Yu, D., Sun, Y., Li, Y., Zhang, L. and Imran, M. (2023) Communication resource allocation of Raft in wireless network. IEEE Sensors Journal, (doi: 10.1109/JSEN.2023.3293715).

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

https://eprints.gla.ac.uk/302436/

Deposited on: 6 July 2023

Enlighten – Research publications by members of the University of Glasgow
https://eprints.gla.ac.uk

Communication Resource Allocation of Raft in Wireless Network

Dachao Yu, Yao Sun, Yuetai Li, Lei Zhang, and Muhammad Imran

***Abstract*— The distributed consensus intends to improve the reliability of critical decision making in wireless connected autonomous systems. The performance of distributed consensus heavily depends on the reliability of wireless links, which should be stochastic with limited communication resources. Therefore, advanced communication resource allocation schemes are needed to achieve high reliability and low latency for the distributed consensus. This article first derives optimized resource allocation schemes for the distributed consensus. The optimal number of nodes for the best reliability performance of the distributed consensus is also investigated to solve the inadequate overall communication resources issue. The revealed derivation and simulation results can provide guidelines to deploy the appropriate paradigm of communication resource allocation in autonomous wireless systems.**

***Index Terms*— Distributed consensus, Reliability, Latency, Resource allocation**

I. Introduction

Industrial scenarios, such as Autonomous Vehicles and Industrial Robots, usually require high reliability and low latency in critical decision-making within the network and essential data processing for distributed sensors and IoT devices. In these scenarios, local nodes from the network can collect data, make initial decisions, and send global consents to the joint nodes in the network. This is especially pertinent in diverse 5G-enabled networks, which include long-term Ultra-Reliable Communication for critical applications, Vehicle-to-Vehicle coordination for enhanced road safety, reliable cloud connectivity for seamless data exchange, and real-time virtualization to enable efficient network services. For example, a decentralized approach has been proposed for decision making in autonomous driving [1], which presents that the local nodes in distributed networks can collect data, make initial decisions and send the global consents (i.e., consensus) to the joint nodes in the distributed network. Because critical decisions making are reliability-intensive and latency-sensitive, a mechanism is required to enhance the reliability of decision-making in critical scenarios, and distributed consensus can work as the fault tolerant protocol for the critical decision making in this scheme [2].

Distributed consensus, which has been prevalently applied to distributed ledger technology (DLT), is defined as a protocol to ensure all normal nodes in the system can achieve the agreements on unified states, even if the network suffers from a certain amount of faulty progress or attack [3]. Therefore, the distributed consensus can work as an interior algorithm that regulates the decision based on the collected information by nodes. In the protocol of a distributed consensus, every participant is capable of transmitting and receiving the command to switch the state of replicas if it can follow specific fault-tolerant protocols. Crash failure and Byzantine failure are two types of errors that may occur in the distributed system. Crash failure refers to the failure that the progress abruptly stops and cannot resume. Crash fault tolerance (CFT) protocol, such as Raft [4] and Paxos [5], aims to manage reliable state duplication and prevent system breakdown from node crash failure. Byzantine failure represents the malicious behaviors given by an adversary, including contradictory commands to the progress, communication abort, and lengthy intentional delays to critical messages, which are more disruptive to the system than crash failures. Corresponding byzantine fault tolerance (BFT) protocols like PBFT [6] and Hotstuff BFT [7] have been introduced to the decentralized systems against the potential malicious attack [8].

In both CFT and BFT protocols, communication acts as a critical enabler to ensure that every node can exchange its state information with others in the distributed consensus. Currently, most of the distributed consensus usually is deployed through stable wired communication [9]. However, the majority of the upcoming generation of IoT networks have the trend to become wireless systems. For example, The protocol of distributed consensus can be deployed in DLT-enabled wireless networks [10]. Unlike the reliable link transmission in a wired network, wireless channels are more stochastic and dynamic. The link transmission failure that occurs in the wireless channel can have the same influence on the state synchronization as the node that has crash or byzantine faults within it. This influence should be addressed when distributed consensus is implemented in the wireless network.

Resource allocation for distributed consensus in wireless networks has been a focal point of research due to its significant impact on consensus performance. [11] have delved into the role of communication resources in the distributed consensus within wireless networks. They demonstrated the feasibility of consensus mechanisms for critical decision-making in distributed wireless communication systems, particularly through the implementation of a consensus-enabled industrial IoT network based on the PBFT protocol. However, wireless networks inherently face challenges such as the risk of link transmission errors and state synchronization loss [12]. The reliability of consensus protocols like Raft is closely tied to the reliability of wireless link transmissions [2]. In scenarios where excessive nodes intensively occupy limited wireless communication resources, there can be a decrease in both link and consensus reliability. This issue is particularly prevalent in massive IoT networks with wireless connections [13]. The above researches indicate that limited communication resources can compromise the reliability of link connections, thereby affecting the reliability of distributed consensus. This problem may increase the frequency of primary node changes, which can cause a longer latency for consensus completion and state synchronization among network nodes.

Therefore, reasonable and practical communication resource allocation methods should be investigated to achieve a better performance of the distributed consensus. [14] proposes the first joint interest, energy, and physical-aware framework for coalition formation among wireless IoT devices and energy-efficient resource allocation in M2M communication, considering mutual interest, energy availability, physical proximity, and communication channel quality, which not only ensures efficient and accurate coalitions but also increases overall system energy efficiency. Other researchers try to use machine learning in the optimization of resource allocation in wireless networks. [15] explores the use of machine learning algorithms for AP selection strategy and found that the Random Forest algorithm demonstrated superior performance in terms of accuracy and complexity in both the training and testing phases. [16] discusses the capacity maximization problem in wireless networks. The authors propose the use of machine learning techniques, specifically support vector machines (SVMs) and deep belief networks (DBNs), for direct approximation of optimal subproblem solutions. However, there are few papers that have systematically analyzed the communication resource allocation to the distributed consensus in wireless networks, which is the motivation of this paper.

In this article, we make efforts to optimize communication resource allocation to improve the reliability and reduce latency of Raft through different algorithms. Our main contributions are summarized as follows.

- We derive an optimal transmit power allocation method through Sequential Quadratic Programming (SQP) to maximize the reliability of Raft.
- The optimal bandwidth allocation method is investigated to minimize the latency in the distributed consensus. We choose Particle Swarm Optimization (PSO) as the optimization algorithm to search for the optimal bandwidth allocation scheme when overall bandwidth is constant.
- We investigate the optimal number of nodes deployed in the wireless network to maximize the reliability of distributed consensus when constant overall communication resources are provided. Relevant analytical proof has been provided to support the conclusion.

The structure of this paper is explained as follows. The protocol of the Raft is given in Section II. Section III introduces the algorithms of nonlinear optimization programming for the performance of the distributed consensus. Section IV proposes the optimized network size for Raft with limited overall communication resources. Section V compares the numerical results of the performance given by different resource allocation methods, which demonstrates the conclusion in Section VI.

