Machine learning aided blockchain assisted framework for wireless networks

Khan, A. S., Zhang, X., Lambotharan, S., Zheng, G., AsSadhan, B. & Hanzo, L.

Author post-print (accepted) deposited by Coventry University's Repository

Original citation & hyperlink:
Khan, AS, Zhang, X, Lambotharan, S, Zheng, G, AsSadhan, B & Hanzo, L 2020, 'Machine learning aided blockchain assisted framework for wireless networks', IEEE Network vol. (In-press), pp. (In-press).
https://dx.doi.org/10.1109/MNET.011.1900643

DOI 10.1109/MNET.011.1900643
ISSN 0890-8044

Publisher: IEEE

© 2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

This document is the author's post-print version, incorporating any revisions agreed during the peer-review process. Some differences between the published version and this version may remain and you are advised to consult the published version if you wish to cite from it.

Machine Learning Aided Blockchain Assisted Framework for Wireless Networks

Amjad Saeed Khan, *Member, IEEE*, Xinruo Zhang, *Member, IEEE*,
Sangarapillai Lambotharan, *Senior Member, IEEE*, Gan Zheng, *Senior Member, IEEE*,
Basil AsSadhan, *Member, IEEE*, Lajos Hanzo, *Fellow, IEEE*

***Abstract*—Inspired by its success in financial sectors, the blockchain technique is emerging as an enabling technology for secure distributed control and management of wireless networks. In order to fully benefit from this distributed ledger technology, its limitations, cost, complexity and empowerment also have to be critically appraised. Depending on the specific context of the problem to be solved, these limitations have been handled to some extent through a clear dichotomy in the blockchain architectures, namely by conceiving both permissioned and permissionless blockchains. Permissionless blockchain requires massive computing power to achieve consensus, while its permissioned counterpart is energy efficient but would require trusted participants. To combine these benefits by gaining trust at a high energy efficiency, a novel mechanism is proposed for automatically learning the trust level of users in a public blockchain network and grant them access to a private blockchain network. In this context, machine learning is a very powerful tool capable of automatically learning the trust level. We have proposed reinforcement learning for bridging the dichotomy of blockchains in terms of striking a trust vs complexity trade-off in an unknown environment. Benefits and limitations of various forms of blockchain techniques are analyzed, followed by its reinforcement-aided evolution. We demonstrate that the proposed reinforcement learning aided blockchain is capable of supporting high-integrity autonomous operation and decision making in wireless networks. The win-win amalgamation of these techniques has been demonstrated for striking a compelling balance between the benefits of permissioned and permissionless blockchain networks through the case-study of the proposed blockchain based unmanned aerial vehicle aided wireless networks.**

I. Introduction

Blockchains have the potential of increasing the resilience, security, reliability, transparency and trust in a wide range of applications including autonomous systems, the Internet of things (IoT) and other communication networks relying on distributed operation and resource allocation [1]. As a benefit of its distributed data storage and security, it has also been proposed for distributed cloud storage and edge computing systems for maintaining ultimate security against malicious attacks. Another fascinating application of blockchain may be found in improving the data security and privacy of software defined networking [2]. Furthermore, the blockchain which is resilient against any unauthorized modification has been proposed for the monitoring and management of IoT devices [3] as well as for energy trading and distribution in addition to improving the security and reliability of smart grid communications [4].

The motivation of relying on blockchain in wireless networks lies in the desire to automate the operations of wireless networks supported by the distributed management of resources. This is particularly important due to the proliferation of heterogeneous IoT devices and the distributed management of tasks, which are normally performed at the edge of the networks. Hence, there is a strong inspiration both for the autonomous execution of smart contracts, as well as for localized identification of malicious activities and for the establishment of trustworthiness of the users/devices [5]. Machine learning (ML) is a powerful technique of learning and improving the trust level among various parties by eliminating the need for human judgment. The family of ML techniques including neural networks, support vector machines, decision trees, naive Bayesian and reinforcement learning (RL) has been widely considered for addressing various challenges in wireless networks [6]. Specifically, ML techniques have been successfully applied for monitoring and performance management, resource allocations, network security, dynamic network access, routing and network connectivity preservation. There has been a growing interest in the convergence of blockchain and machine learning technologies in wireless networks [7]–[9]. The marriage of ML and blockchain technologies has the potential of revolutionizing wireless networking. For example, a machine learning based software-defined blockchain architecture has been proposed in [7] for intelligent management of consensuses in the IoT. The efficiency of blockchains and reinforcement learning for the development of distributed caching schemes has been demonstrated in [8]. In this paper, we critically analyze suitability of permissioned and permissionless blockchains for wireless networks. Particularly, following a rudimentary exposure of the fundamentals of both blockchain solutions and reinforcement learning, we propose a framework to exploit their intrinsic amalgam to strike a balance between the benefits of the permissioned and permissionless blockchain networks. The application of the proposed framework is discussed in the context of a UAV communications case study.

