Reliable Operational Services in MANETs by Misbehavior-Tolerant Quorum Systems

Elisa Mannes, Michele Nogueira, Aldri Santos
Wireless and Advanced Networks Team (NR2)
Federal University of Paraná, Brazil
Email: {elisam, michele, aldri}@inf.ufpr.br

Abstract—Operational services in MANETs, such as resource location and distribution of connectivity information, deal with node mobility and resource constraints to support applications. The reliability and availability of these services can be assured by data management approaches, as replication techniques using quorum systems. However, these systems are vulnerable to selfish and malicious nodes, that intentionally do not collaborate with replication operations or spread malicious data while participating in data replication. $QS^2$, a bio-inspired scheme to tolerate selfish and malicious nodes in replication operation of quorum systems can effectively detect misbehaving nodes in quorum systems, increasing data reliability. This paper evaluates the performance of $QS^2$ considering two realistic scenarios: the distribution of local market information in a downtown city and the distribution of traffic and schedule information in bus lines. Simulation results show that $QS^2$ improves the reliability of replicated data in MANETs and the amount of data compromised by misbehaving nodes inside the replication system is much lower than outdated data due to network topology.

I. INTRODUCTION

Advances in wireless communication technologies have reinforced the development and the use of different wireless networks towards the Future Internet [1]. Mobile ad hoc networks (MANETs) are a branch of those networks and intend to support applications on various areas, such as health care, transportation and entertainment, by a self-organized and distributed approach. MANETs comprise of mobile and portable devices (nodes) that own constrained resources and move into an area following a pattern. MANETs, as well as other wireless and wired networks, are envisioned to converge in order to support Future Internet services and applications. Despite the substantial impact on the society, Future Internet faces numerous challenges and new requirements, as the continuous offer of services despite failures, accidents, attacks or other network conditions [2]. In MANETs, nodes characteristics can easily lead the network to partition, making services unavailable and sustaining outdated information [3].

Towards the design of reliable services on network segments of the Future Internet, data management is essential. This work emphasizes the importance of managing data from different services, such as resource location and mobility information, in order to provide their reliability. Particularly, it focuses on services that support the good performance of applications, named operational services. These services are sensible to the absence or outdated information, being easily compromised. Data replication is broadly used to overcome issues originated by network partition and mobility, providing availability and fault-tolerance on data management. Among all data replication techniques proposed for MANETs, this work highlights quorum systems as an effective way for replication [4]. Quorum systems consist of sets (quorums) of nodes performing as servers. Each quorum must intersect the others in order to guarantee the observation of all read and write operations in the quorum system, being these operations performed in just one quorum. The main advantages of a quorum system compared with classic replication techniques are the low processing and communication costs, making these systems suitable for MANETs. A specific kind of quorum, the probabilistic quorum system, owns relevant characteristics to manage data in MANETs, such as the probabilistic construction of intersections, that reduces the consumption of resources and makes data replication more dynamic [4] [5].

However, probabilistic quorum systems for MANETs show security vulnerabilities that can be explored by selfish and malicious nodes. These nodes intend to induce low data reliability in replication operations [6]. Selfish nodes do not collaborate with data replication operations for saving their resources, whereas malicious nodes aim to disturb the system injecting false data or modifying existing ones. Thus, in order to sustain demanding applications and correctly manage operational services, quorum systems need to tolerate misbehaving nodes that interfere in their performance.

In general, existing solutions to overcome misbehaving nodes assume the existence of fixed infrastructure and reliable channels. However, MANETs can not guarantee these attributes. For MANETs, common approaches to avoid the interaction with misbehaving nodes bring the cost of sharing recommendations, implying in an extra overhead, or employ centralized entities [7] [8]. Hence, it is necessary to provide misbehavior tolerance in quorum systems, as well as a decentralized, autonomous and low cost solution against misbehaving nodes. These characteristics can be naturally found in biological systems, providing inspiration for designing efficient decentralized and autonomous solutions [1].