II. Protocol of Raft

The protocol of distributed consensus has been deployed in many decentralized systems to keep the consistency of the state in nodes. In a system that requires a trusted authority to access (i.e., private blockchain [17]), the possibility that the system suffers from Byzantine fault can be negligible [18]. The crash of nodes and link transmission failure are the main threats to these trusted systems. Therefore, it is appropriate to deploy the CFT protocol in these scenarios. Raft, as a typical CFT consensus algorithm, is generally implemented in a private, trustworthy, distributed system to oppose the breakdown of replicas [4]. The simplicity of Raft has drawn attention to the research about its optimization and applications [19], [20].

Fig. 1 shows that the Raft-enabled distributed network, which is composed of a leader and a group of followers in the stage of log replication. The leader needs to pack the commands in log entries and replicate the entries to all followers ceaselessly through downlink transmission. Depending on the successful reception of log messages, the followers need to reply confirmation packets to the leader through uplink unicast and start to execute the confirmed commands. A successful Raft consensus represents that more than 50% overall followers have received the log entries from the leader and sent the confirmation back to the leader successfully within one term of the consensus. The voting for the leader follows the criteria of first come, first serve, which means the leader candidate with the most reliable wireless connections and lowest latency is most likely to be chosen as a leader.

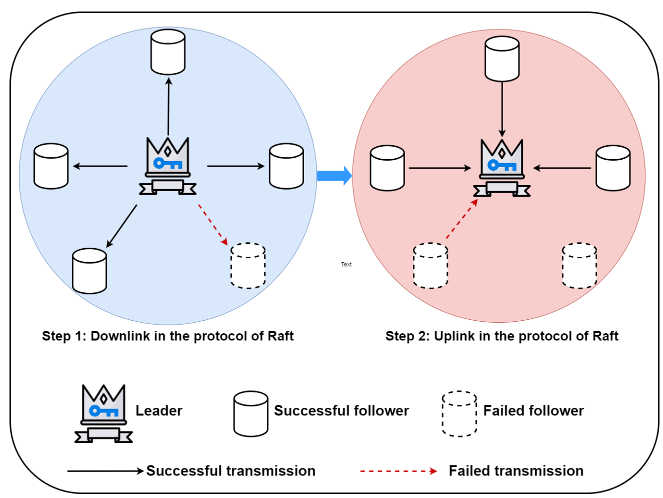


Fig. 1: Communication scheme of Raft

The protocol of Raft indicates that it relies on the inter-node information exchange to achieve the consensus among nodes [11]. Therefore the consensus reliability of Raft heavily depends on the reliability of the link connection between the leader and followers.

III. Communication Resource Allocation Schemes for Raft

Reliability and latency are the important performance metrics for the distributed consensus in wireless networks [12]. The consensus reliability P_C refers to the probability that most trusted nodes complete vote or log replication in a term and the latency of Raft, which includes the time consumed by one round of downlink and uplink transmissions between the leader and all followers and the time of message verification [21]. When the number of nodes in the network is constant, P_C only depends on the link reliability of channels, which refers to the probability of successful link transmission between the leader and followers [2]. Different resource allocation methods and stochastic fading gains may cause variations in the link reliability and transmission time among the channels between the leader and followers. Therefore, varied link reliabilities and latency of wireless channels are determined by a derived wireless link model in this section initially. And relevant optimization problems of resource allocation are solved based on the proposed link reliability and latency.

A. Wireless Link Model

The protocol of Raft is deployed on the considered wireless network that has $N+1$ static nodes, including a leader and N followers. The communication scheme in the protocol of Raft is assumed to be frequency division in this paper. The $2N$ channels, which include N downlink channels and N uplink channels that connect the leader and followers, are characterized by the Rayleigh fading model [22]. Rayleigh Fading is a statistical model for the effect of a propagation environment on a radio signal, such as that used by wireless devices. This model assumes that the magnitude of a signal that has passed through a communication channel will vary randomly, or fade, according to a Rayleigh distribution. It is viewed as a reasonable model in situations where the communication signal may bounce off objects from many directions before reaching the receiver, resulting in a large number of signal paths that can destructively interfere with each other. Rayleigh Fading Model simulates the worst-case scenario for signal distortion by a propagation environment. Therefore it is used extensively in designing wireless networks even if the channels are in terrible conditions. H_k denotes the Rayleigh fading gain of the k^{th} channel that $k \in [1, 2N]$, which follows the complex normal distribution, i.e., $H_k \sim \mathcal{CN}(0,1)$. The channel gains are assumed to be independent and identically distributed (i.i.d.). Therefore, $|H_k|^2$ follows the exponential distribution. When a package is sent through the k^{th} channel with a given transmit power P_{tk}, the signal-to-noise ratio (SNR) in this channel can be indicated as γ_k

$$\gamma_k = \frac{S_k |H_k|^2 P_{tk}}{P_{noise}}, \tag{1}$$

where P_{noise} refers to the white Gaussian noise power, S_k represents the large-scale effect on the k^{th} channel from the environment, such as the path loss and shadowing, and ρ is the SNR threshold. If γ_k is below the threshold ρ, the SNR outage occurs in the k^{th} channel. Consequently, the link reliability P_{lk} of the k^{th} channel can be calculated by the SNR outage probability in this channel [23]

$$P_{lk} = 1 - Pr(\gamma_k < \rho) = exp(-\frac{\rho P_{noise}}{S_k P_{tk}}), \tag{2}$$

which reveals that the transmit power P_{tk} is the communication resource that can affect the link reliability P_{lk} when other parameters keep constant in the wireless link model. Meanwhile, the latency cost by transmission in the k^{th} channel can be represented as

$$t_k = \frac{M}{B_k log(1+\gamma_k)}, \tag{3}$$

where M is the average length of the package sent by the leader or followers, and B_k is the bandwidth used in this channel. When the distributed consensus is implemented in the wireless network, the derived model of link reliability P_{lk} in (2) and time latency t_k in (3) can determine the critical parameters of the performance, such as consensus reliability P_C and the latency of consensus t_c. And the derived model shows that these performance parameters can be improved by optimizing the power and bandwidth allocation.

B. Power Allocation Scheme for Consensus Reliability

The model of wireless channel in (2) is implemented as an example to demonstrate the influence in the consensus reliability P_C given by the allocated transmit power P_{tk}, which is a prevalent type of communication resource that can influence the link reliability in practice. Therefore, P_{tk} is regarded as a variable of the communication resource allocation scheme to pursue the maximum consensus reliability P_C. The procedure of analysis can be similar when other wireless communication models are selected.

With the link reliability given by (2), the consensus reliability P_C can be represented as a function with the transmit power P_{tk}. The communication scheme of Raft in Fig. 1 shows that the successful follower needs to complete both the downlink and uplink transmission. Therefore, the consensus reliability P_C can be calculated as

$$P_C = \sum_{k=\frac{N}{2}+1}^{N} \sum_{Q_k \in \Omega_S} \prod_{w \in Q_k} P_w \prod_{v \in Q_k^C} (1 - P_v), \tag{4}$$

where Q_k refers to the set of k followers that successfully complete both the downlink and uplink transmission. Ω_S refers to the set that over $\frac{N}{2}$ followers have reached the consensus. W is a successful follower that belongs to Q_k and v is a failed follower that belongs to complement of set Q_k. P_w represents the probability that w belongs to the set Q_k

$$P_w = P_{lw}^{DL} P_{lw}^{UL}, \tag{5}$$

which is the product of the downlink reliability P_{lw}^{DL} and uplink reliability P_{lw}^{UL}. Similarly, P_v refers to the probability that nodes from v complete the downlink and uplink transmissions successfully

$$P_v = P_{lv}^{DL} P_{lv}^{UL}, \tag{6}$$

Other parameters in (4) are assumed constant for all $2N$ channels.

