Journal Pre-proof

A probabilistic comparison-based fault diagnosis for hybrid faults in mobile networks

Hazim Jarrah, Peter H.J. Chong, Chris Rapson, Nurul I. Sarkar, Jairo Gutierrez

PII: S0140-3664(19)30744-3
DOI: https://doi.org/10.1016/j.comcom.2020.03.042
Reference: COMCOM 6328

To appear in: *Computer Communications*

Received date : 7 July 2019
Revised date : 29 January 2020
Accepted date : 24 March 2020

Please cite this article as: H. Jarrah, P.H.J. Chong, C. Rapson et al., A probabilistic comparison-based fault diagnosis for hybrid faults in mobile networks, *Computer Communications* (2020), doi: https://doi.org/10.1016/j.comcom.2020.03.042.

A Probabilistic Comparison-Based Fault Diagnosis for Hybrid Faults in Mobile Networks

Hazim Jarrah*, Peter H. J. Chong*, Chris Rapson*, Nurul I. Sarkar†, Jairo Gutierrez†

*Department of Electrical and Electronic Engineering**
Department of Information Technology and Software Engineering†
Auckland University of Technology, New Zealand
peter.chong@aut.ac.nz

Abstract

Fault diagnosis has always been vital to providing a high level of dependability in systems. The comparison approach is one of the most prevalent diagnosis techniques that offers a simple and yet practical way to identify faulty nodes in a system. Even though several comparison-based diagnostic models have already been introduced, the majority of them only diagnose permanent faults in static networks. Nowadays, intermittent faults and dynamic systems are more challenging to diagnose and become more common. This paper, first, proposes a novel comparison-based diagnostic model that deals with hybrid fault model in mobile networks. Both the diagnosable systems and faults under the proposed model have been characterised. Second, this paper proposes an efficient fault diagnosis protocol for hybrid faults in mobile networks. The proposed protocol employs a network coding technique to exchange the diagnosis messages so that it can provide a correct diagnosis with a higher probability of completeness. The correctness and complexity proofs of the proposed protocol are presented, and they show the viability of the proposed diagnostic model and protocol for hybrid faults in mobile networks. Besides, we study and analyse the performance of the proposed protocol under various fault and system parameters using OMNeT++ simulation. The simulation results show that our protocol can diagnose hybrid faults in mobile networks with high accuracy and less overhead.

Keywords: *Dependability; Fault Diagnosis; Intermittent Fault; Mobile Networks; Network Coding*

1. Introduction

Wireless and mobile networks include mobile ad-hoc networks (MANETs), vehicular ad hoc networks (VANETs), wireless sensor networks (WSNs), and wireless mesh networks (WMNs) [1-3]. These networks provide ubiquitous connectivity and services where the deployment of fixed networks is not an option [4, 5]. There are both economic and environmental reasons for the constant increase in demands on mobile networks [6-10]. Due to their ever-changing nature and harsh deployment situations, these networks are highly vulnerable to faults that may incapacitate the components and the operations of a network [11-13]. In other words, faults threaten the dependability of networks and their ability to deliver the required services [14, 15]. As such, researchers have investigated several techniques to improve dependability. In particular, the problem of automatic fault identification has inspired the development of system-level fault diagnosis theory [16] and various diagnosis models [17, 18]. The diagnosis problem aims to form a consensus among faultless nodes about the fault status of nodes in a system and therefore sustain the system's performance and integrity. This consensus on faulty nodes is of high interest for protocols and services being run over distributed systems.

Comparison-based models are the most practical diagnosis models, and they have been used widely in many systems [19, 20]. They enable nodes to diagnose a system on their own and without explicit testing. In brief, these models, at their core, involve assigning tasks to nodes, and the resulting outputs of the same assigned task are compared. A set of comparison outcomes is called a syndrome, which is analysed to identify faulty nodes. The underlying logic of this approach is that faultless nodes performing the same task produce identical correct results. Many comparison-based models have been developed to meet the diagnosis requirement of various systems [21-26]. However, diagnosis models designed for mobile networks are challenging and therefore rare. In general, existing comparison-based models could be categorized into deterministic and probabilistic models. The former provides a correct and complete diagnosis, whereas the latter provides a correct but incomplete diagnosis. Correct diagnosis guarantees that no faultless nodes be classified wrongly as faulty. Complete diagnosis assures identifying all faulty nodes. Deterministic models impose rigid assumptions on fault behaviours and the number of faulty components in a system. On the other hand, probabilistic models consider assumptions that are more realistic, such as no upper bound on the number of faulty nodes, no strings on fault occurrence time, and nodes could be faulty or faultless with probability. It is clear that probabilistic comparison-based models are more appropriate for mobile networks where faults and topology are unpredictable. To the best of our knowledge, our research in this paper introduces the first probabilistic comparison-based model for fault diagnosis in mobile networks.

