International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

HAB&ML Model to Improve the Security & QoS Performance of Industrial IoT

Neha. S. Patil[1], Zaware S. N[2], Shalaka. P. Deore[3], Shrihari Khatawkar[4,] G.J. Navale[5]

Submitted: 12/11/2022 **Accepted:** 14/02/2023

Abstract: Industrial internet of things (IIoT) equipments undergoes a wide variety of quality checks before being accepted for on-site usage. Because of recent advances in machine learning& cloud-based computations, IIoT devices have sufficiently high accuracy & precision performance. Due to which, IIoT network designers focus more towards improving security, privacy and scaliability. A wide variety of models are available for security which uses cryptographic, key-exchange and blockchain implementations. Due to their application specific nature, these models have low scalability, which limits their real-time usability. Hence to overcome this issue, current research proposes hybrid augmented blockchain & machine learning model which help to improve device scalability. This model utilizes side chains and machine learning models for improving QoS while maintaining high security in IIoT networks. It further proposes an interfacing method that is tested on multiple existing IIoT nodes, for enhancing their security & QoS performance. It is observed that the proposed model gives 18%better accuracy,11%better precision and 12%better AUC of attack detection when compared with NR2B, SMHGKA and DQNSBmodels**.** Thus, the proposed model is capable of high security, better QoS, and has superior scalability due to black-box approach for existing devices.

Keywords*: Block-chain, HAB&ML, IIoT, QoS, AUC, Precision, Delay*

1. Introduction

Industrial IoT devices are designed to achieve superior speed with high level of security, which includes data privacy, control signal security, node security, node authentication, etc. In order to design such devices, specialized algorithms for encryption, hashing, secret sharing, privacy preservation, key-exchange, data aggregation, tunneling, etc. are proposed by researchers. These algorithms work at one of the 4 levels of IIoT design, which are [1],

- Device level, wherein data security, privacy, aggregation, etc. algorithms are deployed. These algorithms mainly preserve node's authenticity, and optimize its performance.
- Network & connectivity level, which utilizes algorithms for routing, antenna optimization, etc.
- Infrastructure & platforms, wherein design of Gateways, fog nodes, and other interfacing components is optimized via machine learning, fog computing, and stream processing operations.
- Application level, which assists in monitoring of control devices, maintenance, visualization, etc.

The devices form the core components of any IIoT network require highly complex optimizations in order to achieve better security & QoS performance. In order to perform this task, various machine learning & data security algorithms are proposed by researchers [2, 3, 4]. Survey of these algorithms and their performance, advantages, disadvantages and future scope is discussed in section III. It has been observed from this literature study that, most models proposed are for optimization techniques for new IIoT deployments, which limits their usage capabilities for existing industrial components. Inspired from this study this research work proposed hybrid augmented blockchain &machine learning model(HAB&ML). It works on mutable sidechains and their maintenance mechanism using machine learning approach. The performance of proposed model is evaluated by comparing with existing machine learning methods.Finally, this research concludes with some remarkable observations about the proposed HAB&MLmodel.

2. Literature Review

Blockchain based system models are currently being evaluated for a wide variety of IoT applications. These include industrial IoT, home IoT, and commercial IoT networks. These networks have various security & privacy preservation restrictions, which must be catered by the blockchain model for better deployment capabilities. The work in [5, 6] propose such a framework, wherein researchers have used different types of blockchain models, and used them for adversary node resistant &

[1,2,5]*Department of Computer Engineering, AISSMS Institute of Information Technology,Savitribai Phule Pune University,Maharashtra,India*
[3]*M. E. S College of Engineering, Pune*
[4]*Annasaheb Dange College of Engineering & Technology ,Ashta*

halting recoverable blockchains (NR2B). These models have high good attack detection capabilities, but have low QoS performance due to highly complex attack detection processes. This drawback is removed by [7], wherein researchers have deployed secure sharding blockchain model using hierarchical group key agreement (SMHGKA), and applied it for large-scale networks. The model is capable of providing high QoS with good attack detection & removal capabilities, due to which it can be applied for any sized IoT networks. An application of this model on open blockchains can be observed from [8], wherein smart contracts are deployed for securing node-to-node transactions.