This work presents a reliability evaluation of the use of $QS^2$ (quorum systems, quorum sensing), a bio-inspired scheme we proposed in [9], to support service data management replication in realistic scenarios of MANETs. $QS^2$ is a bio-inspired scheme that provides tolerance to operational services...
through a probabilistic quorum system for MANETs facing misbehaving nodes in replication operations. The evaluation considers two realistic MANETs scenarios: the dissemination of commercial information in a downtown city and the distribution of bus lines information in a transportation environment. Simulation results show that \( QS^2 \) keeps data reliability above 55% facing data injection attacks. Further, we observed that the amount of compromised data is lower than the amount of outdated data, in both simulation scenarios. The reliability ensured by \( QS^2 \) meets requirements for non-inconsistency sensitive applications, i.e., those applications in which eventual inconsistencies are acceptable. Examples of applications supported by \( QS^2 \) are environment monitoring, warnings dissemination and traffic information distribution for vehicles and events information distribution for pedestrians.

This paper proceeds as follows. Section 2 reviews related works. Section 3 defines the system model and assumptions taken by the proposed scheme. Section 4 describes \( QS^2 \), its modules and components. Section 5 features performance and reliability results obtained by realistic scenarios simulations. Section 6 concludes the paper and presents future works.

II. Related Work

Classic data replication techniques [10] use predefined and static servers, requiring reliable delivery of packets and the ordering of replication messages. Misbehavior tolerance in these systems are based on the validation of operations by the majority of nodes. This strategy allows nodes to overpass the amount of misbehaving nodes in the network. However, it results in an overhead on the network, caused by the exchange of messages among nodes for ensuring correct operation of replication. MANETs can not ensure the basic requirements to guarantee fault tolerance in classic replication systems and, hence, these systems are not suitable for this kind of network.

Quorum system replication [11] decreases the use of computing and communication resources and are more suitable for dynamic environments as MANETs. The probabilistic choice of quorums decreases even more the use of resources [4]. However, despite the existence of probabilistic quorum systems that can tolerate misbehaving nodes [11], these systems need the same requirements from classic replication systems, such as reliable delivery of messages. Again, MANETs characteristics make it difficult to implement this kind of quorum to reliably support operational services.

Replication systems for MANETs address security employing misbehavior detection mechanisms, such as reputation systems or intrusion detection [12]. However, many of them use recommendation from other nodes to correctly detect misbehavior, what may be explored by misbehaving nodes. Approaches to detect false data injection [8], [13] are consolidated in sensor networks since their main focus is on data collection. The validation of data is performed by cryptography, by the majority agreement of a data or yet by using firewalls. Although the use of centralized entities are useful in some network topologies, they bring limitations in MANETs.

Despite existing misbehavior detection systems with characteristics as autonomy, decentralization and the efficient use of resources, no one of them consolidate all these characteristics in one solution nor are integrated with data replication techniques. Moreover, there is no solution that can mitigate selfish and malicious nodes together. Attributes like self-organization, autonomy and the use of few resources need to be incorporated in solutions for MANETs. These characteristics are found naturally in bio-inspired solutions such as routing protocols inspired by ant colony [14]. Following these ideas, we proposed in [9] a scheme inspired by biological systems to take advantage of characteristics offered by these systems.

III. System Model

We assume that the network model is composed by a set \( P \) of \( n \) nodes identified by \( \{d_0, d_1 \ldots d_{n-1}, d_n\} \). Every node \( d_i \in P \) has a unique physical address or identification and the same processor and energy capacity. Nodes communicate over a wireless link with a delimited transmission range and same value for all nodes. The communication between nodes is asynchronous, and the transmission time varies and is unknown. Also, the communication channel is not reliable and can discard packets due to collisions or lack of connectivity.

Nodes move accordingly to a movement pattern in a delimited area. Partitions in the network can occur due to movement of nodes, scarce resources, accidents, faults or other issues. Nodes rely on intermediate nodes to route packets since they may not reach all nodes directly due to their coverage area. We also assume that routing and lower layers are not affected by misbehaving nodes. Replicated data, represented by operational data, is managed to be transmitted into single network packets due to its small length. We consider that the network provides a signature scheme, like [15] [16], to authenticate important information sent by the quorum system.