In this paper, we first propose a probabilistic comparison-based diagnostic model considering the design requirements of mobile networks; particularly, their intrinsic characteristics such as unsteady topologies, asynchronous communications, and limited knowledge about the network. Besides, the proposed diagnostic model supports a more realistic hybrid fault model. That is, not only are permanent faults allowable in the system under diagnosis, but so, too, are intermittent faults. The intermittent faults manifest temporarily and at random intervals. In other words, those intermittent faults might be inactive sometime during the diagnosis process; hence, they might be undetectable at that time. Due to their intermittency and random duration, the identification of intermittent faults has a varying degree of certainty. The proposed diagnostic model promises a correct diagnosis with a high probability of completeness. Second, this paper introduces a fault diagnosis protocol for hybrid faults in mobile networks. The proposed protocol implements our proposed probabilistic diagnosis model. This protocol employs a probabilistic broadcasting mechanism to propagate tasks over the network. Also, it leverages a network coding technique to exchange nodes' results of the tests. The protocol can identify both permanent and intermittent faults with higher probability. The proof of correctness and the analysis of the complexity of the protocol are provided. We also study its performance by an extensive set of simulations to evaluate the efficiency of the protocol under different scenarios. Furthermore, the performance of our protocol is compared with other related protocols whenever it is applicable. The results show that our proposed protocol performs better in terms of communication and time complexity.

The remainder of the paper is organised as follows. Section 2 presents an overview of the current comparison-based diagnosis models and highlights the key features of our proposed diagnostic model as compared to them. Section 3 describes the proposed probabilistic comparison model, including the diagnosable systems and faults. Section 4 presents a self-diagnosis protocol for mobile networks and demonstrates its correctness proofs as well as its complexities analysis. Section 5 shows the simulation results and comparative analysis. Section 6 discusses the findings and their implications and concludes the paper.

2. Related Works

Diverse comparison-based models have been proposed for fault diagnosis in various systems. This section investigates existing diagnosis models that employ comparison approaches to identify faulty nodes. Mainly, it focuses on models that have potentials in wireless mobile networks and on models that consider hybrid faults. It provides critical insights into their assumptions and limitations. Besides, it highlights the promising features of our proposed model for diagnosing hybrid faults in mobile networks.

Existing comparison-based diagnosis models can be broadly categorized into deterministic or probabilistic models. The former models identify correctly and completely the set of faults in the system, whereas the latter ones offer a correct diagnosis with a high probability of completeness. The deterministic models, however, impose rigorous requirements on the system's structure and faulty nodes' behaviour, and that limits their usefulness and hinders their scalability. The broadcast comparison model [21] offers a deterministic diagnosis service for multicomputer systems. It involves sending a task to two distinct nodes, and their responses are broadcast to every node in the system using an underlying weak reliable broadcast protocol. Every faultless node compares their outputs, and then, the comparison outcomes are used to identify faulty nodes in a distributed manner. In [22], Chessa and Santi introduced a deterministic comparison-based model for wireless ad-hoc networks. This model takes advantage of broadcast communications in wireless networks to enhance the diagnosis process. Specifically, once a node sends a task, all neighbour nodes receive and execute the task. Also, once a node sends a tasks'result, each neighbour node receives the result. Every node then forms a local diagnosis of its neighbouring nodes. The Chessa and Santi model considers a fixed wireless network with no change of the network topology**.** In addition, it suffers from high message redundancy. Chessa and Santi developed the Static-DSDP diagnosis protocol that can provide a correct and complete diagnosis for fixed wireless ad hoc networks [22]**.** In the Static-DSDP protocol, nodes use Chessa and Santi model to maintain a local view about the status of neighbouring nodes. Next, they exchange their views so that each fault-free node maintains a global perspective of the network. To overcome the shortcomings of the Chessa and Santi model, Elhadef et al. [23] proposed a time-varying comparison model that sends the task and the response together and hence, mobile nodes could be diagnosed not only by the tester nodes but also by any node. In addition, instead of replying to every task, nodes only respond to a limited number of tasks that is more than the minimum connectivity of the network. A Mobile-DSDP protocol is proposed in [22] for mobile networks where nodes generate their local views and then exchange their views to form a global perspective about the network. These models rely on timers to identify faults, and they implicitly assume that a system is synchronous. In [24], Jarrah et al. proposed a time-free comparison model that can diagnose mobile networks with more realistic assumptions. Unlike other models, the time-free comparison model uses no timers and requires no synchronisation. An RLNC-DSDP protocol is proposed to employ the time-free comparison model for diagnosing the status of adjacent nodes and then propagating the local views using a random linear network coding (RLNC) technique. These deterministic models consider permanent faults only and require upper limits on the number of faulty nodes in the system.

