known vulnerabilities in blockchain infrastructure and have succeeded in various hacks and frauds over the years. Here are few examples like Phishing attack, Routing attack, Sybil attack and 51% attack.

**Phishing Attack:** Phishing is a scamming attempt to attain a user's credentials. Fraudsters send wallet key owners emails designed to look as though they're coming from a legitimate source. The emails ask users for their credentials by using fake hyperlinks. Having access to a user's credentials and other sensitive information can result in losses for the user and the blockchain network.

The Nordea Bank Incident- In 2007, Swedish bank Nordea lost over 7 million kronor when phishers managed to send fraudulent emails out to bank customers, luring them to install the "haxdoor" Trojan disguised as anti-spam software.

**Routing Attack:** Blockchains rely on real-time, large data transfers. Hackers can intercept data as it's transferring to internet service providers. In a routing attack, blockchain participants typically can't see the threat, so everything looks normal. However, behind the scenes, fraudsters have extracted confidential data or currencies.

**Sybil Attack:** In a Sybil attack, hackers create and use many false network identities to flood the network and crash the system. Sybil refers to a famous book character diagnosed with a multiple identity disorder.

This type of attack aims to undermine the authority or power in a reputable system by gaining the majority of influence in the network. The fake identities serve to provide this influence.. A notable example is the 2017–2021 attack run by threat actor KAX17. This entity controlled over 900 malicious servers, primarily middle points, in an attempt to deanonymize Tor users.

**51% Attack:** Mining requires a vast amount of computing power, especially for large-scale public blockchains. But if a miner, or a group of miners, might rally enough resources, they might attain more than 50% of a blockchain network's mining power. Having more than 50% of the power means having control over the ledger and the ability to manipulate it.

**Note**: Private blockchains are not vulnerable to 51% attacks.

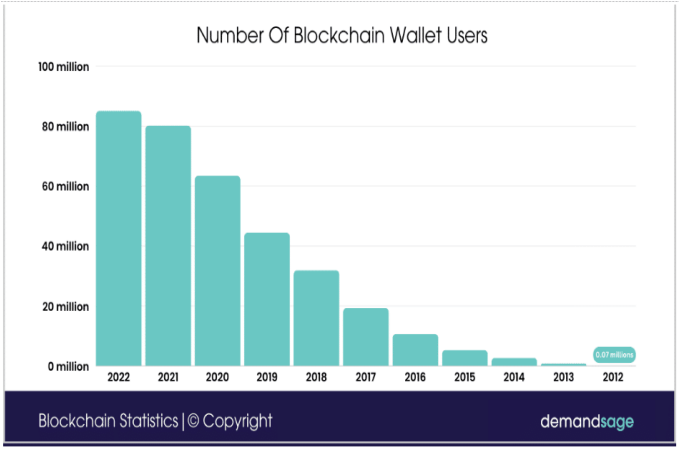


**Fig 2: Graphical Representation**

## III. How to Prevent Cyberattacks and Frauds

### 1. Enhancing Consensus Mechanisms

Blockchain consensus mechanisms like Proof of Work (PoW) and Proof of Stake (PoS) ensure that transactions are validated correctly. However, they can be vulnerable to attacks such as 51% attacks (in PoW) or Sybil attacks (in PoS).

A consensus mechanism is a self-regulatory stack of software protocols written into a blockchain's code that synchronizes a network into agreement about the state of a digital ledger. This is done by upkeeping a single data set — the mutually agreed-upon version of a blockchain's transaction history — rather than employing each node, or in-network computer, to individually maintain their own copy of the database in its entirety.

Although there are a variety of consensus mechanisms to consider when programming a network's standard for verification, each approach is wired to discredit cheaters in their attempts to contradict the record.

**Types of Consensus Mechanisms**

**1. Proof of Work (PoW)**

PoW requires participants (called miners) to solve complex mathematical problems to validate transactions and create new blocks. The first miner to solve the problem gets to add a block to the chain and is rewarded with cryptocurrency.

Used by: Bitcoin, Ethereum

**2. Proof of Stake (PoS)**

In PoS, validators are chosen to create new blocks and validate transactions based on the number of tokens they "stake" (lock up as collateral). The more tokens you stake, the higher your chances of being selected to validate a block.

Used by: Ethereum (post-merge), Cardano, Tezos.

**3. Delegated Proof of Stake (DPoS)**

In DPoS, token holders vote to elect a small number of delegates who are responsible for validating transactions and creating blocks. The elected delegates (validators) perform these tasks on behalf of the network.

Used by: EOS, Tron, Steemit.

**4. Proof of Authority (PoA)**

PoA relies on a small number of trusted validators who are pre-approved by the network to validate transactions and create new blocks. Validators are typically well-known entities and are required to maintain their reputation

Used by: VeChain, private and consortium blockchains.

**5. Proof of History (PoH)**

PoH creates a historical record that proves an event has occurred at a specific time. Nodes can then verify the historical order of transactions without requiring intensive computation.

Used by: Solana.