A. Data management model

We assume a probabilistic quorum system, called PAN, as the data management model. Although other quorum systems could be applied, PAN [4] was chosen because of its approach of asymmetric quorums, reducing the number of replication messages. It is composed by a storage set (\( StS \)), nodes acting as servers, clients and agents. \( StS \) comprises of servers, i.e., nodes elected to replicate data. Server stores data and manage the replication between them. Clients request data from the \( StS \) sending requests to chosen servers. Servers contacted directly by clients are called agents, and are responsible for mediating the operation among clients and the \( StS \).

Quorums formed by the interaction among servers inside the \( StS \) provide read and write operations. Read operations create read quorums \( (Q_r) \) and write operations create write quorums \( (Q_w) \). Although clients partake issuing requests to servers, this work focuses on the \( StS \) when dealing with read and write operations, i.e. clients protocols are not considered.

As part of write operations, all servers have a buffer that stores the latest version of each data. In regular intervals, nodes forward data from the buffer to the \( StS \) using a gossip-based...
proposed protocol [17]. After some time, it is expected that all servers have the most recent data value, and the write quorum will comprise all servers that effectively receive the update.

B. Attack model

In the attack model, misbehaving nodes compromise the data management by affecting data availability and integrity in a quorum system [6]. These nodes are intruders, and they know how the network works. We also assume that these nodes can have valid cryptographic keys and permission to join network operations. Misbehaving nodes can be either selfish or malicious, or even exhibit both behaviors. Selfish nodes do not collaborate with replication operations, representing the lack of cooperation attack. Malicious nodes can delay data forward in write operations, characterizing the timeout manipulation attack, while malicious nodes modifying or injecting malicious data in the replication system are performing the data injection attack. Focusing on the lack of cooperation attack is important since these nodes can be easily found in MANETs, as nodes that aims to save resources for their own use or to damage the network. The timeout manipulation and the data injection attacks are identified as damaging attacks for data replication operations, and it is important to manage the replication to be effective and reliable [18].

IV. OVERVIEW OF QS2 SCHEME

This section describes the QS2 (quorum sensing + quorum system) scheme proposed in [9] to support quorum systems for MANETs facing misbehaving nodes, helping the maintenance of data reliability. By monitoring all replication operations and selecting nodes that act collaboratively, QS2 denies the participation of misbehaving nodes in replication operations. The monitoring is performed by the observation of data operations. The QS2 scheme is composed of two modules: the node monitoring and the cooperation decision.

The node monitoring module is responsible for classifying nodes as reliable or misbehaving. This module is divided into two components: behavior monitoring and node classification. The behavior monitoring component corresponds to the autoinducer counting from bacteria and quantifies the autoinducers (data write and data forward operations) sent by nodes in the system. According to the autoinducer counting and autoinducer threshold equivalent to a reliable node behavior, the node classification component categorizes nodes in one of the three classes: reliable, selfish or malicious.

The cooperation decision module determines the cooperation relationship between two nodes and selects those that are more likely to help. Nodes are chosen by the node selection component based on the node monitoring module. This module also allows nodes to tune the interaction between them. Together, node monitoring and cooperation decision modules choose suitable nodes to participate in quorums.

V. CASE STUDIES

Since we evaluate the performance of QS2 in [9]. in this work we employ QS2 to support operational services in two realistic MANETs scenarios. We aim to comprise different MANETs’ scenarios regarding speed, density and pause time behaviors. Also, we use real data patterns for node’s movement in order to show the viability to use our system in a real scenarios. In the first scenario, we use QS2 in MANETs deployed in downtown cities to distribute local market information among users nearby. In the second scenario, QS2 is used to support replication in a MANET deployed in bus lines to distributed traffic and schedule information among users.

Both scenarios and QS2 scheme are implemented and evaluated by Network Simulator (NS) version 2.33. We modified the PAN implementation in NS to include QS2. The scheme is analyzed considering the presence of misbehaving nodes issuing lack of cooperation, timeout manipulation and data injection attacks in read and write operations. Selfish nodes act in lack of cooperation attacks by not collaborating with data replication operations on quorum systems. The timeout manipulation attack is characterized by malicious nodes delaying data forwards and malicious nodes inject false data during read and write operations in data injection attacks.