Dahbura et al. [27] introduced the first probabilistic comparison-based model. This model assumes that a faulty node produces correct outputs with a probability, p. In addition, it assumes that the number of possible incorrect outputs, m, is extremely large. Therefore, the output of a faulty node matches the output of a fault-free one with probability, p. Comparing the outputs of two faulty nodes, however, produces a match with probability, p^2. In [28], the (p, k)-probabilistic model utilizes tasks with k possible incorrect outputs to identify the faulty nodes. So, the output of a faulty node matches the output of a fault-free node with a probability, $q = 1/k$. This model suits tasks with known possible outputs. In [29], Rangarajan and Fussel proposed a probabilistic model that considers the same assumptions in [27]. However, it runs tasks on a small set of nodes, and no global syndrome is analysed. In other words, the diagnosis process is executed for local nodes. These probabilistic models consider hybrid faults, permanent and intermittent faults, and they impose no limit on the number of faulty nodes. However, they communicate using a one-to-one communication paradigm since they proposed for multiprocessor systems and wired networks. Hence, they are not suitable for wireless and mobile networks. To the best of our knowledge, there is no probabilistic comparison-based model proposed for mobile networks.

In [30], the authors proposed a self-diagnosis protocol for hybrid faults in MANETs. This protocol repeatedly employs the Chessa and Santi comparison model to identify nodes experiencing permanent and intermittent faults. It also maintains a spanning tree connecting nodes in a MANET. This protocol has several shortcomings that lower its efficiency in mobile networks. Particularly, it assumes that the communications are synchronous and reliable. Also, maintaining a spanning tree in mobile networks seems unrealistic. Many diagnosis protocols have proposed for other wireless and mobile networks such as WSNs [31], WMNs [32], and VANETs [33]. However, these protocols have several limitations, such as assuming the existence of infrastructure in a network and synchronous communications, and hence, their usefulness is not guaranteed in infrastructure-less mobile environments.

This paper introduces a novel probabilistic diagnostic model for hybrid fault diagnosis in mobile networks. The proposed model takes account of more realistic assumptions on diagnosable faults and systems. It diagnoses hybrid faults where nodes could be faulty with a probability. In addition, it imposes no upper bound on the number of faulty nodes in a system, and it requires neither synchronisation nor fixed network topologies. These assumptions suit mobile networks where intermittent and dynamic faults are expected to occur frequently. The proposed model, also, employs a network coding technique to exchange the diagnosis messages, and that reduces significantly the messages required to diagnose a system. Table 1 presents the main features of the proposed model against current models.

3. A Probabilistic Comparison-Based Fault Diagnosis

3.1 System Model

This paper considers mobile networks that are dynamic topology systems consisting of an infinite number of mobile nodes. Each run contains a finite set of nodes, $\prod = \{v_1, v_2, \dots, v_n\}$, where v_i is node i and n is the number of nodes. Nodes may enter and depart the system without restraint. The system is asynchronous in a sense that no upper bounds are assumed on node speed, transmission delay and computation time. Nodes communicate via wireless links. Each node is assumed to have a matchless identifier, ID. Further, it is assumed that there exists a link layer protocol that provides essential services for nodes communications such as contention resolution, one-hop reliable broadcasting, and identifying nodes. Some examples of such link layer protocols can be found in [1, 34, 35]. A typical mobile network model is shown in Fig. 1.

Table 1. Comparison among diagnostic models.

Protocol	Deterministic /probabilistic	Fault type	Number of faults	Topology	Global vs local	Communication Paradigm
[27]	Probabilistic	Intermittent	No upper bound	Static	Global	One-to-one
[28]	Probabilistic	Intermittent	No upper bound	Static	Global	One-to-one
[29]	Probabilistic	Intermittent	No upper bound	Static	Local	One-to-one
[21]	Deterministic	Permanent	Upper bound	Static	Global	One-to-one
[22]	Deterministic	Permanent	Upper bound	Static	Global	One-to-many
[23]	Deterministic	Permanent	Upper bound	Dynamic	Global	One-to-many
[24]	Deterministic	Permanent	Upper bound	Dynamic	Global	One-to-many
This paper	***Probabilistic***	***Intermittent &Permanent***	***No upper bound***	***Dynamic***	***Global***	***One-to-many***