II. High Level Description of Blockchain and Reinforcement Learning

Blockchain is a secure distributed ledger technology that facilitates transparent, verifiable, secure and unalterable digital asset transactions with proof of rights and ownership. Based on cryptographic techniques and consensus mechanisms, blockchain employs peer-to-peer computing nodes and a decentralized database for conducting transactions over a trusted layer [10]. The blockchain employs digital signatures

and consensus mechanisms for ascribing digital ownership to its entities, thereby providing them with reliable protection against malicious attacks. Additionally, as a benefit of having multiple cryptographically valid ledger copies across a network, blockchain eliminates the curse of a single point of failure or hacking attacks or control by a single element, hence providing superior security compared to conventional centralized mechanisms. The smart contract features of blockchain have enabled autonomous distributed applications on networks without any fraud, control or interference from a third party. Additionally, smart contracts are capable of automating agreements and provisions between access nodes, networks and subscribers, and have the ability to authorize devices to directly negotiate the acquisition of the best possible service.

A. Consensus Mechanisms and Their Limitations

Popular cryptocurrencies such as Bitcoin and Ethereum rely on the so-called *Proof of Work* (PoW) mechanism for combining transactions into Merkle tree-based blocks, where participants termed as miners compete for finding a random instant that produces a hash digest of a predefined range of complexity [11]. The first miner who won the right to create a block earns a block reward. Once the new block is broadcast into the network, each miner verifies the solution and appends the block to its blockchain. However, the PoW mechanism has many limitations in terms of its scalability, latency and energy consumption, in addition to the to overwhelming 51%-attack or majority hash rate attack, as seen in recent events with Bitcoin, Peercoin, Litecoin, Namecoin, etc [12]. In order to address these issues, the new generation of blockchains - including Ethereum and Cardano - relies on the more powerful *Proof of Stake* (PoS) mechanism [11], in which the participants' chances of successful mining action (block creation) increase with the number of coins they have. Compared to the PoW mechanism, PoS is more energy-efficient and less prone to attacks of the networking miners. It can support high transaction throughput, hence, the blockchain is capable of validating thousands or even hundreds of thousands of transactions per second using the PoS. Nevertheless, the PoS is not a perfect protocol either. For example, the richer only gets richer, and in the long run the network can converge to a highly centralized state. To circumvent this impediment, several other consensus mechanisms have been designed, as presented in [11]. However, to the best of our knowledge, little to no consensus mechanisms have been designed for exploiting the joint benefits of both the PoW and PoS protocols. In this context, we propose a machine learning aided consensus mechanism for amalgamating the benefits of both protocols.

B. Permissioned vs permissionless Blockchains

In contrast to the permissionless blockchains that are completely open for everyone to participate in, which are also known as public blockchains, permissioned blockchains have an access control layer for allowing a limited number of participants to read, write and access information on them. The most well-known permissioned blockchains are Hyperledger fabric, Quorum, Ripple, and R3 Corda [12]. They allow the network to appoint a group of participants to partake in the consensus mechanism or to validate blocks of transactions. Permissioned blockchains are typically operated by an organization or consortium that only allows trusted participants to join their enterprise blockchain. This makes the permissioned blockchains as less transparent and more centralized in nature. Given the fact that permissioned blockchains are designed for solving very specific business-oriented problems and for maintaining only a limited number of participating nodes, they are much faster than permissionless blockchains. However, the presence of trust among the participants is one of the pivotal factors in the efficient operation of blockchains, hence they are more vulnerable to malicious activities than their public counterparts. On the other hand, permissionless blockchains rely on a large number of participants/miners contributing to the transaction verifications and mining processes. They are transparent and resistant to malicious activities, but they are very expensive and slower.

Hence, there is a need to make blockchain technology not only transparent, secure and decentralized like permissionless blockchains, but also faster and inexpensive, like permissioned blockchains. In this paper, we propose a reinforcement learning framework that allows us to acquire the benefits of both the permisioned and permissionless blockchains. The marriage of reinforcement learning and blockchain is able to accelerate and simplify the process of selecting reliable participant(s) from a wider pool of volunteering users to form a permissioned blockchain network.

C. Overview of Reinforcement Learning

Reinforcement learning (RL) [13] has a huge potential in optimizing sequential decision making in real-time communications under unknown network environments. It has been widely considered in wireless networks for spectrum access, data rate selection and transmit power control problems, particularly for throughput maximization and energy consumption minimization. Typically, RL can be viewed as an agent vs. environment interaction mechanism, which is conventionally modeled by a Markov decision process, as shown in Fig. 1. The agent interacts with the environment constituted for example by a set of states, which the agent attempts to influence through its choice of instantaneous actions. In the environment, the agent receives an observation that typically includes a specific reward. The agent then embarks on a particular action in the environment, which is chosen from the set of legitimate actions. As a result, the environment evolves to a new state, yielding a particular reward for each transition, which is then fed back to the agent. This enables RL to find optimal policies for resource allocation problems, including channel allocation, network routing and admission control in modern wireless networks which are more decentralized, ad-hoc and autonomous in nature.