The scheme of power allocation aims to maximize the consensus reliability P_C when the overall transmit power P_{sum} is fixed. In the protocol of Raft, the overall transmit power P_{sum} is allocated to all $2N$ channels. Therefore, the problem of optimization for the power allocation scheme can be formulated as

$$\begin{aligned} \min_{P_t} \quad & 1 - P_C \\ \text{s.t.} & \sum_{k=1}^{2N} P_{tk} \le P_{sum}, \end{aligned} \tag{7}$$

This optimization problem has $2N$ variables of transmit power. The channels from 1 to N represent the downlink channel of N followers, and channels from $N+1$ to $2N$ are the corresponding uplink channel of N followers. Sequential quadratic programming (SQP) is implemented to solve the nonlinear programming in this resource allocation scheme, which aims to transform the original optimization problem into an optimal quadratic problem and find the appropriate descent direction d. The transformed quadratic optimal problem can be formulated as follows:

$$\begin{aligned} \min_{d} \quad & f(P_{tk}) + \nabla f(P_{tk})^T d + \frac{1}{2} d^T \nabla^2 L(P_{tk}, \lambda) d \\ & \text{s.t.} \nabla g(P_{tk}) d + g(P_{tk}) = 0, \end{aligned} \tag{8}$$

where $f(P_{tk})$ represents the objective function $1 - P_C$ with a vector of transmit power P_{tk} allocated to all $2N$ channels, $\nabla f(P_{tk})^T$ denotes the gradient of the transpose of $f(P_{tk})$, $g(P_{tk})$ denotes the constriant and $L(P_{tk}, \lambda)$ denotes the Lagrangian multiplier,

$$L(P_{tk}, \lambda) = f(P_{tk}) - \sum \lambda g(P_{tk}). \tag{9}$$

The objective function of the transformed quadratic optimal problem in (8) is the first three terms of the Taylor series from the original optimization problem [24]. The remainder R_n of the Taylor series [25] can be calculated as:

$$R_n = \sum_{n=3}^{+\infty} \frac{\nabla^n f(P_{tk})}{n!} d^n. \tag{10}$$

If the descent direction d is small in each iteration, the remainder R_n will converge to zero, which means the transformed optimization problem in (8) is equal to the original nonlinear optimization problem. Therefore, the solution to the optimization problem (7) is identical to the convergence of the result from SQP. However, consensus reliability P_C from (4) shows that the overall probability is the summation of the product of link reliabilities from $2N$ channels, which can exponentially increase the complexity of nonlinear programming. The high complexity can be impractical to deploy the scheme of communication resource allocation in a large-scale wireless network.

C. Comparison of Optimal Power Allocation and Other Power Allocation Schemes

Two power allocation methods, which can be practical to implement in reality, are proposed to compare with the performance of the optimal power allocation scheme from SQP. The first method is allocating the transmit power equally to each channel,

$$P_{tk}{}^1 = \frac{P_{sum}}{2N}. \tag{11}$$

With identical communication resources, the channel with better channel gain will have higher link reliability to complete transmission.

The second power allocation method aims to ensure all channels receive appropriate transmit power P_t to reach the same link reliability P_l, which follows the proportion of the channel fading gain S_k in each channel to the summation of channel fading gains from all $2N$ channels. The link reliability in (2) indicates that the transmit power P_{tk} is inversely proportional to S_k when link reliability P_{lk} is constant. Therefore, the link reliability in this power allocation method should be

$$P_{tk}{}^2 = \frac{P_{sum} \sum_{k=1}^{2N} S_k}{S_k}. \tag{12}$$

According to the inversely proportional relationship between the transmit power P_t and fading gain S_k when the link reliabilities of all channels tend to be identical with this allocation method, more transmit power should be compensated to the communication channel with lower S_k to keep the identical link reliability. These two power allocation methods have lower complexity than the result of SQP, which means they can replace the optimal power allocation method from the nonlinear optimization if the gap between their performances can be tolerated.

D. Bandwidth Allocation Scheme for Consensus Latency

Besides reliability, latency is also critical to the performance of distributed consensus. Consensus reliability and transmission time are two factors that can influence the overall latency of distributed consensus in a wireless network. Optimal consensus reliability indicates that the protocol of Raft has the maximum probability of preventing a new leader election and spending extra time on this stage. Therefore, an optimal consensus latency means the reliability of consensus needs to reach the maximum, which means the power allocation method in this condition should be optimal, and it follows the result of SQP, then the only factor that can change the consensus latency is the transmission time cost by nodes. Based on the model in (3), the consensus latency can be reduced by minimizing the transmission time through an optimal bandwidth allocation method. In this section, we aim to investigate this optimal bandwidth allocation scheme to pursue the minimum value of consensus latency.

The protocol of Raft indicates that each follower needs to receive a downlink message from the leader and respond with confirmation through uplink transmission in one term of consensus. The time that $\forall n \in 1, 2..., N$ follower spends in completing the consensus can be represented as

$$\begin{aligned} t_n &= t_n^{DL} + t_n^{UL} + t_v \\ &= \frac{M^{DL}}{B_n^{DL} log(1 + SNR_n^{DL})} + \frac{M^{UL}}{B_n^{UL} log(1 + SNR_n^{UL})} + t_v, \end{aligned} \tag{13}$$

which is the summation of delays caused by the downlink t_n^{DL}, uplink transmissions t_n^{UL} and verification time t_v. M^{DL} and M^{UL} refer to the package length during downlink and uplink transmission. In the same round of communications, the protocol of Raft indicates that M^{DL} and M^{UL} are identical for all downlink and uplink channels, respectively. All nodes are assumed to have the same ability to handle the verification, so the verification time t_v of all N followers is the same. The derived model of latency in (13) shows the bandwidth allocated to n^{th} channel is the communication resource that can influence the transmission latency t_n besides the SNR of channels. The consensus ends up the term when the last follower completes its transmission. Therefore, the longest latency cost by the follower can be considered as the latency t_c of distributed consensus.

$$t_c = \max \{t_1, t_2, ..., t_N\}, \tag{14}$$

which derives an optimization problem to solve the minimum value of t_c when the overall bandwidth B_{sum} is constant.

$$\begin{aligned} \min_{B} \quad & t_c \\ \text{s.t.} \sum_{k=1}^{2N} & B_k \le B_{sum}. \end{aligned} \tag{15}$$

where SNR in all downlink and uplink channels of the followers are based on the result of SQP in the section III-B, which means the consensus reliability P_C converges to the theoretical maximum value in this scheme. The overall bandwidth B_{sum} is the constraint for this optimization problem. Table. I shows the notations of major parameters used in the proposed resource allocation schemes.