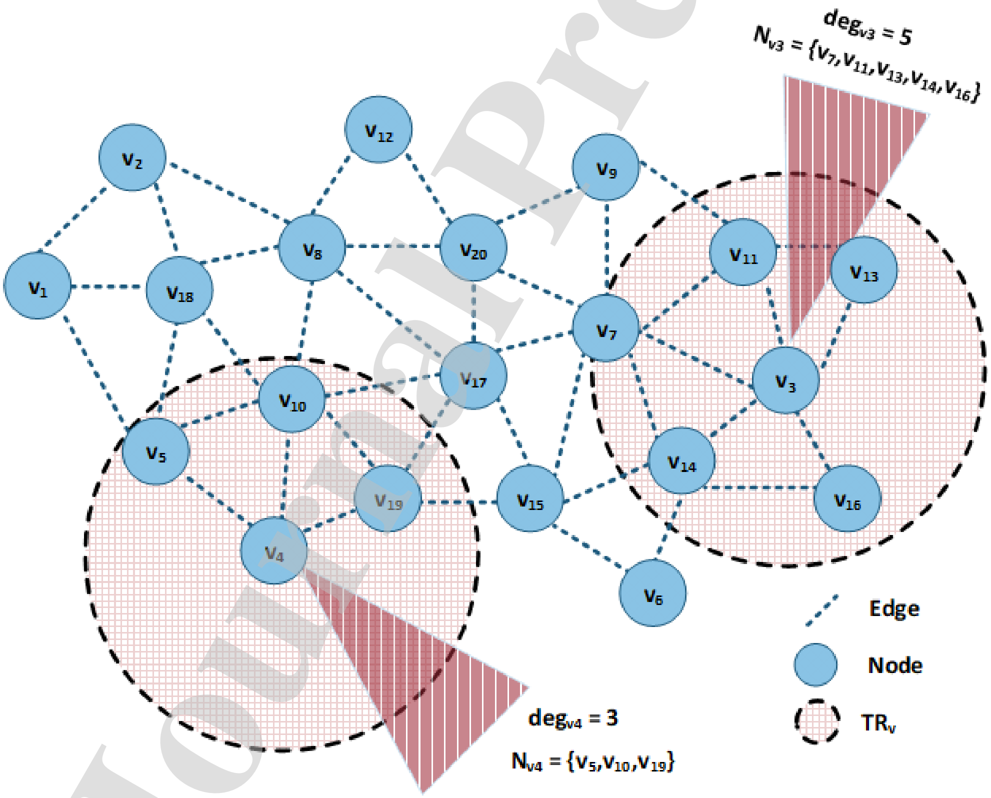

Figure 1. A mobile network model.

The topology of a system is represented by a communication graph, $G_t = (V, L_t)$, where $V = \prod$ is the set of mobile nodes, and $L_t \subseteq V \times V$ is the set of links at time t. TR_v represents the transmission range of each node $v \in V$. There is a link $(u, v) \in L_t$ for every two nodes $u, v \in V$ if u is within the TR_v at time t. N_v represents the set of neighbour nodes within the transmission range, TR_v, at time t. G_t is undirected, and hence $(u, v) \in L_t$ if $(v, u) \in L_t$. The degree of a node v, $deg_v = |L_v|$. Neighbour nodes may change over time since their movement is unrestrained. However, each node knows the identities of its neighbours at any time t.

3.2 Fault Model

In this subsection, we define the fault model considered in this research. Mainly, we describe diagnosable faults that may manifest in the system in terms of the fault's type and behaviour.

Our research focuses on identifying faulty nodes in mobile networks. Faultless nodes exhibit the correct responses to the same stimulus, whereas faulty nodes manifest various responses according to the faults they undergo. This fault model considers faults from different perspectives as follows. First, from a fault duration perspective, this research assumes that nodes are subject to both permanent and intermittent faults. While a permanent fault demands external intervention for repair or removal, an intermittent fault appears and disappears frequently and unpredictably. Second, based on the behaviour of a faulty node, this research considers soft faults that do not interrupt the communications with other nodes, such as omission and timing faults. Hard faults, such as crash faults, which prevent communication with the rest of the system, are excluded in this paper. Third, based on fault occurrence time, a fault could be static or dynamic. A dynamic fault occurs during the diagnosis session, whereas a static fault exists before the session and lasts to the end of the session. Here, both static and dynamic faults are considered. Nodes undergoing permanent and static faults operate all the time incorrectly, whereas nodes experiencing intermittent and dynamic faults may operate correctly sometimes and fail at other times. We model the behaviour of a node as a Bernoulli trial as follows. Let p_{vj} be the probability of a faulty node, v, operates correctly at the time, j. Hence, $1 - p_{vj}$ is the probability that v operates incorrectly at the same time, j. Let p_v denote the average values of probabilities over all times. For simplicity, it is assumed that this probability, denoted by p_v, is the same for all times, and all nodes. This assumption could be easily extended. In addition, it is assumed that faults in different nodes occur independently. Dependent faults and malicious faults are beyond the scope of this paper and a matter of future research.