D. Blockchain for Trust and Security in Wireless Networks

Security and privacy are vital for wireless networks. Given the proliferation of heterogenous devices and applications that generate a high volume of diverse data, supporting

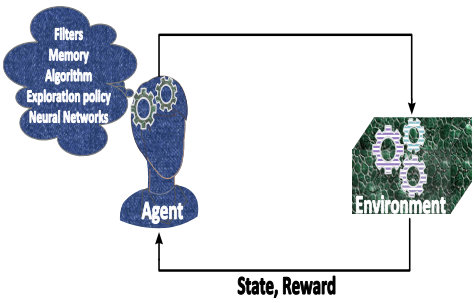

Fig. 1. Fundamental principle of reinforcement learning

centralized access control is a challenging task. Innovative blockchain solutions come to rescue for enhancing trust in next-generation wireless networks (NGNs) [14]. There are several promising blockchain solutions for establishing trust and security, including those conceived for edge computing, spectrum sharing, content caching, network virtualization and interference management [2], [14]. Moreover, the blockchain technology is attractive for IoT networks, since it can facilitate their decentralized operation for supporting scalability and for avoiding a single point of failure. This will also maintain anonymity and security, while supporting diverse applications and a massive number of connections. However, most of the IoT devices are resource-limited and will require low-latency communications. Therefore, the computationally intensive and hence time-consuming mining tasks will severely hinder the applications of blockchains in wireless networks, particularly in IoT networks [3]. To circumvent the above impediments, we will demonstrate that the marriage of blockchain and machine learning has the potential of maintaining the trust and security offered by blockchain at a modest complexity, whilst relying on both decentralized operation and distributed resource sharing [8]. The content caching problem of a secure permissioned blockchain-enabled framework was solved in [15] and a deep RL technique was used for maximizing system utility in terms of caching-resource-sharing. Similarly, a blockchain-enabled architecture was proposed in [9] for secure data sharing among vehicles and RL was used for mitigating the transmission overhead.

The above-mentioned treatises on blockchain and ML combined the benefits of decentralized operation, with the high-level security and trust offered by the blockchain with the aid of network performance optimization relying on ML. In contrast to these methods, our aim is to conceive bespoke ML for enhancing the operation of blockchain itself in support of NGNs. Specifically, we introduce a novel blockchain architecture for intrinsically amalgamating blockchains with an improved miner selection process by applying ML for automatically learning the trust level of users in a public blockchain network and then grant them access to a private blockchain network. Hence, we can combine the benefits of enhanced trust and security offered by the permissionless network with the low complexity offered by the permissioned blockchain. In a nutshell, this is a compelling solution for NGNs.

E. Reinforcement Learning for Blockchains: A Proposal

Reinforcement learning is capable of intrinsically amalgamating the benefits of permissioned and permissionless blockchains. For example, within the framework of permissioned blockchain, instead of allowing participants to contribute for the entire duration of blockchain operation during which the trust level of the participants might change, RL can be embedded as a consensus building mechanism for the selection of participants in the blockchain from a wider pool of volunteering users. This may be achieved by critically appraising all the selected participants and then exploiting the knowledge gained during the appraisal of the nodes in terms of their trustworthiness to contribute to the blockchain. This way, the RL aided blockchain quantifies the reputation of participants, whilst optimizing itself against malicious elements. As a benefit, the membership of malicious participants can be terminated autonomously. It also opens up new opportunities for a large number of users to earn profits as a benefit of their contribution in the blockchain. Since the proposed RL aided blockchain framework interacts with the publicly available volunteers willing to partake in the network, while admitting only a limited number of participants as shown in Fig. 2, it can be treated as a semi-private blockchain.

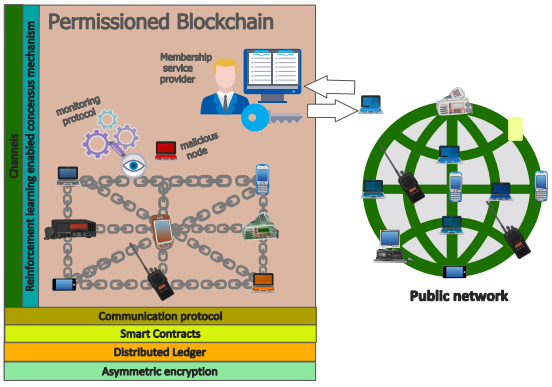

Fig. 2. Proposed framework of reinforcement learning enabled blockchain.