Scalable blockchains are proposed in [9, 10, 11], wherein researchers have discussed scalable distributed ledgers (SDLs), trust models for blockchain sharding, and mobile edge computing based secure self-organized blockchains (MECSSOB) which reduce computational complexity of IoT nodes by offloading mining operations to mobile edge devices. Further the performance of these stratigies are improved by using some deep learning methods for effective task scheduling. Such deep learning models are proposed in [12, 13, 14], where Deep Q Network Shard based Blockchain (DQNSB), blockchain-based software defined network (BSDN), and fault tolerant blockchain models are discussed. These models assist in improving fault tolerance performance of blockchains, while reducing redundant calculations during blockchain mining, and verification process. Due to which they are applicable for large-scale IIoT networks with millions of nodes. But overall throughput of these models is lower, and reduces further as numbers of attacks in the network are increased. In order to improve attack resilience of blockchains, work in [15, 16, 17] proposes extensible blockchain-based data provenance, decentralized privacy over IIoT nodes using deep federated learning on mobile edge devices. These models aim at generating security tokens, which reduce authentication & access control attacks in the system. Performance of these models must be further evaluated on a large variety of deployments, some of which are discussed in [18, 19, 20, 21], and include highly intrusive networks, 5G Vehicular Adhoc Networks (VANETs), search-based environments with encrypted cloud deployments, and medical data sharing applications with highly sensitive environments. Estimation of performance on these environments will assist researchers in deploying scalable and high-efficiency blockchain networks. Such networks are described in [22, 23, 24], where effective key-management, privacy preservation, models that ensure fairness & reliability across multiple applications are discussed. These applications include smart car parking, home automation, hierarchical network scenarios, etc. It is observed that sidechain-based models are highly effective in such scenarios, and thus must be used for large scale application deployments. Further extensions of sidechains to larger networks can be observed from [25, 26], wherein Deep Reinforcement Learning based Spatial Crowdsourcing models, and sidechaining using machine learning for smart city application are discussed. These models suggest that existing protocols for IoT security are highly complex and require large computational power in order to deploy privacy preservation, encryption, and intrusion detection mechanisms. Hence, sidechains were developed, but most of the sidechain models are highly complex, and thus cannot be applied to large-scale IIoT networks. To overcome this drawback, our research proposes a security & QoS improvement for industrial IoT devices using hybrid augmented blockchain & machine learning (HAB&ML) model in next section. This is followed by its parametric validation, and probable recommendations for further extending its performance.

3. Method

From the literature review, it can be observed that existing protocols for IoT security require complex computations in order to deploy privacy preservation, encryption, and intrusion detection mechanisms. Due to the complexity of these protocols, QoS performance of the underlying IoT deployment reduces, which reduces its deployment capabilities for large-scale networks. In order to avoid this drawback, blockchain based models were deployed, which assisted in providing security using transparency, immutability, traceability, and distributed computing capabilities. To add data into these chains, mining algorithms are deployed, which depend on chain length, and hash complexity. Delay for mining increases exponentially w.r.t. the chain length, and impacts network as number of blocks are added. In order to overcome this drawback, sidechains were invented, which assist in dividing the main blockchain into multiple smaller chains. While using sidechains, each chain is managed separately by a group of selected miner nodes. Thus, as number of sidechains increase, computational complexity of managing these chains also increases. This complexity can be managed via machine learning-based sidechains, which allow for intelligent chain division. Such a model that uses hybrid augmented blockchain withmachine learningfor sharding is discussed in this section. The proposed model is capable of selecting optimum sidechain length depending upon multiple parameters including, main blockchain length, delay for block addition, consensus delay, and attack prevention probability. Initially the current blockchain is given to a stochastic modelling block, which generated multiple sidechain configurations, and evaluates various sidechain parameters. These parameters are fused using a fitness evaluation block, which decides whether to ‘modify’ or ‘pass’ these solutions. Finally, a process checker block is used for evaluation of quality of

service (QoS), and security parameters, which assists in selecting a given sidechain configuration.