While the nodes in these networks use wireless links to communicate, and these links, by nature, are lossy, we assume that messages are reliably delivered. This assumption can be achieved by an underlying data link layer protocol that can provide a reliable broadcast even in the presence of unpredictable behaviours as described above in the system model. It is also noteworthy that the proposed diagnostic model in the next subsection can tolerate delays that may be caused by such data link layer protocol because it uses no timers and requires no synchronicity.

It is assumed further that there is no upper bound on the number of faults since hard faults are excluded. Soft faults, on the other hand, are included, but they have no impact on the system's connectivity.

3.3 Diagnostic Model

This subsection introduces a comparison-based diagnostic model for fault diagnosis in mobile networks. The proposed model respects the hybrid fault model characterised in the previous subsection; hence, it relaxes stringent assumptions imposed by previous models. Also, it employs the comparison approach to identify the set of faulty nodes; that is, the agreements and disagreements of output results obtained by distinct nodes for identical tasks are the basis for identifying the state of nodes. In this model, each faultless node carries out all the comparisons in the system. The fault identification process is performed in multiple rounds to detect dynamic and intermittent faults with a high probability. In addition, this model leverages network coding to deal with the high communication overhead caused by multiple testing rounds.

This model utilises a comparison approach to identify faulty nodes. That is, instead of nodes testing one another, a task is assigned to two distinct nodes, and their results are compared. We assume that tasks can only detect faults existing at the testing time in a testee node. It is assumed that the number of possible incorrect results is huge, and faulty nodes produce random and independent results. The design of these tasks is beyond the scope of this research since it may vary depending on the system under consideration. The outcome of the comparison is **0** if the results are matched, and **1** otherwise. In our model, the same task is assigned to every node in the system, and this strategy eliminates the overhead that may be caused by generating different tasks at each node. Using the same task enables nodes to perform the diagnosis process in a fully distributed fashion. In addition, this strategy maintains consistency, exposing nodes to the same task. Given that, nodes may be faulty with different probability due to considering intermittent faults; this model governs the generation of comparison outcomes in nodes based on the following assumptions.

A1: A faultless node comparing its output with a faultless node's output always produces a match outcome, **0**.

A2: A faultless node comparing its output with a faulty node's output produces a match outcome, **0** with a probability, p, and a matchless outcome, **1** with a probability $1-p$.

A3: A faulty node comparing its output with a faultless node's output produces a match outcome, **0** with a probability, p, and a matchless outcome, **1** with a probability $1-p$.

A4: A faulty node comparing its output with a faulty node's output produces a match outcome with a probability, P_b:

$$P_b = p^2 + \frac{(1-p)^2}{m} \tag{1}$$

where m represents the number of possible incorrect outputs. We assume m is extremely large, and hence $P_b \approx p^2$. In other words, P_b is insignificant even though it may occur in the case where both produce the correct output for the same task.

The **A1** assumption states that faultless nodes behave consistently and run as expected. As such, their outputs are always identical; hence, the comparison outcome is **0**. In this sense, faultless nodes are continually diagnosed correctly by other faultless nodes. On the other hand, the assumption **A4** states that faulty nodes misbehave repeatedly, and their outputs are different. Therefore, the outcome of comparing two faulty nodes is **1**. When one of the nodes under comparison is faulty, and another one is fault-free, the assumptions **A2** and **A3** state that there is a probability, p, that the faulty node is acting correctly at that time; hence, both nodes generate identical results. Therefore, faulty nodes may be undetectable at that time.

To overcome this issue, our model uses multiple rounds to improve the probability of identifying nodes experiencing such hybrid faults. That is, a set of tasks, $J = \{j_1, \ j_2, \dots \ j_r\}$, is to be executed by nodes in a system up to r rounds. Clearly, the complete diagnosis, where all faults are identified, is still not guaranteed in a finite time. That is, a node is diagnosed as a faultless node if it performs all the tasks correctly. This might happen with a probability, p^r. On the other hand, a faulty node is identified correctly if it performs at least one task incorrectly and this happens with a probability, $1-p^r$.