TABLE I: Notation used in resource allocation of Raft-enabled Network

Notation	Definition
N	Number of Nodes within network
S_k	Large Scale Effect of the k^{th} channel
H_k	Rayleigh Fading Gain of the k^{th} channel
P_{sum} (dBm)	The overall transmit power
B_{sum} (MHz)	The overall bandwidth
P_{tk} (dBm)	Transmit Power allocated to the k^{th} channel
B_k (MHz)	Bandwidth allocated to the k^{th} channel
P_{lk}	Link reliability of the k^{th} channel
P_C	Consensus reliability
t_k (s)	Transmission time of the k^{th} channel
t_c (s)	Transmission time cost by consensus
N_{max}	Number of node with maximized consensus reliability

The optimization problem presented in equation (15) is non-linear, and its objective function lacks an explicit closed-form solution, implying that the solution is complex and cannot be obtained through straightforward mathematical methods. Thus, we have employed Particle Swarm Optimization (PSO) to iteratively resolve this optimization problem and find the minimum value of t_c. The PSO algorithm, renowned for its prowess in global optimization, enables us to evade suboptimal solutions [26]. In the context of our study, Algorithm.1 represents the application of PSO in bandwidth allocation within the Raft consensus algorithm. The position of the particles in this algorithm corresponds to the bandwidth distributed to the wireless channels. The PSO's inertia weight w, along with acceleration constants $c1$ and $c2$, guide the particle's movements and drive it towards the historically optimal and collective optimal position. The position of a particle gets

Algorithm 1 PSO algorithm for t_c

```
Initialize population
for m = 1 : Iterations do
  for i = 1 : n do
    t_{i,m} = f(B_{i,m})
    if t_{i,m} < t_{i,h} then
      t_{i,h} = t_{i,m}
      B_{i,h} = B_{i,m}
    else
      t_{i,h} = t_{i,h}
      B_{i,h} = B_{i,h}
    end if
    t_{i,opt} = min(t_{i,m})
    B_{i,opt} = B_{min(t_{i,m})}
  end for
  for i = 1 : n do
    v_i(m+1) = w v_i(m) + c_1 r_1 (B_{i,opt} - B_i) + c_2 r_2 (B_{i,h} - B_i)
    B_i(m+1) = B_i(m) + V_i(m+1)
    if V_i(m+1) > V_max then
      V_i(m+1) = V_max
    else if V_i(m+1) < V_min then
      V_i(m+1) = V_min
    end if
    if B_i(m+1) > B_{i,max} then
      B_i(m+1) = B_{i,max}
    else if B_i(m+1) < B_{i,min} then
      B_i(m+1) = B_{i,min}
    end if
  end for
end for
```

updated iteratively through the combination of its inertia weight and acceleration constants. After sufficient iterations, we are able to derive $t_{i,opt}$ as the maximum value of consensus latency t_c, a testament to PSO's effectiveness in exploring and converging towards an optimal solution in a complex problem space.

In the protocol of Raft, all followers need to occupy a constant overall bandwidth. A reasonable expectation of the optimization result is that most of the followers' latency t_n tends to be close when the optimal bandwidth allocation method is implemented because a non-optimal bandwidth allocation method can cause some followers to cost more time to complete the transmission, which increases the overall latency of distributed consensus in wireless networks. However, the stochastic wireless channel between the leader and some followers may have extremely terrible conditions, which can occupy a large proportion of communication resources and limit the optimal performance of the distributed consensus.

IV. Limited overall Communication Resource and Optimal Number of Nodes

The algorithms of nonlinear optimization proposed in Section. III can solve the optimization problem of the communication resource allocation to achieve the maximum consensus reliability P_C and minimum consensus latency t_c. However, if the overall communication resources are not adequate, even the optimal consensus reliability and latency cannot reach the requirement of high reliability and low latency in specific scenarios. This section aims to investigate the solution to the problem of inadequate overall communication resources in resource allocation. Firstly, the criteria of adequate communication resources for the distributed consensus Raft is defined. Then we find out the solution based on the feature of fault tolerance in the distributed consensus to improve the performance of the optimized consensus reliability and latency from the perspective of network size.

A. Limited Overall Communication Resource for Raft

In the assumption of this article, the allocated communication resources to the wireless channels and channel gains are the parameters that can influence link reliability P_l and the consensus reliability P_C. Therefore, the link reliability P_l and consensus reliability P_C can be reasonable criteria to judge the condition of overall communication resources when the wireless channel gain is determined. The reliability of information delivery and synchronization changes in different applications. These reliability requirements correspond to the consensus reliability if the distributed consensus is implemented. The dotted lines in Fig. 2 denote the target consensus reliability in multiple 5G scenarios, including URC over the long term, V2V wireless coordination, Reliable cloud connectivity, and Real-time Virtualization [27] [28].

The optimization problem in (7) indicates that even though the power allocation method is optimized by SQP, adequate overall transmit power should also be provided if the consensus reliability needs to be improved to reach the requirement of a specific scenario. Otherwise, an alternative solution should be implemented to improve the consensus reliability of Raft in the wireless network.

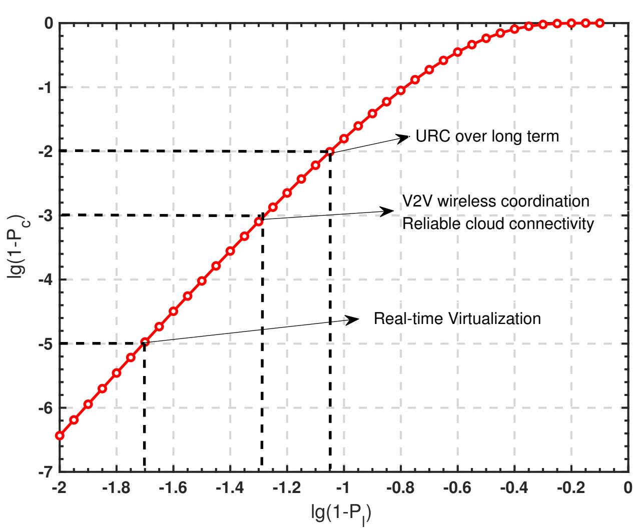


Fig. 2: Reliability requirements in different scenarios

B. Optimal Number of Nodes

When the overall communication resource is constant, the number of nodes that participate in the distributed consensus can influence the performance of distributed consensus because more nodes should occupy the limited communication resources, and each node is expected to take fewer resources for the transmission. Specifically, the performance of the resource allocation method will be damaged when the overall communication resources are inadequate because some channels cannot gain enough resources to achieve the target performance.

A reasonable solution to this problem is eliminating the redundant consensus nodes that are linked with terrible communication channels. However, the increasing size of the network represents that the distributed consensus can tolerate more crash faults or byzantine fault nodes [29]. These two controversial characteristics can cause the maximum global value for the reliability of consensus P_C with a dynamic number of nodes but constant communication resources for a local wireless network. The corresponding number of nodes N to the maximum of P_C can be determined by Proposition 1. It shows that when the overall communication resources are inadequate for a distributed network, the number of nodes engaged in this network should be less than the value of N_{max}. The maximum value of the consensus reliability P_C indicates that excessive consensus nodes can damage the reliability of Raft. Therefore, a large-scale network can abandon some nodes that have terrible communication channels to converge the number of nodes N to N_{max} if the overall communication resource is rare, which can improve the consensus reliability of Raft. For example, In a multiple-layer consensus network [30], the network size in the consensus layers can be optimized based on the communication resource allocated to them, which helps the whole network achieve the highest performance.