3.1. Stochastic modelling for generation of sidechain configurations

In order to select the most optimum sidechain lengths, a stochastic modelling is approach is used. This approach initially generates sidechains with varying lengths, and then evaluates each sidechain configuration using QoS & security parameters. To generate these sidechains, the following process is used,

- Initialize input parameters,
    - Number of configurations (N_c)
    - Number of iterations (N_i)
    - Machine learning rate (ML_r)
    - Mark each configuration as 'to be changed'
- For each iteration in 1 to N_i
    - For each configuration in 1 to N_c
        - If this configuration is marked as 'not to be changed', then go to the next configuration
        - Else, divide the blockchain into 'k' random sidechains, where 'k' is evaluated using equation 1,
- $k = random(1, L_B * ML_r)$ ... (1)

Where, L_B represents length of the original blockchain.

- Using this division, identify the sidechain with minimum length, and add a block into this sidechain.
- While adding a new block, evaluate delay for block addition using equation 2,

- $D_{new} = D_{read} * L_i + D_{write} + D_{hash} * L_i + D_{check} * (L_i - 1)$ ... (2)

Where, $D_{new}, D_{read}, D_{write}, D_{hash}, and\ D_{check}$ represents delays for adding a new block, reading the existing blockchain, writing a block to the current chain, hashing the block, and checking the block for uniqueness in the current chain respectively, while L represents length of the currently selected sidechain.

- Based on this delay, evaluate fitness of this chain, using equation 3,

- $F_i = \sum_{j=1}^{k,k\neq i} \frac{D_{new_j}}{k-1} + D_{new_i}$ ... (3)

o Evaluate fitness for all sidechain configurations, and then evaluate fitness threshold using equation 4,

- $f_{th} = \frac{\sum_{i=1}^{N_c} F_i}{N_c} * ML_r$ ... (4)

o Discard all solutions where fitness is more than threshold by marking them as 'to be changed', while mark other solutions as 'not to be changed'

- At the end of last iteration, generate a sidechain evaluation table (similar to table 1), which will be used for QoS & security evaluation of the selected configurations.

Sidechain Configuration	**Fitness**	**Block addition Delay**
S1 (100 blocks), S2 (50 blocks), S3 (175 blocks), S4 (15 blocks)	F1	D1

Table 1. Sidechain evaluation table generated via stochastic modelling

This table is given to the sidechain evaluation block, which estimates multiple parameters for each sidechain configuration, and assists in selection of the best sidechain structure for block processing.

3.2. Sidechain evaluation using multiple parameters

The selected sidechain configurations are given to a sidechain evaluation block, wherein its hashing delay, sidechain length, consensus evaluation delay, and attack detection probabilities are estimated. These metrics are evaluated using different mechanisms, which are stated as follows,

3.2.1. Evaluation of hashing delay& sidechain length

For each configuration in table 1, the sidechain length can be evaluated from 'sidechain configuration' column, where different sidechains and their respective lengths are captured. Delay of hashing is evaluated by adding a new block to the minimum length sidechain, and then finding its block addition delay using equation 2. This new block addition delay is used in equation 5 to evaluate new hashing delay as follows,

$$D_{hash} = \frac{1}{L_i}[D_{new} - (D_{read} * L_i + D_{write} + D_{check} * (L_i - 1))] \quad ... (5)$$

The values of sidechain length, and hash delays are stored and used in section 3.3 for sidechain quality evaluation,

3.2.2. Evaluation of consensus delay

A random set of 'k' miner nodes are selected, and the sidechains are stored on these nodes. Each of the nodes are

then inquired based on Proof-of-work (PoW) consensus model. After this inquiry, each of the miner nodes respond with their respective consensus decisions. These decisions are combined to evaluate consensus delay via equation 6 as follows,

$$D_{consensus} = \frac{\sum_{i=1}^{k} D_{check_i}}{k} \dots (6)$$

Where, D_{check_i} represents delay needed by miner node i to verify current block. This delay is given for sidechain quality evaluation in the next sub-section.