The probability, λ, of a faultless node producing a mismatch with a faulty node is equal to the probability of producing identical outputs for the previous $i-1$ tasks multiplied by the probability of producing different outputs for the i^{th} task as follows.

$$\lambda = p^{i-1}.(1-p) \tag{2}$$

The minimum number, r_{min}, of tasks that a faulty node should perform to be detectable with an accuracy, $c = 1 - p^r$, is

$$r_{min} = \left\lceil \frac{ln\,(1-c)}{ln\,(p)} \right\rceil \tag{3}$$

This formula indicates that the detection accuracy of faulty nodes is proportional to the probability of a node being faultless. In addition, the detection accuracy increases with the number of rounds. Fig. 2 shows the values of r_{min} for different values of p and c. It is clear that the larger the value of p, the larger the value of r. In other words, the identification of a faulty node that produces correct results with high probability requires a relatively high number of rounds.

This model aims to reach a consensus about comparison outcomes among faultless nodes; hence, each node takes the responsibility of diagnosing the system based on a syndrome generated. Therefore, the following assumptions are considered at each round.

A5: Nodes cooperatively distribute tasks so that every node receives a task at each round.

A6: Each faultless node correctly receives any task results broadcast in the system.

A7: Nodes cooperate in exchanging messages using a network-coding paradigm.

A8: Each node maintains an updated syndrome considering comparison syndromes generated in earlier rounds.

These assumptions assure the consistency of comparison outcomes produced by all faultless nodes in the system. In other words, faultless nodes reach a consensus about comparison outcomes. Hence, the correctness is guaranteed, and the completeness could be reached with probability approaching one in the long run. The probability of completeness comes from considering intermittent faults that might be undetectable. That is, complete diagnosis is non-deterministic and proportional to the fault's probability. On the other hand, faulty nodes may produce inconsistent outcomes, and hence, their behaviours are untrustworthy. In this model, each node takes the responsibility of diagnosing the system, postulating itself as a faultless node. In this sense, the diagnosis algorithm is executed in a fully distributed manner.

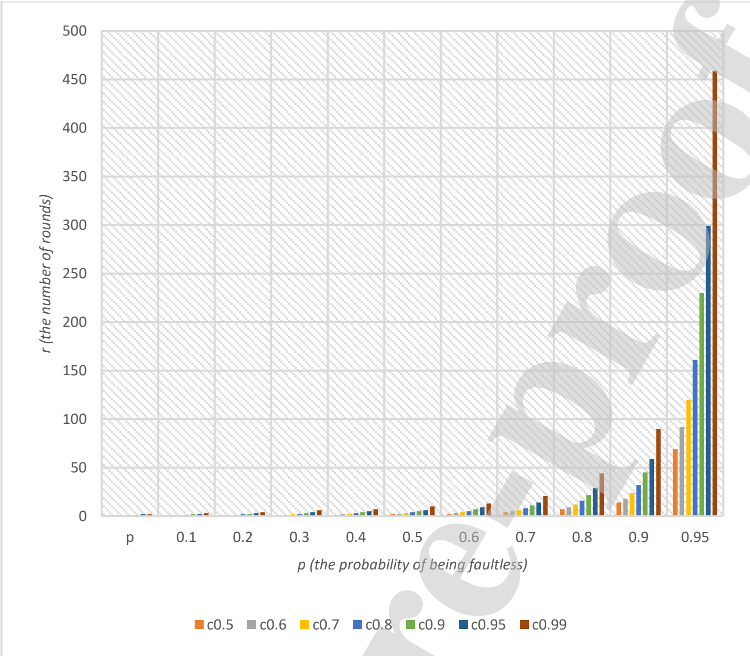

Figure 2. The number of rounds required for specific fault probabilities.

4. A Self-Diagnosis Protocol for Hybrid Faults

In this section, we introduce a distributed self-diagnosis protocol using a network coding-based comparison model for hybrid faults identification in mobile networks. It is called a network coding-based comparison distributed self-diagnosis protocol (NCBC-DSDP). Our proposed protocol can identify permanent and intermittent faults regardless of the change of the network topology in mobile networks. Thus, it supports more realistic assumptions about the fault models and network structure.