In order to evaluate the use of QS2 in realistic scenarios, we employ four metrics: reliability degree (Rd), detection ratio (TDx) and outdated data ratio (TDx) out. Rd assesses the performance of QS2 while TXf and TXf ratios are used to quantify the efficiency of QS2. We also employ a set A comprising all misbehaving interactions and a set B representing reliable interactions in read and write operations. These sets are composed of a tuple (d, a), where d represents the class of a node detected by QS2 and a is the real class of that node.

The reliability degree (Rd) quantifies the performance of QS2 and represents the amount of correct reads obtained by read operations. Correct readings are those that retrieve data corresponding to a previously performed write operation or to a writing that is still in progress. Rd is defined by Eq. 1, where CT represents reads with correct results, and R is the total amount of reading requests issued by clients.

\[ Rd = \frac{\sum CT}{R} \]  

Detection ratio (TDx) represents the amount of detected misbehaving nodes. It is accounted for the lack of cooperation and data injection attacks, and it is calculated by Eq. 2.

\[ TDx = \sum D_i \forall i \in A, \text{ where } D_i = \begin{cases} 1 & \text{if } d_i = a_i \\ 0 & \text{if } d_i \neq a_i \end{cases} \]  

False and outdated data (TXmis and TXout) are used to quantify the amount of false and outdated data replied to clients. Eq. 3 quantifies the amount of false data replied to clients, where CW represents the amount of false data replied by nodes from the SS in read operations and R represents the total amount of issued reading in the quorum system.

\[ TXmis = \frac{\sum CW}{R} \]  

Equation 4 calculates the amount of outdated data, where Rd represents the amount of reliable data and TXmis represents the amount of false data.
To assess these metrics, we assume the existence of an oracle observing and collecting the class of nodes and detection occurrence. However, this oracle is not part of the solution.

A. Urban environment - downtown city

The first realistic scenario corresponds to a classical downtown city environment, like the ones found in Central Europe. It is based on [19], and consists of pedestrians and cyclists moving toward common places, such as shopping centers and theaters. In [19], authors proposed a network to share information between restaurants, shops, bus systems and mobile devices owned by users. It is supposed that the information shared concern about users interest, like discounts, dish of the day, traffic condition, etc. In order to receive messages, users are invited to subscribe to the service, and messages are received while the user is within the range of other devices. For this purpose, it is necessary to secure operational services, such as resource location and mobility management services, to avoid that malicious devices compromise the network.

From [19] we borrow movement traces and nodes behavior. The downtown city scenario consists of points of interest, such as churches, universities and restaurants, interconnected by streets. People follow the streets toward one point of interest, following the graph walk model. This model is different from the commonly used random waypoint model in the way that instead of follow streets to reach a point of interest, in random waypoint model nodes follow a straight line.

In order to evaluate the impact of network density, we considered scenarios composed by 50, 100 and 150 nodes, with maximum speed varying from 3 to 5 km/h. The area of 2462 x 1733m is divided by 75 points of interest interconnected by 116 streets. The pause time in points of interest is considered between 12 and 20 minutes, representing pedestrians stopping in central stations or shopping centers. We used AODV as routing protocol and results are the average of 35 simulations of 1500 seconds, with a confidence interval of 95%.

Read quorums \((Q_r)\) are composed by 5 servers, including the agent, and write quorums \((Q_w)\) are composed by all nodes that participate in the dissemination of some data. Each node disseminates data for 4 servers, every \(T = 200\text{ms}\). The storage set (StS) is composed by 25 nodes, randomly chosen. The writing and reading intervals are modeled following a Poisson distribution, as found in [4], with the expected writing rate \(\lambda = 100\) and \(\lambda = 36\) for reading requests. The writing threshold for each node is \(k_{env}^{w}_{\text{max}} = 0.018\) writes per second, and the minimum forward threshold is \(k_{env}^{w}_{\text{min}} = 0.15\) per second.

Attacks are divided into two scenarios: the one with data injection attacks and the one with lack of cooperation, timeout manipulation and data injection attacks together. Scenarios with data injection attacks are configured with 5, 7 and 9 misbehaving nodes, representing 20%, 28% and 36% of the nodes in StS. Scenarios with all attacks considered 5, 10 e 15 misbehaving nodes, representing 20%, 40% and 60%. These scenarios apply lack of cooperation in write and read operation, timeout manipulation \((T = 3000\text{ms})\) and data injection in write and read operations. Each attack is performed by 20% of the total amount of misbehaving nodes.