3.2.3. Evaluation of attack detection probability

Evaluation of attack detection probability examines security strength for the given sidechain configuration. This strength is estimated using the following process,

- A testbed of 'k' random attack nodes is generated, which perform the following attacks on each of the miner nodes,
    - Man in the middle attack
    - Masquerading attack
    - Spoofing attack
- After attacking these nodes, their individual blockchains are examined, and a metric for attack detection probability (ADP) is evaluated using equation 7,

$$ADP = \frac{1}{N_a * N_n} \sum_{i=1}^{N_a} \sum_{j=1}^{N_n} CV_j * A_{i,j} \dots (7)$$

Where, $N_a, N_n, CV, and\ A_{i,j}$ represents number of attacks injected, number of miner nodes, chain verification status, and presence of attack i on the chain present at node j

- The chain verification status is a flag with binary values, and indicates if the current chain is validated via hash checking or not.

3.3. Fitness evaluation & Process checking

Using the evaluated metrics from equation 5, 6, and 7, a sidechain rank is estimated for each sidechain configuration of table 1. This rank is estimated using equation 8 as follows,

$$SC_{rank} = \frac{var(S_{lengths})}{ADP} * \left[\frac{D_{hash}}{\max(D_{hash})} + \frac{D_{consenus}}{\max(D_{consensus})}\right] \dots (8)$$

All the sidechain configurations are ranked according to SC_{rank} in ascending order. Sidechains with maximum rank indicate configurations that have highest delays with lower attack detection probabilities. Similarly, sidechains with lower ranks indicates configurations with higher attack detection rate, and lower delays. Thus, this block selects the sidechain configuration with lowest rank, and uses it for future network operations. Due to use of these machine learning-basedsidechain configuration modelling processes, overall accuracy of attack detection improves, while delay needed for blockchain mining, and overall consensus reduces. This improves QoS performance of the chain, while making the selected sidechain highly secure under different types of attacks. Performance of this model is evaluated in terms of accuracy of attack detection, precision of attack removal,delay for blockchain mining, and area under the curve (AUC) for attack detection. This performance is compared with various state-of-the art models, and is tabulated in the next section of this text.

4. Results and Discussion

It has been observed from our experimentation that performance of hybrid augmented blockchain &machine learning (HABC & ML) model is better than previous models.This model is able to achieve high accuracy in detection of attack, good precision in removal of attack , reducetion in delay for blockchain mining and better area under the curve (AUC) performance. This performance is evaluated by comparing with NR2B [6], SMHGKA [7], &DQNSB [12]. This experimentation is performed on IEEE Data Port and TON IoT datasets, which consist of over 3 millionIoT data sensing and attack instances. Further these instances were splitted into 70:30 ratio for training &testing purpose.Accuracy in attack detection (Accuracy) is be observed as per table 2.

From table 2, it is observed that the proposed model has 21%,14%&18% better accuracy than NR2B [6],SMHGKA [7] and DQNSB [12] when compared on the same dataset. This is due to the fact that the proposed model uses blockchain, which improves probability of attack detection via incorporation of immutability and traceability, this helps in improving attack detection accuracy for small, medium, and large communication packets.

From table 3, it is observed that the proposed HAB&ML model has 20%,3%,9% better precision than NR2B [6], SMHGKA [7] and DQNSB [12], when compared on multiple number of requests. This performance improvement is due to use of sidechain with machine learning approach which assists in improving efficiency of attack detection via adaptively modifying sidechain lengths.