4.1 Protocol Description

Faultless nodes start diagnostic sessions at regular intervals. Each session includes multiple rounds of testing that run in sequence. The diagnostic session ends when each node completes the protocol execution. The protocol diagnoses nodes using various diagnostic messages, such as test request messages and test response messages, which have to be transmitted during a diagnostic session. Nodes, in each testing round, conform to the following procedures:

- **Task Assigning**
  A single task, j_i, is generated by a node, u, at the i^{th} test round. The task, j_i, is then assigned to every node, including u itself, in a network by retransmitting j_i among nodes, if necessary. Note that, given the broadcast nature of wireless networks, retransmission of the task could be dispensable. Since the goal is to have each node executing the same task, diverse mechanisms [36, 37] could be adopted to convey and assign j_i to each node with less overhead. Here, a simple probabilistic broadcasting mechanism [36, 38] is employed. The retransmission of j_i from a node depends on the forwarding factor, d_v,:

$$d_v = \frac{l}{|N_v|} \tag{4}$$

  where l is the number of times a node v has received the task j_i, and $|N_v|$ is the number of neighbour nodes for node v. In other words, d_v is inversely proportional to $|N_v|$. It has been proved mathematically and through simulations that the value of l should be ~ 6-8 in probabilistic forwarding and ~ 3 in case of network coding. This mechanism has been adopted widely to reduce the number of redundant rebroadcasts [38].

Using the same task in every test round is of interest becuase it reduces the complexity of task assignment and generation. Also, it enables nodes to perform the comparison in a fully distributed manner.

- **Task Execution**
  Upon receiving j_i, a node, v, executes the task and generates an output result, R_i^v. Afterwards, node v calculates the result's checksum, C_i^v , which will be broadcast to every node in the system. It is assumed that nodes collaborate with each other to disseminate their checksums through the network. This step helps each node to diagnose all other nodes in a fully distributed manner. However, it imposes high communication overheads. Therefore, this protocol employs a network coding technique to exchange the checksums among nodes [39, 40]. Network coding combines the checksums and sends out a single encoded message instead of transmitting them individually [41]. Once a node collects enough encoded messages, it can decode them and retrieve the original checksums. In the following, we describe the RLNC technique implemented in this protocol.

  Each node v generates and propagates encoded packets instead of sending original information. An encoded packet, e, is the total of multiplying each packet with the corresponding value in the coding vector, e as given by

$$e = \sum_{i=1}^{n} x_i \; y_i \tag{5}$$

  where $x = (x_1, x_2, \ldots\ldots, x_n)$ is a coding vector that comprises coefficients selected randomly from a finite field, F_{2^8}, n is the number of nodes, and $y = (y_1, y_2, \ldots\ldots, y_n)$ is a vector of original packets. In the case of sending C_i^v,e_v encoded packet from node v has one at v's location and zero at all other locations. Every node has a decoding matrix, M_v, which stores the coding vectors it receives. The size of M_v depends on the generation size. Upon receiving an innovative packet that increases the rank of M_v, v may generate and propagate linear combination of packets received based on a forwarding factor, d_v [38]. The non-innovative packets are ignored. In this protocol, packets for each task are combined linearly with each other. In other words, replies for a task could be encoded and decoded together. Once the decoding matrix becomes full rank, i.e., n linearly independent combinations have been received, the original checksums could be retrieved by performing a decoding process, namely Gaussian eliminations. It has been shown that the probability of the matrix being un-decodable tends to **0** and hence negligible [41-43].

- **Syndrome Generation**
  Once a node, v retrieves the original checksums, it compares their outputs with its own checksum and produces a comparison syndrome, assuming itself to be faultless. Here, node v maintains an updated syndrome, S, that maps nodes and their faulty status, $S < ID,\ status,\ cb >$, where ID is the identifier of the node, $status$ is the faulty status of the node, and cb is a parameter that describes whether a node status has changed during the diagnostic session. $status$ is set to 0 (faultless) for nodes reporting identical checksums with v and to **1** (faulty) for nodes reporting non-matching checksums. cb is initialized to **0** indicating that the node with ID has not changed its status from the previous test round. Both $status$ and cb are set to **1** if the node's $status$ has been changed from previous test rounds. The parameter cb is used to differentiate between permanent and intermittent faults.

After executing r test rounds, nodes, assuming themselves, to be faultless, take the responsibility of diagnosing the network depending on their updated syndromes. In this sense, the NCBC-DSDP provides a fully distributed diagnosis approach. Faultless nodes continuously report correct outputs, matching other faultless nodes. Therefore, $status = 0$ for every faultless node and $cb = 0$ indicates the consistency of $status$ over all tasks. On the other hand, nodes experiencing permanent faults report continuously incorrect matchless outputs. Thus, their $status$ in the updated syndromes is 1, and their $cb = 0$ because they show constant dissimilarity. Nodes undergoing intermittent faults, however, may manifest inconsistent outputs. That is, they might report correct outputs matching faultless ones for some tasks, and for other tasks, they report incorrect matchless outputs. Thus, their $status$ is 1 in the updated syndromes, and their cb is 1 indicating their changing behaviour.