Proposition 1: If N_{max} is assumed as the number of followers that can reach the maximum of consensus reliability,

$$N_{max} = \lceil M_a \rceil = \lfloor M_b \rfloor. \tag{16}$$

M_a and M_b correspond to the value of function

$$M_a = \frac{\widetilde{P} - \sqrt{\widetilde{P}^2 - 4\widetilde{P} + 1}}{\frac{1}{2} - 2\widetilde{P}}$$
$$M_b = \frac{1 - 3\widetilde{P} - \sqrt{\widetilde{P}^2 - 4\widetilde{P} + 1}}{\frac{1}{2} - 2\widetilde{P}}, \tag{17}$$

where $\widetilde{P} = (1 - P_l^2)P_l^2$ and P_l denotes the average link reliability of channels

Proof: See Appendix A ■

The computational complexity of the model revolves around the calculation of N_{max}, which is the optimal number of nodes that can reach the maximum consensus reliability. Calculating N_{max} involves solving the equation (17), which is the function of link reliability $P(N)$. $P(N)$ is a function with $2N$ variables, which means the calculation of $P(N)$ can involve iterating over all $2N$ variables at least once. Therefore, the computational complexity for $P(N)$ will be

$O(N)$. Subsequently, N_{max} is calculated from $P(N)$ with the equation (17), which are operations with constant computational complexity. Therefore, the overall computational complexity of the model primarily depends on the calculation of $P(N)$ and is $O(N)$.

While the proposed model's computational complexity is linear in the size of network, the feasibility of real-time or near-real-time implementation of the proposed model depends on the number of nodes N and environmental effects. If N is large in the network, the calculation of link reliability $P(N)$ can be computationally intensive, which makes real-time implementation challenging. Moreover, the dynamic change of the communication environment causes a varied distribution of link reliability among nodes, and the Raft-enabled network has to frequently recalculate the optimal resource allocation scheme, which may pose influence the real-time implementation of the proposed model. Therefore, an ideal condition for the real-time deployment of the proposed model should contain an appropriate number of nodes within the network and a stable communication environment.

V. Simulation Results

In this section, the proposed resource allocation schemes for Raft are simulated in MATLAB R2019b. Based on the Rayleigh Fading model, we assume the channel fading gain H_k and large-scale effect S_k of $2N$ channels from (1) are in the Gaussian distribution [22]. The nodes are set as static nodes, and the number of them N in the wireless network is set to 13. The overall power P_{sum} ranges from 20 dBm to 36 dBm for the transmit power allocation. The Coefficient of Variation (CV), which refers to the ratio of standard derivation to mean of channel fading gain H and large-scale effect S in the wireless model, is implemented in the simulation to represent the dispersion in the probability distribution of wireless channel fading gains and large-scale effect. A higher CV means that part of channels have more probabilities of suffering terrible fading gain H and large-scale effect S, which influence the performance of proposed resource allocation schemes.

The optimal reliability of the distributed consensus P_C from SQP is compared with the other two transmit power allocation methods. The numerical results of three transmit power allocation methods are presented in Fig. 3 when the channel gains S_k has a high coefficient of variation ($CV = 1.303$). The consensus reliability given by the three allocation methods is significantly different. The output P_C from the equal power method in (11) is closer to the optimized result of SQP, which reveals that the equal power allocation method has better performance than the equal link reliability method when the variation of channel gains is large. Even though the complexity of SQP will rise when the size of the network increases, the transmit power allocation method derived by SQP is still the best allocation method to use in this case.

Moreover, Fig. 4 shows that when the channel fading gain is more concentrated ($CV = 0.388$), the curves of equal power and equal link reliability methods will converge to the optimized consensus failure rate $1 - P_{Copt}$, which means three power allocation methods will have similar performances when the conditions of wireless channels are close. Therefore, two practical transmit power allocation methods in (11) and (12) can substitute the optimal power allocation method derived by SQP in this case.

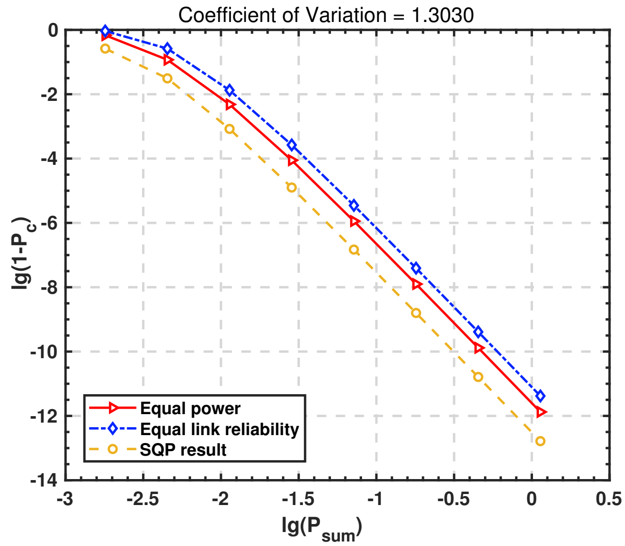


Fig. 3: Performance of three power allocation methods with a high coefficient of variation in channel gains

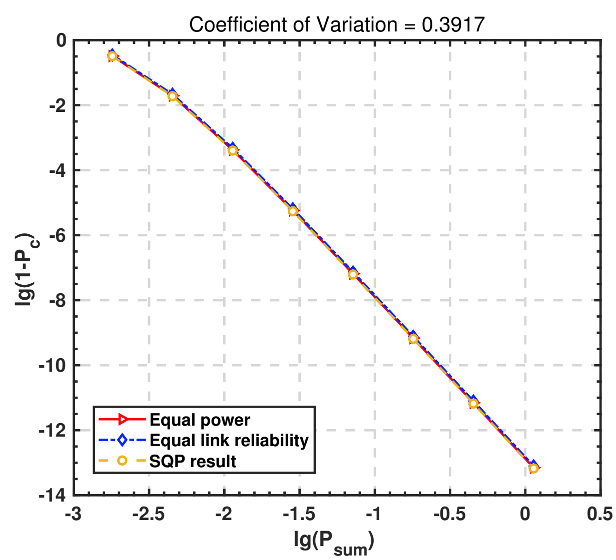


Fig. 4: Performance of three power allocation methods with low coefficient of variation in channel gains

Fig. 5 illustrates the influence of the varied channel gains in consensus reliability where P_C denotes the consensus reliability derived by the two practical power allocation methods in (11) and (12), P_{Copt} is the optimal consensus reliability from SQP, and Reliability Gap (RG) represents the ratio of consensus failure rate between $1 - P_C$ and $1 - P_{Copt}$. The difference among the three allocation methods gradually increases when the CV of channel gains is rising. All three methods have approximate results when the CV is less than 0.5, which means the other two power allocation methods can replace the optimal power allocation method derived by SQP with a small compromise performance. In practice, the CV of wireless channel gain can be reduced by abandoning some nodes with bad channel conditions (e.g., low large-scale effect S, etc.) to achieve a near-optimal power allocation scheme, which is supported by the feature of fault tolerance in the distributed consensus.