**2. Multi-signature (Multisig) Wallets**

A multi-sig wallet requires multiple private keys to approve a transaction, making it harder for attackers to compromise funds.

Multi-sig wallets are an advanced security tool that aims to provide an added layer of protection for cryptocurrency users. They may be particularly useful for businesses or groups that need to safeguard their digital assets. However, like any tool, they need to be used properly to be effective. Understanding how multi-sig wallets work, their potential benefits, and potential risks is crucial for anyone considering their use.

When used properly, a multi-sig wallet aims to offer additional security by eliminating the single point of failure risk associated with having one private key. It makes it difficult for hackers to steal funds from a wallet, because they must have the different keys to complete any action. This feature is especially desirable when the assets belong to multiple parties in a company.

**3. Decentralized solution**

A blockchain allows the data in a database to be spread out among several network nodes—computers or devices running software for the blockchain—at various locations. This not only creates redundancy but maintains the fidelity of the data. For example, if someone tries to alter a record at one instance of the database, the other nodes would prevent it from happening because they compare block hashes. This way, no single node within the network can alter information within the chain. Because of this distribution—and the encrypted proof that work was done—the information and history (like the transactions in cryptocurrency) are irreversible.

**4. Cryptographic hashing**

The cryptographic hash function in Blockchain is a way to secure the message block and is used to connect the blocks in a chain. Briefly, In the blockchain, each block contains its own block hash and a hash of its previous block. It helps them to form a cryptographically secured linear chain of blocks. The figure below shows an example of how cryptographic hashing works.

Hashing is converting an original piece of data into a digest or hash. The process uses cryptographic hash functions for the irreversible conversion of the message.

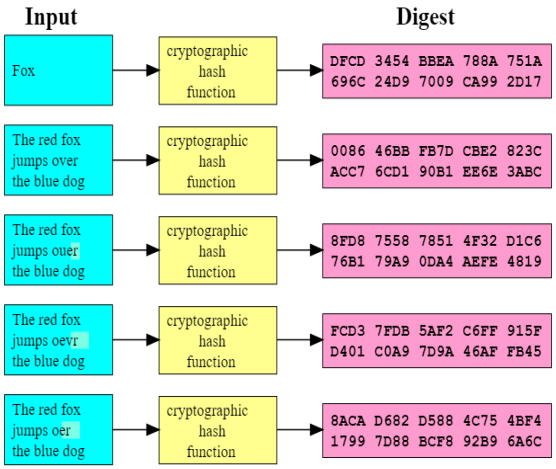


**Fig 3: Conceptual Structure and Output**

Now, that we have understood ways to prevent fraud and attacks, it is also essential to develop predictive model for the system so that the system is much transparent and trustworthy.

**Need to develop a predictive model:**

Predictive analytics on the blockchain helps mitigate uncertainty. By quantifying potential risks and forecasting demand shifts, businesses can make informed decisions, minimize losses from stockouts or oversupply, and confidently capitalize on emerging opportunities within the dynamic decentralized market.

Developing a predictive model for blockchain in cybersecurity can provide significant advantages in terms of threat detection, anomaly identification. and overall system security.

## IV. The Role of Predicting Analytics in Blockchain for Cybersecurity:

### 1. Data: The backbone of prediction

Predicting models work best with large amounts of good quality data. Blockchain provides a rich source of on-chain data such as transaction histories and smart contract executions. Off-chain data like market trends & social media activity, should also be included to strengthen predictions. The challenge is verifying and integrating this external data while maintaining the model's accuracy and security.

### 2. Statistical Foundations

Analyzing historical data and forecasting future trends are crucial for analyzing on-chain data, helping to identify patterns like fluctuations in gas fees or periods of high network demand. These methods establish a baseline for more advanced predictions.

### 3. Machine Learning Advantage

Machine learning models like neural networks and decision trees are good at uncovering hidden patterns within decentralized systems. By analyzing vast amounts of on-chain and off-chain data, these models reveal complex relationships that traditional methods might miss, such as correlations between token staking behaviours and market trends.

### 4. Probabilistic Outcomes for Risk Management

Predictive models in blockchain should offer a range of possible scenarios with associated probabilities. This helps businesses manage risk and automate responses through smart contracts when thresholds are reached, such as adjusting inventory or liquidity.

## V. Methods for Predicting an Attack on a system or user

**Federated learning**

Federated learning (often referred to as collaborative learning) is a decentralized approach to training machine learning models. It doesn't require an exchange of data from client devices to global servers. Instead, the raw data on edge devices is used to train the model locally, increasing data privacy.

**How does federated learning works?..**

A baseline model developed as standard is held in the central server. The copies of this model are given to the client devices which then, train the models from the local data they generate. Through this process, the models on the various devices become unique and meet user needs more effectively.

In the next step, the local training updates (model parameters) are securely aggregated and sent back to the central server into a single model. This architecture takes several inputs in a particular form and processes them in order to 'learn' and make other new combinations. When the data comes from different sources, a better generalization can be obtained in the model.

When the centralized model has been re-trained with a new set of parameters, it is once more released to the client devices for the next iteration. In every cycle, the models acquire a different quantity of information and continues getting better even in the presence of privacy concerns.