Similarly performance in terms of AUC of attack detection (AUC) is as depicted in table 4.It is observed from table 4 that the HAB&ML model has 20%,6%,10% better AUC than NR2B [6], SMHGKA [7], and DQNSB [12], when compared on different requests. This

performance improvement is observed due to implementation of machine learning with sharding, which assists in reducing overall complexity of attack detection.

From table 5, it is observed that the performance of HAB&ML model in terms of delay in block chain mining is 15% ,18% and 16 % lower than NR2B [6], SMHGKA [7], andDQNSB [12], when compared on a large number of communications.

Number of IIoTcommunications	Accuracy (%) NR2B [6]	Accuracy (%) SMHGKA [7]	Accuracy (%) DQNSB [12]	Accuracy(%) HAB&ML
500	66.26	67.92	75.25	84.60
1000	67.87	69.00	75.45	85.77
1500	68.18	70.16	75.65	86.45
2000	68.58	70.85	76.05	87.05
2500	68.88	71.57	76.36	87.58
5000	69.08	71.94	76.66	87.93
7500	69.90	73.09	76.96	88.85
10000	70.40	73.90	77.23	89.49
12500	70.90	74.72	77.53	90.15
15000	71.41	75.53	77.83	90.80
17500	71.91	76.35	78.12	91.44
20000	72.41	77.16	78.42	92.10
22500	72.91	77.97	78.72	92.75
25000	73.42	78.79	79.01	93.40
27500	73.91	79.60	79.31	94.05
30000	74.42	80.42	79.60	94.71
32500	74.92	81.22	79.90	95.35
35000	75.43	82.04	80.19	96.00
37500	75.92	82.86	80.49	96.66
40000	76.43	83.67	80.79	97.31
42500	76.93	84.49	81.08	97.96
45000	77.43	85.29	81.38	98.61
47500	77.93	86.11	81.67	99.26
50000	78.44	86.92	81.97	99.91

Table 2. Accuracy of event classification of different models

Performance in terms of Precision in attack detection (P) is as shown in table 3 .

Number ofIIoTcommunications	Precision (%) NR2B [6]	Precision (%) SMHGKA [7]	Precision (%) DQNSB [12]	Precision (%) HAB&ML
500	62.94	71.32	67.72	75.55
1000	64.48	72.46	67.90	76.61
1500	64.77	73.67	68.08	77.24
2000	65.16	74.40	68.45	77.80
2500	65.44	75.14	68.72	78.29
5000	65.63	75.54	68.99	78.61
7500	66.41	76.75	69.27	79.45
10000	66.88	77.60	69.51	80.04
12500	67.36	78.46	69.78	80.64
15000	67.83	79.31	70.04	81.23
17500	68.32	80.16	70.32	81.83
20000	68.79	81.02	70.58	82.44
22500	69.27	81.87	70.84	83.03
25000	69.74	82.73	71.11	83.63
27500	70.22	83.58	71.38	84.22
30000	70.70	84.44	83.59	89.29
32500	71.17	85.29	83.89	89.90
35000	71.65	86.14	84.20	90.51
37500	72.12	87.00	84.52	91.13
40000	72.61	87.85	84.83	91.75
42500	73.08	88.71	85.13	92.36
45000	73.56	89.57	85.45	92.97
47500	74.03	90.42	85.76	93.59
50000	74.52	91.27	86.07	94.20

Table 3. Precision of attack detection for different models

Number of IIoTcommunications	AUC (%) NR2B [6]	AUC (%) SMHGKA [7]	AUC (%) DQNSB [12]	AUC (%) HAB&ML
500	66.26	71.41	73.32	82.13
1000	67.87	72.54	73.51	83.27
1500	68.18	73.76	73.71	83.95
2000	68.58	74.49	74.10	84.54
2500	68.88	75.23	74.40	85.06