III. Blockchain for Communications: A Case Study

Let us now consider a case study employing the popular multi-arm bandit algorithm for the development of permissioned blockchain. The proposed blockchain is employed as a distributed monitoring framework for service level agreement (SLA) enforcement between a pair of trading parties, including a business agent (customer) and an unmanned aerial vehicle (UAV) (service provider), as described below.

A. System model and proposed framework

Consider a group of IoT devices located in a remote geographical area distributed randomly according to a binomial point process in a circular area, where it is almost impossible for classic terrestrial BSs on the ground to provide network coverage. In this context, a business agent who could be a third party agent is interested to have network coverage of a certain quality (e.g., data rate) in the entire area. Hence,

a UAV is employed for achieving the desired service quality in return for a payment, as shown in Fig. 3. After successful negotiation between the business agent and the UAV operator, an SLA is signed between the two parties. The SLA specifies the terms and conditions for both parties, including: the agreed quality of service (QoS) level, service duration, amount of payment to be remitted by the business agent if the agreement has been fulfilled, the terms under which the UAV is agreed to be penalized, etc.

During the service time, it is important for the UAV to maintain the target rate for all possible channel conditions. Without an appropriate mechanism for SLA monitoring, the case of SLA violations may occur. For example, a selfish UAV operator may misbehave for increasing its revenue by dropping the bit rate either by transmitting signals at an inadequate power or by allocating insufficient radio frequency bandwidth to IoT users, hence failing to achieve the desired data rate. Thus, in order to address this issue, a distributed monitoring framework is conceived for supporting the ground users that can record all the transactions generated in the wireless network, and by employing private blockchain techniques they provide autonomous SLA monitoring. The transactions may contain received signal strength and/or bit rate measurements, feedback alerts, and so on. The proposed framework is constituted by the set of volunteering ground users (e.g., IoT devices) that have good local area connectivity among themselves and that are willing to offer their services for SLA monitoring in return for a monetary reward.

The overarching aim is to form a blockchain network that will monitor the SLA of the UAV. Accordingly, consider a large set of volunteering ground users (e.g. N=1000). It is possible that these individual ground users are not entirely reliable: each of them is associated with an unknown trust-level and has a certain probability of being faulty or malicious throughout the blockchain operation. Hence, we aim for choosing a smaller subset of these ground users who would act as members of a private blockchain network. In order to avoid the malicious behaviour of the users participating in the private network, we use RL to only allow trustworthy participants to join, whilst removing the malicious participants from the private blockchain network. Among the private blockchains, we use the hyperledger fabric of [16] for the development of our distributed monitoring framework, comprised of three types of participants: monitors, endorsers and orderer, as shown in Fig. 3. As for the distributed automation, the framework contains various smart contracts, namely: SLA contract, monitoring, endorsing, alert, enroll, reinforcement and revoke contract. Monitors are the subset of users within the private blockchain who monitor the data rate offered by the UAV through invoking a monitoring contract. In our work, we use the signal power level received at the monitoring terminal for quantifying the data rate. Similarly, the endorsers are a subset of the blockchain network users excluding the monitors, who run an endorsing contract for analyzing the measured data and for making a verdict about the QoS. Additionally, endorsers contain digital ledgers for blockchain storage. The orderer is a member of the blockchain network, who is however not part of the monitors or endorsers, and is responsible for generating a block of transactions.

As shown in the Fig. 3, once the service starts, the UAV submits a transaction (e.g., transmit power level) for notifying the quality of its service delivery during the contract. The monitors use peer-to-peer communication for forwarding their received power level measurements to endorsers as transactions. After validating the transactions received from the monitors, each endorser invokes the endorsing contract that employs an SVM that has been trained using a trusted UAV for processing the received measurements (e.g. received power levels) for evaluating the data rate to check whether the UAV is indeed providing the agreed data rate. After evaluating the QoS, the endorsers share their verdict with the orderer. The orderer may choose the Practical Byzantine Fault Tolerance consensus mechanism [11] for making a final decision about the QoS experienced, and then generates a block of transactions (in the form of a service log, including the final verdict about the UAV-aided QoS). This final verdict is then broadcast back to all the endorsers for first validating all the transactions included in the block, followed by the orderer's decision about the QoS. Finally the endorsers update their ledger. In case of inadequate service delivery, an alert contract is invoked for instructing the UAV to improve its service quality either by raising its transmission power or by allocating a wider bandwidth for example.

B. RL-Empowered Users Selection Protocol

This protocol works both as an access mechanism encouraging ground users to contribute to the private blockchain as semi-trustworthy participants as well as a consensus mechanism for estimating the unknown trust-levels of the participants. We consider N volunteering ground users and K of them are used for establishing the blockchain network. The protocol is capable of selecting the most trustworthy participant as an orderer. This problem is formulated as the maximization of the long-term average trust-level of the private blockchain network. Since the role of the orderer is of salient importance in the blockchain's operation, we will evaluate the trustworthiness of participants to operate as an orderer. However, our solution may be readily extended to establishing the trust-level of both the monitors and the endorsers.