1) Reliability degree: First, we address the performance of PAN in normal conditions, as presented in Fig. 1(a), where PAN provides a data reliability above 80%. However, when the network faces data injection attacks (Fig. 1(b)) and all attacks (Fig. 2(a)), the reliability is lower than 10%. This analysis is consistent with previous studies, and shows that it is not possible to support operational services facing misbehaving nodes without security mechanisms. Applying QS\(^2\) to tolerate misbehaving nodes does not affect PAN performance when it is not under attack, but provide an increase in the reliability when facing them. Fig. 1(c) presents results obtained by QS\(^2\) facing false data injection attacks. It is possible to observe that \(R_d\) is not changed as the density of nodes increase, considering the same amount of misbehaving nodes. Scenarios with all
attacks together, shown in Fig. 2(b), have the same behavior. The mix of different kinds of attacks results in greater values for $R_d$. This is because the misbehaving nodes includes the lack of cooperation and timeout manipulation, that alleviate the impact on data injection attackers.

2) Efficiency: Fig. 1(d) shows that the detection ratio obtained is greater than previous simulations. $T_{x_{det}}$ is greater than 90%, and increases together with the amount of misbehaving nodes. However, this behavior is not observed in scenarios with all attacks together, described in Fig. 2(c), where $T_{x_{det}}$ varies between 85% and 90%. This is mainly because $QS^2$ does not employ specific autoinducers for timeout manipulation attacks and the detection of attackers is delayed.

3) False and outdated data: Due to the discrepancy between higher $T_{x_{det}}$ and lower $R_d$, we suspected that the reason why $R_d$ is lower in such scenarios would not be the action of misbehaving nodes or the lack of efficacy of $QS^2$, but a normal loss due to network characteristics. In order to verify this assumption, we quantified the ratios of false data ($T_{x_{mis}}$) and outdated data ($T_{x_{out}}$) in reading operations. For both attacks, the assumption proved to be true, as seen in Fig. 1(e) and Fig. 1(f). While $T_{x_{mis}}$ grows about 1% with the increase of the number of nodes in the system, $T_{x_{out}}$ grows 10% on average. This behavior is repeated in scenarios with all the attacks, as shown in Fig. 2(d) and 2(e). This is an evidence that lower $R_d$ in these scenarios happens partly due to network topology, as the size of the movement area and density of nodes in the network. Moreover, $QS^2$ obtained good detection ratios and $R_d$, the latter being superior to that achieved by PAN without the use of $QS^2$ in validation scenarios.

B. Transportation environment - metro bus system

This scenario is based on the implementation provided by [20]. The scenario is deployed in a transportation environment, supported by a routing architecture called Ad hoc City. It is used to distribute information among vehicles, users and public and private transportation. Users can participate using wireless devices such as notebooks, mobile phones and tablets. In this kind of scenario, mobility and lack of infrastructure are evident and must be managed to guarantee proper message delivery application for vehicles and pedestrians. The use of quorum systems to provide fault tolerance also needs to be robust in order to support the proper functioning of the network.

The transportation scenario is characterized by a backbone consisting of buses and delivery vehicles, which covers a specific area. The movement of buses were obtained by [20] through the observation of buses during the morning and the afternoon for two weeks. Originally in [20], there are approximately 850 buses moving in an area of 5100 km$^2$. Although the scenario was adapted to 150 nodes distributed in 1500x2000 meters, we kept the real movement patterns. This reduction is justified since we aim at evaluating $QS^2$ under high speed with a support infrastructure and real movement traces. Nodes move between 0 and 90 km/h, and are supported by 8 fixed stations distributed proportionally, according to [20].