5000	69.08	75.63	74.69	85.40
7500	69.90	76.84	74.99	86.31
10000	70.40	77.69	75.26	86.94
12500	70.90	78.55	75.55	87.58
15000	71.41	79.41	75.83	88.22
17500	71.91	80.26	76.12	88.86
20000	72.41	81.11	76.41	89.51
22500	72.91	81.97	76.70	90.15
25000	73.42	82.83	76.98	90.78
27500	73.91	83.68	77.28	91.42
30000	74.42	84.54	83.69	94.35
32500	74.92	85.40	83.99	95.00
35000	75.43	86.25	84.30	95.66
37500	75.92	87.10	84.62	96.30
40000	76.43	87.96	84.93	96.95
42500	76.93	88.82	85.24	97.59
45000	77.43	89.67	85.55	98.25
47500	77.93	90.53	85.86	98.90
50000	78.44	91.38	86.17	99.54

Table 4. AUC of attack detection for different models The observations related to delay of blockchain mining (Delay) is depicted in table 5.

Number of IIoTcommunications	**Delay (s) NR2B [6]**	**Delay (s) SMHGKA [7]**	**Delay (s) DQNSB [12]**	**Delay (s) HAB&ML**
500	0.19	0.21	0.21	0.16
1000	0.40	0.42	0.43	0.33
1500	0.61	0.66	0.66	0.50
2000	0.81	0.88	0.88	0.66
2500	1.02	1.11	1.10	0.84
5000	2.04	2.23	2.20	1.68
7500	3.09	3.40	3.32	2.55
10000	4.15	4.59	4.44	3.42
12500	5.23	5.80	5.58	4.30
15000	6.32	7.03	6.71	5.20
17500	7.42	8.28	7.86	6.12
20000	8.54	9.57	9.02	7.04
22500	9.68	10.88	10.18	7.98
25000	10.83	12.22	11.35	8.93
27500	11.99	13.57	12.53	9.89
30000	13.17	14.97	14.82	11.14
32500	14.36	16.37	16.11	12.15
35000	15.57	17.81	17.41	13.17
37500	16.80	19.27	18.73	14.21
40000	18.04	20.76	20.04	15.25
42500	19.29	22.27	21.37	16.32
45000	20.55	23.81	22.71	17.39
47500	21.84	25.37	24.06	18.47
50000	23.14	26.96	25.42	19.57

Table5. Delay of blockchain mining using different number of communications

Thus, this comparative study of HAB&ML model with NR2B [6], SMHGKA [7], and DQNSB [12] proves the capability of HAB&ML model to support high-speed, high-accuracy, and high AUC performance on multiple request types efficiently.

5. Conclusion

Due to incorporation of machine learning for sidechain configuration selection, the proposed model is capable of reducing blockchain mining delay, while showcasing good precision, accuracy and AUC performance for attack detection. The proposed model is able to provide21%,14%,18% better accuracy than NR2B [6], SMHGKA [7], and DQNSB [12], when compared on the same dataset. While it has 15% lower delay than NR2B [6], 18% than SMHGKA [7], and 16% than DQNSB [12], when compared on a large number of communications. Due to such a high performance in terms of both delay of mining, and accuracy of attack detection, the proposed model is useful for a wide variety of highly useful IIoT applications. The model further showcases 20% better precision than NR2B [6], 3% better precision than SMHGKA [7], and 9% better precision than DQNSB [12], when compared on multiple number of requests. In future, researchers can use Q-learning, reinforcement learning, and other deep learning models for improving computational efficiency of sidechain creation. Researchers can also explore utility of different consensus models including Proof-of-Stake (PoS), Proof-of-Authority (PoA), etc., then, apply it for selection of sidechain configuration, and observe its effect on QoS & security performance.

Author contributions

N.S.Patil: Conceptualization, Methodology, Software, Field study **S.N.Zaware:**Methodology,Data curation, Writing-Original draft preparation, Software, Validation., Field study **S.P.Deore:** Visualization, Investigation **S.Khatavkar**:Writing-Reviewing and Editing **G.J.Navale:**Writing-Reviewing and Editing

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] M. Li, H. Tang, A. R. Hussein and X. Wang, "A Sidechain-Based Decentralized Authentication Scheme via Optimized Two-Way Peg Protocol for Smart Community," in IEEE Open Journal of the Communications Society, vol. 1, pp. 282-292, 2020, doi: 10.1109/OJCOMS.2020.2972742.