Since each individual ground user is deemed to behave either maliciously or honestly according to its unknown trust-level, our problem can be modeled as Bernoulli bandit problem [13]. Briefly, the Bernoulli bandit problem is a classic single-state RL problem that models a system of K actions, and when played, each action yields either success with an unknown but stationary probability of θ_u or failure with a probability of $1 - \theta_u$ (binary reward). An agent makes sequential decisions as to which specific actions to choose with the objective of maximizing the cumulative number of successes [13]. In our case-study, the K participants of our private blockchain network may be viewed as representing K actions, while the success probability θ_u corresponds to the unknown trust-level of the participant u. Note that the instantaneous reward generated by opting for a specific action is one if the selected orderer is trustworthy and zero otherwise.

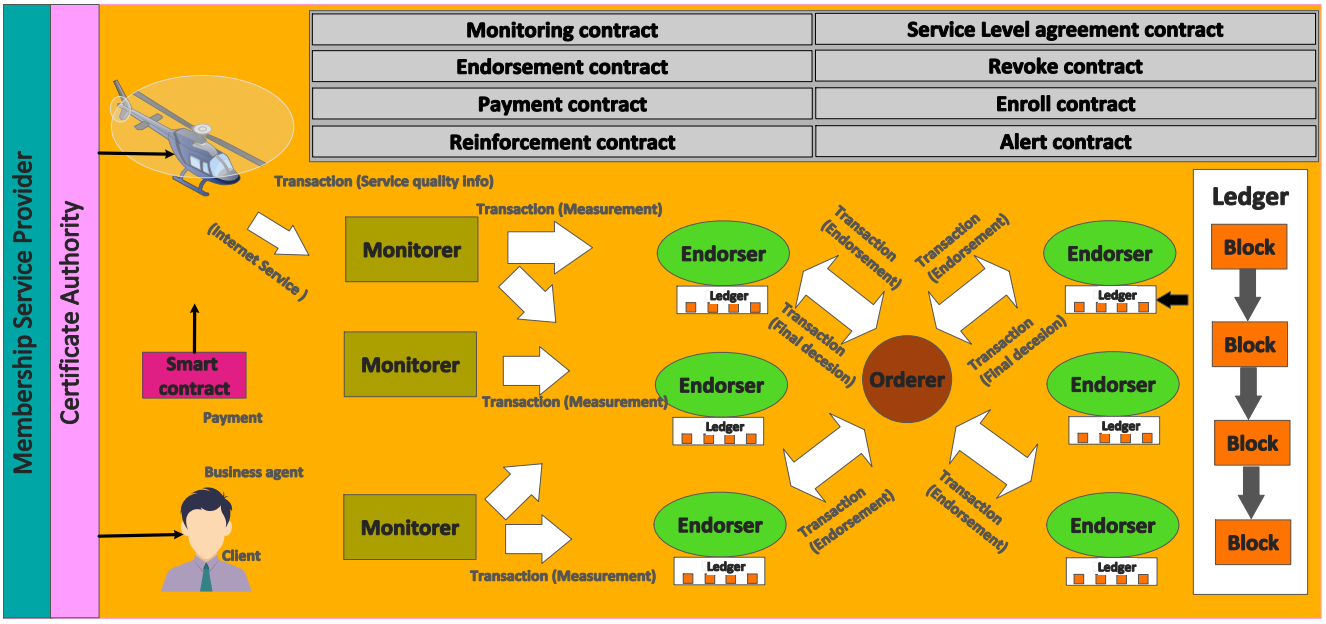

Fig. 3. Transactions flow diagram of the proposed blockchain infrastructure.

The main objective of the proposed technique is to maximize the accumulated probabilities of selecting the trustworthy orderer over time. Therefore, the protocol exploits enroll, revoke and reinforcement contracts iteratively. For instance as shown in Fig. 4, a reinforcement contract is used for appointing the participants as an orderer, as monitors as well as endorsers, and for estimating their unknown trust-levels with the aid of the Thompson Sampling method. This contract may also call upon the enroll and revoke contract for admitting and terminating the membership of participants respectively, depending on their trust-levels.