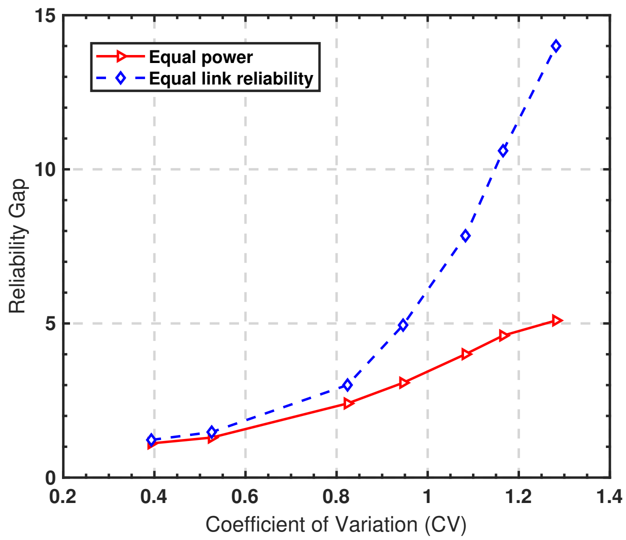


Fig. 5: The performance comparison among optimal consensus reliability and other two methods with different CVs in wireless channel gains

The simulation of bandwidth allocation assumes that the amount of overall bandwidth B_{sum} ranges from 8 to 14 MHz, and the number of nodes $N = 13$. The model of the wireless channel is the same as the previous transmit power allocation, and the SNRs of all channels are set based on the optimal result of transmit power allocation scheme from SQP. The iteration rounds are set to 500 in the PSO algorithm. The curve of the fitness function in the proposed optimization problem should be presented first. Fig. 6 shows the convergence of the optimal consensus latency when different overall bandwidths are used in the same wireless network. The convergence of consensus latency decreases when more overall bandwidth is provided for the communication. The number of iterations that the result of PSO converges to the minimum consensus reliability is between 100 to 150.

The transmission time cost by all followers is evaluated in Fig. 7 when the optimal bandwidth allocation scheme is exploited. Because the definition of consensus latency refers to the longest time cost by the follower from the whole wireless network, the simulation result matches the expectation that the transmission time cost by most of the followers is close when the consensus latency t_c reaches a minimum value.

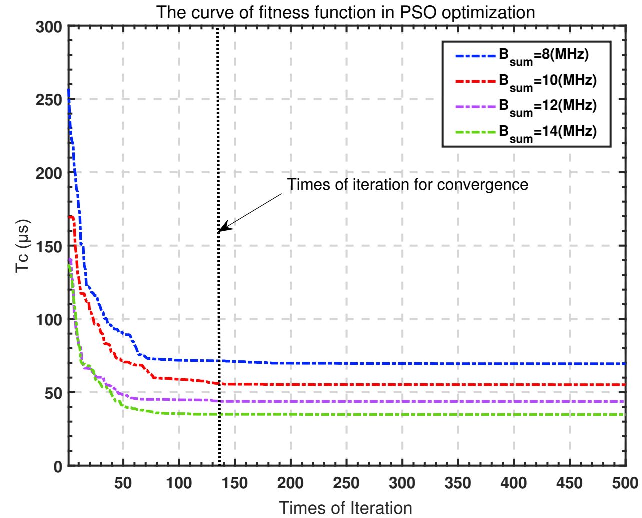


Fig. 6: The curve of fitness function in PSO

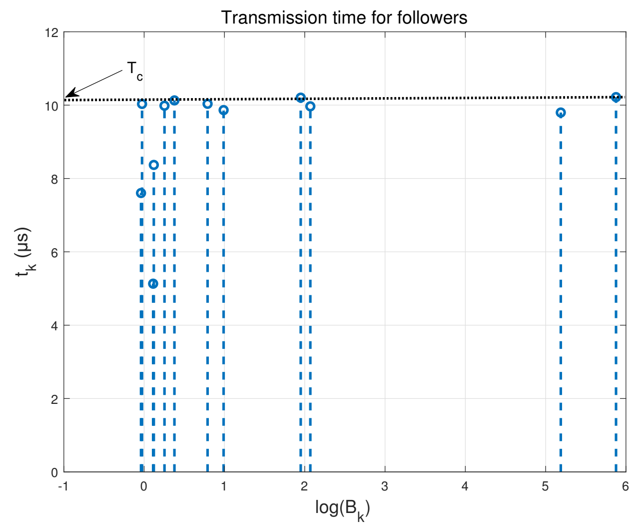


Fig. 7: The transmission time used by followers with optimized bandwidth allocation scheme

The stochastic wireless channels between the leader and followers have variable channel gains, which can have a significant influence on consensus latency. Fig. 8 aims to indicate the tendency of optimized consensus latency t_c with an increased coefficient of variation CV in channel gain S_k. The results show that when CV increases from 0.74 to 1.56, the optimal consensus latency t_c dramatically rises from 1 μs to 10^5 μs. This numerical result reveals a larger variation of channel gain can increase the optimal latency of Raft in the wireless network.

The simulation of the optimal number of nodes is presented in Fig. 9, which illustrates the change in the consensus reliability when the number of nodes in the network increases. The number of nodes N is assumed to range from 4 to 40, and the overall communication resource keeps constant. The trend of consensus reliability increases first and then drops when the number of followers reaches the optimal network size and finally increases. The number of nodes that corresponds

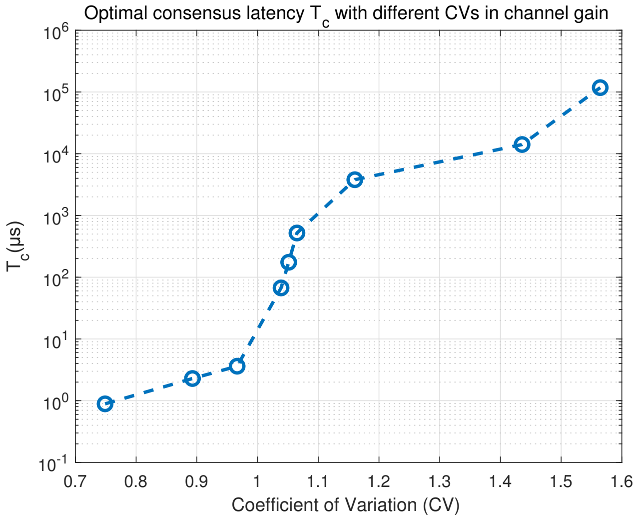


Fig. 8: The optimal consensus latency with different CV in the channel gains

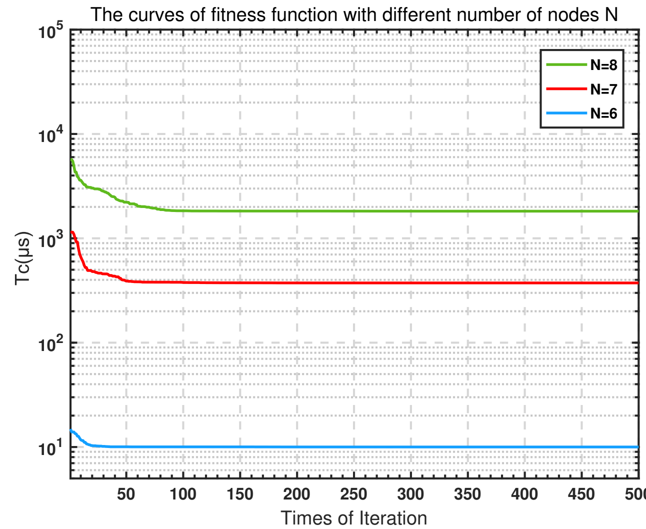


Fig. 10: The convergence of consensus latency with different numbers of followers