[2] S. Sun, R. Du, S. Chen and W. Li, "Blockchain-Based IoT Access Control System: Towards Security, Lightweight, and Cross-Domain," in IEEE Access, vol. 9, pp. 36868-36878, 2021, doi: 10.1109/ACCESS.2021.3059863.

[3] V. Agarwal and S. Pal, "Blockchain meets IoT: A Scalable Architecture for Security and Maintenance," 2020 IEEE 17th International Conference on Mobile Ad Hoc and Sensor Systems (MASS), 2020, pp. 53-61, doi: 10.1109/MASS50613.2020.00017.

[4] Y. Miao, M. Zhou and A. Ghoneim, "Blockchain and AI-Based Natural Gas Industrial IoT System: Architecture and Design Issues," in IEEE Network, vol. 34, no. 5, pp. 84-90, September/October 2020, doi: 10.1109/MNET.011.1900532.

[5] Firoozjaei, MD, Lu, R, Ghorbani, AA. An evaluation framework for privacy-preserving solutions applicable for blockchain-based internet-of-things platforms. *Security and Privacy*. 2020; 3:e131. https://doi.org/10.1002/spy2.131

[6] Xu, Yibin& Huang, Yangyu. (2020). An n/2 byzantine node tolerate blockchain sharding approach. 349-352. 10.1145/3341105.3374069.

[7] Naresh, VS, Allavarpu, VVLD, Reddi, S, Murty, PSR, Raju, NVSL, Mohan, RNVJ. A provably secure sharding based blockchain smart contract centric hierarchical group key agreement for large wireless ad-hoc networks. *Concurrency ComputatPractExper*. 2021;e6553. https://doi.org/10.1002/cpe.6553

[8] Y. Tao, B. Li, J. Jiang, H. C. Ng, C. Wang and B. Li, "On Sharding Open Blockchains with Smart Contracts," 2020 IEEE 36th International Conference on Data Engineering (ICDE), 2020, pp. 1357-1368, doi: 10.1109/ICDE48307.2020.00121.

[9] S. R. Niya, R. Beckmann and B. Stiller, "DLIT: A Scalable Distributed Ledger for IoT Data," 2020 Second International Conference on Blockchain Computing and Applications (BCCA), 2020, pp. 100-107, doi: 10.1109/BCCA50787.2020.9274456.

[10] M. N. Halgamuge, S. C. Hettikankanamge and A. Mohammad, "Trust Model to Minimize the Influence of Malicious Attacks in Sharding Based Blockchain Networks," 2020 IEEE Third International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), 2020, pp. 162-167, doi: 10.1109/AIKE48582.2020.00032.

[11] A. Asheralieva and D. Niyato, "Reputation-Based Coalition Formation for Secure Self-Organized and Scalable Sharding in IoT Blockchains With Mobile-Edge Computing," in IEEE Internet of Things Journal, vol. 7, no. 12, pp. 11830-11850, Dec. 2020, doi: 10.1109/JIOT.2020.3002969.

[12] J. Yun, Y. Goh and J. -M. Chung, "DQN-Based Optimization Framework for Secure Sharded Blockchain Systems," in IEEE Internet of Things Journal, vol. 8, no. 2, pp. 708-722, 15 Jan.15, 2021, doi: 10.1109/JIOT.2020.3006896.

[13] L. Liu et al., "BS-IoT: Blockchain Based Software Defined Network Framework for Internet of Things," IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2020, pp. 496-501, doi: 10.1109/INFOCOMWKSHPS50562.2020.9163070.