The fundamental principle of Thompson Sampling is that it processes independent *a priori* beliefs of each θ_u and updates its *a posteriori* distribution, as additional observations are gathered. More specifically, each *a priori* belief is assumed to be beta-distributed with parameters α_u and β_u. With the benefit of ever-increasing knowledge about the volunteering ground users, the distribution is updated according to Bayes' rule and its shape becomes more concentrated, which represents a reduced uncertainty. As seen in Fig. 4, since the blockchain network has no *a priori* knowledge of the trust-levels of public users, an initial trust level of 0.5 is assigned to all participants. Prior to building each block, an estimated user trust-level $\hat{\theta}_u$ is drawn from the beta distribution and the orderer is elected based on $\hat{\theta}_u$. This procedure strikes a trade-off between exploring the participants associated with higher uncertainties in their trust-levels, and exploiting the participants associated with the highest estimated trust-level based on past observations. Upon electing the orderer, the rest of the participants will be divided into a pair of non-overlapping groups and assigned as monitors, as well as endorsers accordingly. Following the transactions flow (i.e., from monitors to orderer and then back to endorsers for block validation) as exhibited in Fig. 3, each endorser rates

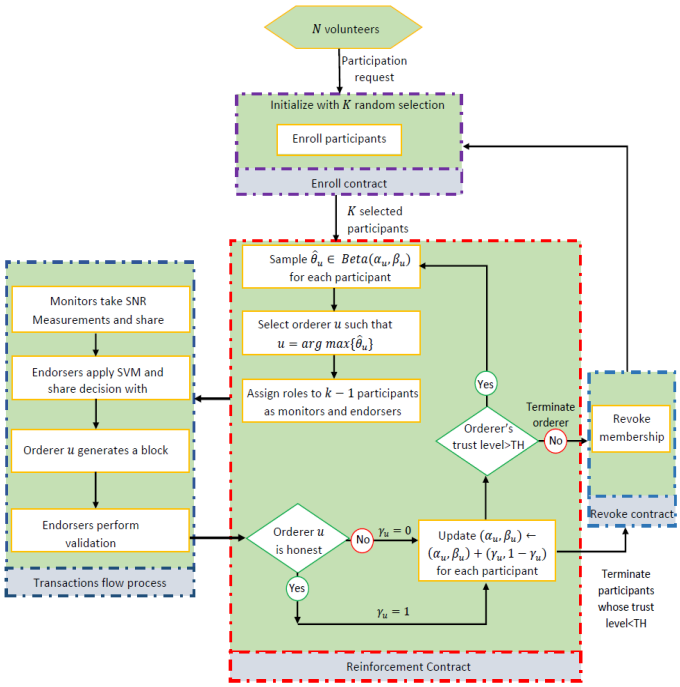

Fig. 4. Flow diagram of the proposed reinforcement learning.

the orderer's service and makes a binary decision indicating adequate or inadequate service provision. This information is then fed back to the reinforcement contract for making the final decision about the orderer's service based on the majority of votes as well as for updating its trust-level of service (as a resultant reward). In case if the orderer's trust-level falls below a certain threshold, revoke contract will be invoked for terminating its membership, and a new member will be enrolled (selected randomly) from the $(N - K)$ volunteers

using enroll contract.

C. Results

Consider a private blockchain network, where a UAV serves $N = 1000$ randomly positioned ground users. At each time $K = 50$ ground users are issued licenses to join the private blockchain network and at most 25 of them whose estimated trust-level fall below a threshold value of $\psi = 0.5$ can be replaced by some random ground users (out of 950 users) in the public network. Two benchmark schemes are employed in our simulation, namely the exploration-based and the conventional greedy benchmark schemes. The former explores the environment by selecting a random participant as the orderer at each epoch, whilst the latter chooses the orderer purely based on its past observations of the participants. An optimal scheme where the actual trust levels of users are perfectly known to the private blockchain network is further employed to indicate the performance upper bound. For fair comparison, simulation parameters such as the actual trust levels of the individual users are set identical for all the schemes.

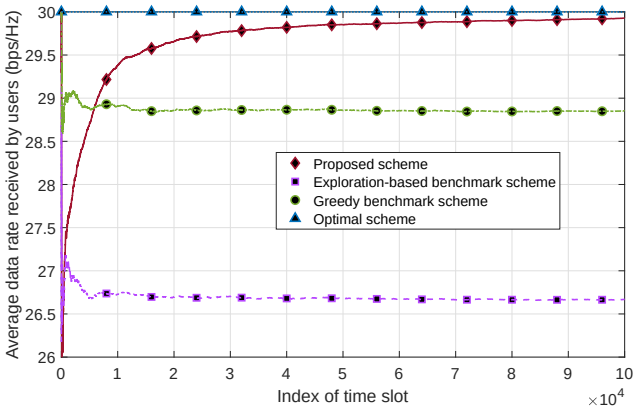

Fig. 5. Average received data rate for various schemes.

The data rates achieved by the users for all three schemes are compared in Fig. 5, where the average agreed data rate of users is assumed to be 30 bps/Hz. The proposed scheme is capable of compelling the UAV operator to provide the required QoS, while the greedy scheme and the exploration based scheme are only capable of attaining a throughput of 28.7 bps/HZ and 26.8 bps/Hz, respectively. One of the intuition is that choosing a block orderer purely based on the past knowledge without involving any learning process for exploring the environment - as in the greedy scheme - may get stuck in some sub-optimal actions. By contrast, selecting the orderer completely at random as in the exploration scheme squanders the historical data. Naturally, upon acquiring ever-increasing knowledge of the malicious behaviours of the ground users, the elected orderer of the proposed scheme is more likely to be trustworthy.