to the maximum consensus reliability matches the result of the optimal number of nodes in Proposition 1. S represents the rounds of synchronization processed during the Raft consensus protocol. When more rounds of synchronization S are implemented to the distributed consensus protocol, the maximum value of consensus reliability P_C will increase. But the eventual tendencies of all curves still remain the same.

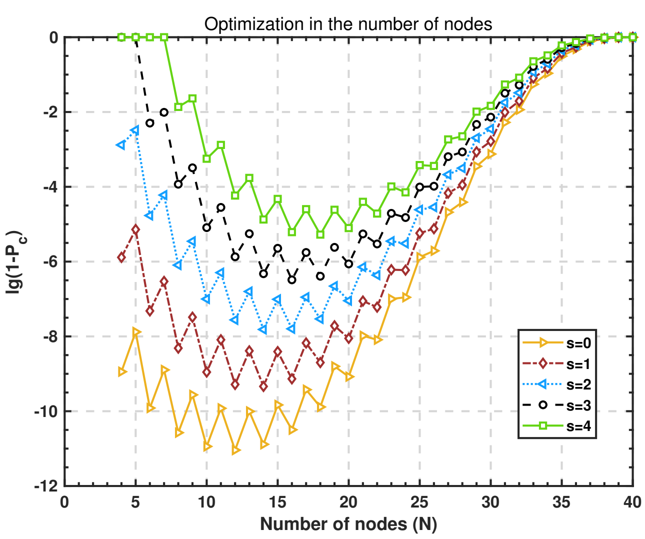


Fig. 9: Optimal network size for Raft

The negative influence on consensus latency from varied wireless channel gains indicates that if the consensus latency needs to be improved, the node with terrible channel gain should be removed. Fig. 10 compares the numerical result of optimal consensus latency t_c before and after the followers with the worst channel gains are eliminated from the network. The number of followers $N = 8$ in the initial network. The channel gains S_k of all nodes follow the normal distribution. The convergence of optimized t_c is close to 2000 μs when no followers are removed. The convergence of t_c drops to the region between 300 and 400 μs when one follower with the worst channel gain is removed. And t_c will keep dropping to 10 μs after two followers are removed from the network, which proves this method is also efficient in reducing the consensus latency.

VI. Conclusion

In this article, optimal power and bandwidth allocation methods are proposed to improve reliability and reduce latency for the distributed consensus Raft in a wireless network. Both power and bandwidth allocation methods, which are derived through two different optimization algorithms, can reach near-optimal performance when the overall communication resource is constant. Moreover, an optimized network size is defined to provide the solution to the scenario that the overall resources are inadequate to reach the required performance. These results can provide a guideline for the deployment of resource allocation schemes when consensus Raft is implemented in the distributed wireless network.

Appendix

The dominant term of consensus reliability P_C from (4) is a discrete function, which means P_C cannot determine its tendency through derivation. If the Raft consensus with N followers can reach the maximum consensus reliability $P_C(N)$, $P_C(N)$ should be less than the consensus reliability of the network that contains $N-2$ and $N+2$ followers.

$$\begin{cases} P_C(N) > P_C(N+2) \\ \\ P_C(N) > P_C(N-2). \end{cases} \tag{18}$$

In the problem of communication resource allocation, if the network with N followers can reach the minimum consensus failure rate, the overall communication resource can be regarded as adequate for this network, which means the dominant term of (4) can replace the whole consensus reliability P_C. Therefore, the difference among the average link reliability P_l in the network of N, $N-2$, and $N+2$ followers can be negligible. The dominant term of the consensus failure rate is substituted into (18) to solve the N_{max}

$$\begin{cases} \frac{\binom{N}{f+1}(1-{P_l}^2)^{f+1}({P_l}^2)^{N-f-1}}{\binom{N-2}{f}(1-{P_l}^2)^{f}({P_l}^2)^{N-f-2}} < 1 \\ \\ \frac{\binom{N}{f+1}(1-{P_l}^2)^{f+1}({P_l}^2)^{N-f-1}}{\binom{N+2}{f+2}(1-{P_l}^2)^{f+2}({P_l}^2)^{N-f}} < 1. \end{cases} \tag{19}$$

Eventually, the conclusion in Proposition 1 can be derived by replacing the number of fault tolerant nodes $f = \frac{N}{2}$ in (19) when the distributed consensus protocol is Raft.

References

[1] C. Feng, Z. Xu, X. Zhu, P. Valente Klaine, and L. Zhang, "Wireless distributed consensus in vehicle to vehicle networks for autonomous driving," *IEEE Transactions on Vehicular Technology*, 2022.

[2] D. Yu, W. Li, H. Xu, and L. Zhang, "Low reliable and low latency communications for mission critical distributed industrial internet of things," *IEEE Communications Letters*, 2020.

[3] L. Lamport, "Generalized consensus and paxos," 2005.

[4] D. Ongaro and J. Ousterhout, "In search of an understandable consensus algorithm," in *2014 {USENIX} Annual Technical Conference ({USENIX}{ATC} 14)*, pp. 305–319, 2014.

[5] L. Lamport and M. Massa, "Cheap paxos," pp. 307–314, 2004.

[6] M. Castro, B. Liskov, *et al.*, "Practical byzantine fault tolerance," in *OSDI*, vol. 99, pp. 173–186, 1999.

[7] M. Yin, D. Malkhi, M. K. Reiter, G. G. Gueta, and I. Abraham, "Hotstuff: BFT consensus with linearity and responsiveness," in *Proceedings of the 2019 ACM Symposium on Principles of Distributed Computing*, pp. 347–356, 2019.

[8] M. Baudet, A. Ching, A. Chursin, G. Danezis, F. Garillot, Z. Li, D. Malkhi, O. Naor, D. Perelman, and A. Sonnino, "State machine replication in the libra blockchain," *The Libra Assn., Tech. Rep*, 2019.

[9] M. Van Steen, *Distributed systems*. Citeseer, 2017.

[10] Y. Sun, L. Zhang, G. Feng, B. Yang, B. Cao, and M. A. Imran, "Blockchain-enabled wireless internet of things: Performance analysis and optimal communication node deployment," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 5791–5802, 2019.

[11] L. Zhang, H. Xu, O. Onireti, M. A. Imran, and B. Cao, "How much communication resource is needed to run a wireless blockchain network?," *IEEE Network*, pp. 1–8, 2021.