Given the network topology, some parameters related to PAN quorum system have been adapted. Read quorums ($Q_r$) consists of four servers, including the agent, and write quorums ($Q_w$) are formed by all nodes that receive a data writing. Data is disseminated to four servers, every $T = 200ms$. The storage set (StS) consists of 30 nodes, chosen randomly. The interval of writes and reads is modeled following the Poisson distribution with $\lambda = 100$ for data writes and $\lambda = 36$ for data reads, and is given in seconds. The maximum rate of write operations per node is equal to $k_{env_{max}} = 0.018$ writes per second, and forwarding rate must exceed $k_{enc_{min}} = 0.15$ packets per second. The used pattern of movement refers to an interval of fifteen minutes from the records obtained by [20]. To simulate the routine of the buses in the morning, it was considered the records obtained from 07:15 to 07:30, and for the afternoon, we used the records obtained between 17:15 and 17:30. We used AODV as the routing protocol and the results are the average of 35 simulations of 900 seconds, with a confidence interval of 95%. The attack scenarios follow he
same description as in the previous downtown scenarios. We show results related to scenarios in the morning since results obtained with scenarios in the afternoon are much similar.

1) Reliability degree: The results of PAN quorum system performance in normal conditions shows a reliability greater than 85%. As the previous scenarios evaluations, the reliability in the system without using QS2 and facing misbehaving nodes is lower than 10%. This applies to scenarios with data injection attacks and with all attacks, as shown in Fig. 3(a). The reliability achieved by using QS2 is over 60%, as shown in Fig. 3(b). We can also note that scenarios with data injection attacks and all attacks are similar in the reliability results obtained, contrary to the results obtained with the validation scenario, in which scenarios with all attacks the Rd is superior to 92%. This may be due to the pattern of motion of the nodes, whose speed varies irregularly. Because of the regular bus stops to pick up passengers, these vehicles have an irregular acceleration, which can cause loss of data packets in the wireless medium. Moreover, the delay of these vehicles also have a large variation, and as previously seen, prolonged pause times do not favor the spread of autoinducers and their respective accounting for all network nodes. Also, it is worth noticing that data delivery rates can be improved by changing some network parameters unrelated to our solution: increasing network density or spreading more fixed station could improve data delivery rates and consequently our solution would be able to provide better reliability.

2) Efficiency: The detection efficiency (Tdet) of QS2 reflects the results of Rd obtained. Fig. 3(c) shows that while Tdet for scenarios with data injection attacks is greater than 65% and grows as the number of misbehaving nodes increases, Tdet for scenarios with all attacks is less than 80%. The detection efficiency of misbehaving nodes with all attacks is lower due to the lack of autoinducers to correctly represent the timeout manipulation attack, and due to the delay for the detection of selfish nodes. Moreover, due to this specific scenario, nodes can face difficulties delivering packets, and thus the detection of nodes may delay. This probably is the cause of lower rates in the detection of misbehaving nodes.

3) False and outdated data: Similar to previous scenario, the ratios of false data and outdated data show that the amount of false data obtained by nodes through readings in the quorum system is less than the amount of outdated data. Figure 3(d) presents Txdet and Figure 3(e) shows the results for Tdet. We observe that in both scenarios, approximately 50% of the data obtained by nodes are discarded because they are not updated, and approximately only 10% are false data, injected by malicious nodes. These scenarios also show a low percentage of false data obtained by nodes and becomes evident that the amount of outdated data returned by read operations are the majority.

VI. Conclusion

Due to the imminent importance of Future Internet applications, it is crucial to provide mechanisms to guarantee the continuous offer of services despite failures, accidents, attacks or other network conditions. In MANETs, a network segment for wireless communication in Future Internet, nodes characteristics can easily lead the network to partition, making services unavailable and sustaining outdated information. MANETs also need to observe misbehaving entities looking forward to disrupt operational services. This work evaluates QS2, a scheme to tolerate misbehaving nodes in quorum systems supporting operational services in realistic MANETs scenarios. Results show that QS2 allows an increase in the reliability of a quorum system for MANETs facing misbehaving attacks in read and write operations and QS2 detected misbehaving nodes with a high efficiency. The reliability achieved by the use of QS2 together with a probabilistic quorum system provides a secure way to support operational services in which eventual inconsistencies are acceptable. As future works, we intend to characterize operational services data traffic in MANETs to use reliable autoinducers thresholds and to validate QS2 using the real operational data settings.
REFERENCES