Table. I compares the average likelihood of the selected orderer being honest at the 10^5-th block for all three schemes against the number of participants K. It is evident from Table. I that amongst all the schemes, our proposed scheme has attained the best and most stable performance. This is because a well-balanced exploration-exploitation trade-off is implemented. By contrast, the greedy benchmark scheme may only attain a good performance, if the users associated with high trust levels are selected as the orderer in the initial epochs. Hence, the performance of the greedy benchmark scheme is somewhat unstable and highly depends on the initial observations of the behavior of the orderer.

Table I
Performance comparison against benchmark schemes

No. Participants	Average likelihood of honest orderer		
	Proposed scheme	Greedy scheme	Exploration-based scheme
300	0.9958	0.9055	0.6855
200	0.9958	0.9882	0.6848
100	0.9957	0.9045	0.6861
50	0.9957	0.9057	0.6866
30	0.9959	0.8205	0.6863
10	0.9955	0.8198	0.6873

IV. Conclusion

We have discussed the fundamentals of the blockchain technology and demonstrated both its potential and its key limitations in the context of communication networks. The advantages and disadvantages of private and public blockchain technologies have been analyzed and a reinforcement learning technique has been proposed for harnessing the power of both of these technologies. A case study based on a UAV-IoT network has been presented to demonstrate the potential of the proposed machine learning aided blockchain technology.

Acknowledgments

This work was supported in part by the Engineering and Physical Sciences Research Council under Grants EP/R006385/1, EP/N007840/1 and EP/P003990/1 (COALESCE), in part by the Royal Society's Global Challenges Research Fund Grant, in part by the European Research Council's Advanced Fellow Grant QuantCom, and in part by the International Scientific Partnership Program (ISPP) at King Saud University under Grant ISPP 134.

References

[1] K. Zhang, Y. Zhu, S. Maharjan, and Y. Zhang, "Edge intelligence and blockchain empowered 5G beyond for the industrial internet of things," *IEEE Network*, vol. 33, no. 5, pp. 12–19, Sep. 2019.

[2] P. K. Sharma, S. Singh, Y.-S. Jeong, and J. H. Park, "Distblocknet: A distributed blockchains-based secure SDN architecture for IoT networks," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 78–85, Sep. 2017.

[3] A. Dorri, S. S. Kanhere, R. Jurdak, and P. Gauravaram, "Blockchain for IoT security and privacy: The case study of a smart home," in *proc. IEEE international conference on pervasive computing and communications workshops*, Kona, HI, USA, May 2017, pp. 618–623.

[4] N. Z. Aitzhan and D. Svetinovic, "Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams," *IEEE Transactions on Dependable and Secure Computing*, vol. 15, no. 5, pp. 840–852, Sep. 2018.

[5] "Blockchain's real promise: Automating trust," Accessed: Jun 2020. [Online]. Available: https://www.technologyreview.com/2019/06/13/102979/blockchains-real-promise-automating-trust

[6] C. Jiang, H. Zhang, Y. Ren, Z. Han, K. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, April 2017.

[7] J. Wu, M. Dong, K. Ota, J. Li, and W. Yang, "Application-aware consensus management for software-defined intelligent blockchain in IoT," *IEEE Network*, vol. 34, no. 1, pp. 69–75, 2020.

[8] Y. Dai, D. Xu, S. Maharjan, Z. Chen, Q. He, and Y. Zhang, "Blockchain and deep reinforcement learning empowered intelligent 5G beyond," *IEEE Network*, vol. 33, no. 3, pp. 10–17, May 2019.

[9] Y. Lu, X. Huang, K. Zhang, S. Maharjan, and Y. Zhang, "Blockchain empowered asynchronous federated learning for secure data sharing in internet of vehicles," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 4298–4311, 2020.

[10] T. Salman, M. Zolanvari, A. Erbad, R. Jain, and M. Samaka, "Security services using blockchains: A state of the art survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 858–880, 2018.

[11] W. Wang, D. T. Hoang, P. Hu, Z. Xiong, D. Niyato, P. Wang, Y. Wen, and D. I. Kim, "A survey on consensus mechanisms and mining strategy management in blockchain networks," *IEEE Access*, vol. 7, pp. 22 328–22 370, Jan. 2019.

[12] M. Belotti, N. Božić, G. Pujolle, and S. Secci, "A vademecum on blockchain technologies: When, which and how," *IEEE Communications Surveys & Tutorials*, vol. 21, pp. 3796–3838, July. 2019.

[13] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[14] T. Nguyen, N. Tran, L. Loven, J. Partala, M.-T. Kechadi, and S. Pirttikangas, "Privacy-aware blockchain innovation for 6G: Challenges and opportunities," in *2nd 6G Wireless Summit (6G SUMMIT)*. IEEE, March 2020, pp. 1–5.