[14] C. Mao, A. -D. Nguyen and W. Golab, "Performance and Fault Tolerance Trade-offs in Sharded Permissioned Blockchains," 2020 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), 2020, pp. 1-3, doi: 10.1109/ICBC48266.2020.9169425.

[15] Sigwart, M., Borkowski, M., Peise, M. *et al.* A secure and extensible blockchain-based data provenance framework for the Internet of Things. *Pers UbiquitComput* (2020). https://doi.org/10.1007/s00779-020-01417-z

[16] Y. Qu et al., "Decentralized Privacy Using Blockchain-Enabled Federated Learning in Fog Computing," in IEEE Internet of Things Journal, vol. 7, no. 6, pp. 5171-5183, June 2020, doi: 10.1109/JIOT.2020.2977383.

[17] Y. Zhao et al., "Privacy-Preserving Blockchain-Based Federated Learning for IoT Devices," in IEEE Internet of Things Journal, vol. 8, no. 3, pp. 1817-

1829, 1 Feb.1, 2021, doi: 10.1109/JIOT.2020.3017377.

[18] O. Alkadi, N. Moustafa and B. Turnbull, "A Review of Intrusion Detection and Blockchain Applications in the Cloud: Approaches, Challenges and Solutions," in IEEE Access, vol. 8, pp. 104893-104917, 2020, doi: 10.1109/ACCESS.2020.2999715.

[19] L. Xie, Y. Ding, H. Yang and X. Wang, "Blockchain-Based Secure and Trustworthy Internet of Things in SDN-Enabled 5G-VANETs," in IEEE Access, vol. 7, pp. 56656-56666, 2019, doi: 10.1109/ACCESS.2019.2913682.

[20] Y. Yang, H. Lin, X. Liu, W. Guo, X. Zheng and Z. Liu, "Blockchain-Based Verifiable Multi-Keyword Ranked Search on Encrypted Cloud With Fair Payment," in IEEE Access, vol. 7, pp. 140818-140832, 2019, doi: 10.1109/ACCESS.2019.2943356.

[21] H. Jin, Y. Luo, P. Li and J. Mathew, "A Review of Secure and Privacy-Preserving Medical Data Sharing," in IEEE Access, vol. 7, pp. 61656-61669, 2019, doi: 10.1109/ACCESS.2019.2916503.

[22] M. Ma, G. Shi and F. Li, "Privacy-Oriented Blockchain-Based Distributed Key Management Architecture for Hierarchical Access Control in the IoT Scenario," in IEEE Access, vol. 7, pp. 34045-34059, 2019, doi: 10.1109/ACCESS.2019.2904042.

[23] C. Zhang et al., "BSFP: Blockchain-Enabled Smart Parking With Fairness, Reliability and Privacy Protection," in IEEE Transactions on Vehicular Technology, vol. 69, no. 6, pp. 6578-6591, June 2020, doi: 10.1109/TVT.2020.2984621.

[24] A. Saini, Q. Zhu, N. Singh, Y. Xiang, L. Gao and Y. Zhang, "A Smart-Contract-Based Access Control Framework for Cloud Smart Healthcare System," in IEEE Internet of Things Journal, vol. 8, no. 7, pp. 5914-5925, 1 April1, 2021, doi: 10.1109/JIOT.2020.3032997.

[25] H. Lin, S. Garg, J. Hu, G. Kaddoum, M. Peng and M. S. Hossain, "Blockchain and Deep Reinforcement Learning Empowered Spatial Crowdsourcing in Software-Defined Internet of Vehicles," in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 6, pp. 3755-3764, June 2021, doi: 10.1109/TITS.2020.3025247.

[26] P. Kumar et al., "PPSF: A Privacy-Preserving and Secure Framework Using Blockchain-Based Machine-Learning for IoT-Driven Smart Cities," in IEEE Transactions on Network Science and Engineering, vol. 8, no. 3, pp. 2326-2341, 1 July-Sept. 2021, doi: 10.1109/TNSE.2021.3089435.