[12] H. Seo, J. Park, M. Bennis, and W. Choi, "Communication and consensus co-design for distributed, low-latency, and reliable wireless systems," *IEEE Internet of Things Journal*, vol. 8, no. 1, pp. 129–143, 2021.

[13] Z. Sun, Z. Wei, N. Yang, and X. Zhou, "Two-tier communication for UAV-Enabled massive IoT systems: Performance analysis and joint design of trajectory and resource allocation," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 4, pp. 1132–1146, 2021.

[14] E. E. Tsiropoulou, S. T. Paruchuri, and J. S. Baras, "Interest, energy and physical-aware coalition formation and resource allocation in smart iot applications," in *2017 51st Annual conference on information sciences and systems (CISS)*, pp. 1–6, IEEE, 2017.

[15] D. Militani, S. Vieira, E. Valadão, K. Neles, R. Rosa, and D. Z. Rodríguez, "A machine learning model to resource allocation service for access point on wireless network," in *2019 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pp. 1–6, 2019.

[16] X. Cao, R. Ma, L. Liu, H. Shi, Y. Cheng, and C. Sun, "A machine learning-based algorithm for joint scheduling and power control in wireless networks," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4308–4318, 2018.

[17] 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*, pp. 1–15, 2018.

[18] L. Lamport, R. Shostak, and M. Pease, "The byzantine generals problem," in *Concurrency: the Works of Leslie Lamport*, pp. 203–226, 2019.

[19] S. Pedersen, H. Meling, and L. Jehl, "An analysis of quorum-based abstractions: A case study using gorums to implement raft," in *Proceedings of the 2018 Workshop on Advanced Tools, Programming Languages, and PLatforms for Implementing and Evaluating Algorithms for Distributed systems*, pp. 29–35, 2018.

[20] J. Polge, J. Robert, and Y. Le Traon, "Permissioned blockchain frameworks in the industry: A comparison," *ICT Express*, vol. 7, no. 2, pp. 229–233, 2021.

[21] E. Sakic and W. Kellerer, "Response time and availability study of raft consensus in distributed SDN control plane," *IEEE Transactions on Network and Service Management*, vol. 15, no. 1, pp. 304–318, 2017.

[22] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[23] N. C. Beaulieu and J. Hu, "A closed-form expression for the outage probability of decode-and-forward relaying in dissimilar rayleigh fading channels," *IEEE Communications Letters*, vol. 10, no. 12, pp. 813–815, 2006.

[24] P. T. Boggs and J. W. Tolle, "Sequential quadratic programming," *Acta numerica*, vol. 4, pp. 1–51, 1995.

[25] M. Kline, *Calculus: an intuitive and physical approach*. Courier Corporation, 1998.

[26] R. Poli, J. Kennedy, and T. Blackwell, "Particle swarm optimization," *Swarm intelligence*, vol. 1, no. 1, pp. 33–57, 2007.

[27] P. Popovski, "Ultra-reliable communication in 5G wireless systems," in *1st International Conference on 5G for Ubiquitous Connectivity*, pp. 146–151, 2014.

[28] S. Zhang, X. Xu, Y. Wu, and L. Lu, "5G: Towards energy-efficient, low-latency and high-reliable communications networks," in *2014 IEEE international conference on communication systems*, pp. 197–201, IEEE, 2014.

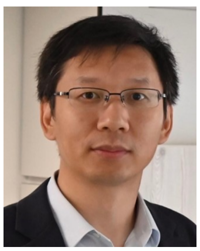
[29] D. Yu, H. Xu, L. Zhang, B. Cao, and M. A. Imran, "Security analysis of sharding in the blockchain system," in *2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1030–1035, IEEE, 2021.

[30] W. Li, C. Feng, L. Zhang, H. Xu, B. Cao, and M. A. Imran, "A scalable multi-layer PBFT consensus for blockchain," *IEEE Transactions on Parallel and Distributed Systems*, vol. 32, no. 5, pp. 1146–1160, 2021.

Dachao Yu received the B.S. degree in Electronic and Electrical Engineering from University of Electronic Science and Technology of China in 2019. He is currently pursuing the Ph.D. degree in Electronics and Communication Engineering at University of Glasgow. His research interest includes the performance analysis and optimization to the Crash fault tolerance and Byzantine Fault tolerance consensus in wireless network, security analysis of wireless blockchain system.

Yao Sun is currently a Lecturer with James Watt School of Engineering, the University of Glasgow, Glasgow, UK. Dr. Sun has extensive research experience and has published widely in wireless networking research. He has won the IEEE Communication Society of TAOS Best Paper Award in 2019 ICC, IEEE IoT Journal Best Paper Award 2022 and Best Paper Award in 22nd ICCT. He has been the guest editor for special issues of several international journals. He has served as TPC Chair for UCET 2021, and TPC member for a number of international flagship conferences, including ICC 2022, VTC spring 2022, GLOBECOM 2020, WCNC 2019. His research interests include intelligent wireless networking, semantic communications, blockchain system, and resource management in next generation mobile networks. Dr. Sun is a senior member of IEEE.

Yuetai Li is currently an undergraduate major in communication engineering from the university of Glasgow and the University of Electronic Science and Technology of China (UESTC). His current research interests include distributed consensus, blockchain, information security, and distributed intelligent systems.



Lei Zhang (Senior Member, IEEE) is a Professor of Trustworthy Systems at the University of Glasgow. He has academia and industry combined research experience on wireless communications and networks, and distributed systems for IoT, blockchain, autonomous systems. His 20 patents are granted/filed in 30+ countries/regions. He published 3 books, and 150+ papers in peer-reviewed journals, conferences and edited books. Prof. Zhang is an associate editor of IoT Journal, IEEE Wireless Communications Letters and Digital Communications and Networks, and a guest editor of IEEE JSAC. He received the IEEE Internet of Things Journal Best Paper Award 2022, IEEE ComSoc TAOS Technical Committee Best Paper Award 2019 and IEEE ICEICT'21 Best Paper Award. Dr. Zhang is the founding Chair of IEEE Special Interest Group on Wireless Blockchain Networks in IEEE Cognitive Networks Technical Committee (TCCN). He delivered tutorials in IEEE ICC'20, IEEE PIMRC'20, IEEE Globecom'21, IEEE VTC'21 Fall, IEEE ICBC'21 and EUSIPCO'21.

Muhammad Ali Imran is a professor of Wireless Communication Systems with research interests in self organised networks, wireless networked control systems, and the wireless sensor systems. He heads the Communications, Sensing and Imaging CSI Research Group, University of Glasgow. He is an affiliate professor with the University of Oklahoma and a visiting professor with 5G Innovation Centre, University of Surrey, United Kingdom. He has more than 20 years of combined academic and industry experience with several leading roles in multi-million pounds funded projects. He has filed 15 patents; has authored/co-authored more than 400 journal and conference publications; was editor of three books and author of more than 20 book chapters; has successfully supervised more than 40 postgraduate students at doctoral level. He has been a consultant to international projects and local companies in the area of self-organised networks. He is a fellow of the IET and a senior fellow of the HEA.