[15] Y. Dai, D. Xu, K. Zhang, S. Maharjan, and Y. Zhang, "Deep reinforcement learning and permissioned blockchain for content caching in vehicular edge computing and networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 4312–4324, 2020.

[16] E. Androulaki, A. Barger, V. Bortnikov, C. Cachin, K. Christidis, A. De Caro, D. Enyeart, C. Ferris, G. Laventman, Y. Manevich *et al.*, "Hyperledger fabric: a distributed operating system for permissioned blockchains," in *Proceedings of the Thirteenth EuroSys Conference*, no. 30. Porto, Portugal: ACM, Apr. 2018.

Amjad Saeed Khan received the B.Eng. degree in computer engineering from the COMSATS Institute of Information Technology, Pakistan, in 2010, and the M.Sc. degree in digital signal processing and intelligent systems and the Ph.D. degree in communication systems from Lancaster University, U.K., in 2013 and 2018, respectively. From 2018 to 2020, he was working as a Research Associate in signal processing for 5G networks with the Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University, U.K. He is currently a Lecturer with the School of Computing, Electronics and Mathematics, Coventry University, U.K. His research interests include 5G networks, network coding, secure wireless communication, digital signal processing, nonorthogonal multiple access, embedded systems design, machine learning, and blockchain technology.

Xinruo Zhang (S'15-M'19) received the B.Eng and the M.Sc degrees in electronics engineering and satellite communications engineering from Beihang University, China and University of Surrey, U.K. in 2010 and 2012, respectively, and the Ph.D. degree in Telecommunications Research from King's College London, U.K., in 2018. She is currently a Lecturer in the School of Computer Science and Electronic Engineering, University of Essex. Prior to that, she was a Research Associate with Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University. Her research interests include machine learning for wireless communications, radio resource allocation and mobile edge intelligence.

Sangarapillai Lambotharan (SM'06) received the Ph.D. degree in signal processing in 1997 from Imperial College London, London, where he remained until 1999 as a postdoctoral research associate. He was a visiting scientist with the Engineering and Theory Centre of Cornell University, USA in 1996. He is a Professor of Digital Communications and the Head of Signal Processing and Networks Research Group with the Wolfson School Mechanical, Electrical and Manufacturing Engineering, Loughborough University, Loughborough, U.K. Between 1999 and 2002, he was with Motorola Applied Research Group, U.K. and investigated various projects including physical link layer modeling and performance characterization of GPRS, EGPRS, and UTRAN. He was with Kings College London and Cardiff University as a Lecturer and Senior Lecturer, respectively, from 2002 to 2007. His current research interests include 5G networks, MIMO, radars, smart grids, machine learning, network security, and blockchain technology. He has authored more than 200 journal and conference articles in these areas. He currently serves as an Associate Editor for the IEEE Transactions on Signal Processing.

Gan Zheng (S'05-M'09-SM'12) received the BEng and the MEng from Tianjin University, Tianjin, China, in 2002 and 2004, respectively, both in Electronic and Information Engineering, and the PhD degree in Electrical and Electronic Engineering from The University of Hong Kong in 2008. He is currently Reader of Signal Processing for Wireless Communications in the Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University, UK. His research interests include machine learning for communications, UAV communications, mobile edge caching, full-duplex radio, and wireless power transfer. He is the first recipient for the 2013 IEEE Signal Processing Letters Best Paper Award, and he also received 2015 GLOBECOM Best Paper Award, and 2018 IEEE Technical Committee on Green Communications & Computing Best Paper Award. He currently serves as an Associate Editor for IEEE Communications Letters.

Basil AsSadhan received the M.S. degree in electrical and computer engineering from the University of Wisconsin and the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University. He is currently an Associate Professor with the Electrical Engineering Department, King Saud University. His research interests are in the areas of cybersecurity, network security, network traffic analysis, and anomaly detection.

Lajos Hanzo (M'91-SM'92-F'04) (http:www-mobile.ecs.soton.ac.uk, https:en.wikipedia.orgwikiLajos_Hanzo) (FIEEE'04, Fellow of the Royal Academy of Engineering F(REng), of the IET and of EURASIP), received his Master degree and Doctorate in 1976 and 1983, respectively from the Technical University (TU) of Budapest. He was also awarded the Doctor of Sciences (DSc) degree by the University of Southampton (2004) and Honorary Doctorates by the TU of Budapest (2009) and by the University of Edinburgh (2015). He is a Foreign Member of the Hungarian Academy of Sciences and a former Editor-in-Chief of the IEEE Press. He has served several terms as Governor of both IEEE ComSoc and of VTS. He has published 1900+ contributions at IEEE Xplore, 19 Wiley-IEEE Press books and has helped the fast-track career of 123 PhD students. Over 40 of them are Professors at various stages of their careers in academia and many of them are leading scientists in the wireless industry.