**How to implement it in cybersecurity via blockchain technology?**

Each participant trains a local model on its own data and shares only model updates, which are securely recorded and validated on the blockchain. This ensures privacy, transparency, and integrity of the learning process, preventing malicious tampering. Blockchain's decentralized nature also facilitates trust and incentivizes participation through rewards. By combining the privacy-preserving benefits of federated learning with blockchain's security and transparency, this approach enhances collaborative cybersecurity while safeguarding data.

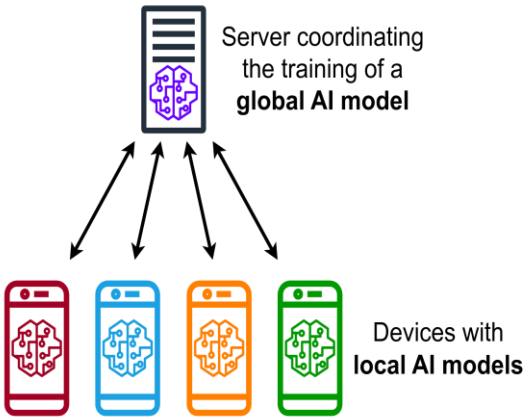


**Fig 4: The above picture indicates the flow of data update from local models to a single common global model**

**Deception technology**

Deception technology is a category of cybersecurity solutions that detect threats early with low rates of false positives. The technology deploys realistic decoys (e.g., domains, databases, directories, servers, apps, files, credentials, breadcrumbs) in a network alongside real assets to act as lures. The moment an attacker interacts with a decoy, the technology begins gathering intel that it uses to generate high-fidelity alerts that reduce dwell time and speed up incident response.

In simple words, Deception technology is a strategy to attract cyber criminals away from an enterprise's

true assets and divert them to a decoy or trap

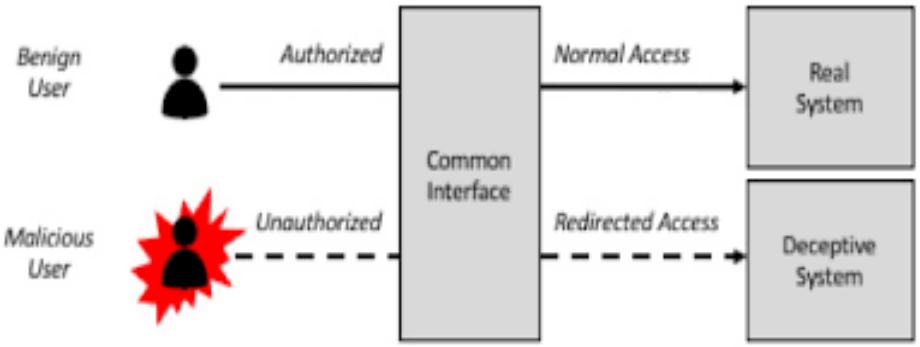


**Fig 5: Basic diagram of how deception technology is implemented.**

## VI. Security Matrix….What is security matrix?

Developing and validating a security matrix for blockchain involves identifying potential risks and vulnerabilities, defining controls to mitigate these risks, and establishing criteria to evaluate the effectiveness of the controls. The matrix can serve as a comprehensive guide to securing a blockchain network or application by addressing its unique characteristics.

Components that are in security matrix-

1. Consensus Mechanism: Proof-of-Work (PoW), Proof-of-Stake (PoS), etc., each having different attack vectors like 51% attacks or staking collusion.
2. Network Security-Ensuring the integrity of peer-to-peer communication and data transmission.
3. Data Privacy- Protecting sensitive data within transactions, using techniques like encryption or zero-knowledge proofs.
4. Smart Contracts Security- Validating smart contracts to prevent bugs, reentrancy attacks, and other vulnerabilities
5. Identity and Access Management (IAM)- Ensuring that only authorized entities interact with the blockchain or associated systems.
6. Governance- Addressing the decentralized nature of blockchain governance mechanisms and their impact on security.
7. Scalability- Managing security as the blockchain scales with more users and nodes.

## VII. Conclusion

This paper has presented an overview of the application of Blockchain technology in the cyber security domain,Moreover, it has also highlighted the cyberattack prevention method which is federated learning. Blockchain technology continues to evolve and find more use cases in the modern world. The Blockchain infrastructure makes it highly practical in addressing the existing security challenges in areas such as IoT devices, networks, and data in transmission and storage. The paper has evaluated the applicability of the Blockchain technology from the perspective of 30 researchers reviewed by Taylor et al. It has been observed that most Blockchain security researchers are concentrating a lot on the adoption of Blockchain security for IoT devices. Alongside this, other major areas of Blockchain security are networks and data

**References**

1. How Consensus Algorithms Solve Issues with Bitcoin's Proof of Work, http://www.coindesk.com/stellar-ripple-hyperledger-rivals-bitcoinproof-work/
2. What is Multi-Sig, and What Can It Do? https://coincenter.org/entry/what-is-multi-sig-and-what-can-it-