It is noteworthy that intermittent faults are tricky, and they might be unobservable because they are likely to be exercised when they are inactive. In this case, nodes experiencing intermittent faults act like faultless nodes. Therefore, this model cannot guarantee a complete diagnosis.

4.2 Protocol Correctness and Analysis

In this subsection, we first prove the correctness of the proposed protocol and then we analyse its complexity in terms of communication, time, decoding, and energy complexities.

4.2.1 Proof of Correctness

This protocol guarantees a correct but incomplete diagnosis because nodes in a system may undergo intermittent faults, and these faults could be undetected after many tasks. Hence, in this subsection, we prove that the proposed protocol provides a 100% correct diagnosis and a complete diagnosis with high probability. That means the proposed protocol can diagnose all faultless nodes correctly; however, faulty nodes can be diagnosed correctly with a high probability. It is noteworthy that incomplete diagnosis is the best strategy for intermittent faults diagnosis.

We establish the correctness of the NCBC-DSDP based on two main properties: correct testing and correct responding. Correct testing is achieved if each node in the system receives the testing tasks in each testing round. Correct responding is achieved if the tasks' results are received correctly by each faultless node in the system. To prove the correctness of the NCBC-DSDP, we need to prove that it satisfies both the correct testing and the correct responding properties. Lemma 1 shows the correct testing property, followed by its proof.

Lemma 1. *Let $G_t = (V, L_t)$ be a communication graph that represent a mobile network at time t. Let $u \in V$ be a node that starts a diagnosis session, then tasks are generated for multiple test rounds and each node $v \in V$ will receive the same task in each test round.*

Proof. The NCBC-DSDP employs a probabilistic mechanism to broadcast tasks among nodes in a network. In this mechanism, each node, u, broadcasts a task received with a forwarding factor, $d_u = \frac{l}{|N_u|}$. Hence, we need to prove that this forwarding factor is enough to propagate the tasks to every node in the network. First, we need to analyse the expected additional coverage of rebroadcasting a test task. Let us consider the simple case where C_v and C_u are the circle areas of radius r that are covered by the transmissions of two neighbour nodes v and u respectively. When node v rebroadcasts the test task, which was broadcast by u, then the area, $|C_{v \cap u}|$ that is covered by both C_u and C_v is $4\int_{\partial/2}^{r} \sqrt{r^2 - x^2}\, dx$, where ∂ is the distance between u and v. Hence, the expected additional coverage when the node v broadcasts the test task is $|C_{u-v}| = \pi r^2 - 4\int_{\partial/2}^{r} \sqrt{r^2 - x^2}\, dx$. When $\partial = r$, $|C_{u-v}|$ is the largest coverage area and it approximately equals $0.61\pi r^2$. In other words, the maximum additional coverage is 61% of a circle area. However, we are interested in the average value of $|C_{u-v}|$ because neighbour nodes can be distributed randomly within the transmission range of the node u. The average value can be obtained by integrating the above value over the circle of radius x in $[0, r]$: the average value of $|C_{u-v}| = \int_0^r \frac{2\pi x.[\pi r^2 - 4\int_{x/2}^{r} \sqrt{r^2-x^2}]}{\pi r^2}\, dx \approx 41\%$. In other words, a rebroadcast provides, in average, about 41% additional coverage of a circle area. In [], it has been shown through simulations that if a node w, which was overheard the test task from both u and v, rebroadcasts the test task, then the additional coverage is about 19%. Further, the additional coverage tends to be 0 when a test task is overheard about 8 times [36, 38]. Accordingly, the NCBC-DSDP broadcasts the tasks with a forwarding factor d_v where $l = 8$. It is noteworthy that this value of l is independent of the network density.

Now, we prove that the NCBC-DSDP satisfies the correct responding property. This property is given in Lemma 2, followed by its proof.

Lemma 2. *Let $G_t = (V, L_t)$ be a communication graph that represent a mobile network at time t. Let the diagnosis session be initiated and a node $u \in V$ has generated a test response, then each faultless node $v \in V$ will receive and retrieve the u's response within a finite time.*

Proof. Here, we need to prove that the test responses are received and retrieved by every faultless node in the network. That is, the response messages that include task results generated by nodes are successfully received and retrieved by each faultless node in the network. In essence, this problem is an all-to-all broadcasting problem. Therefore, our proposed protocol employs the efficient broadcasting algorithm presented in [38]. In our protocol, each node transmits its results to its neighbour nodes. Intermediate nodes, then, transmit a random linear combination over a finite field F_{2^8} of other results, which have been previously received for the same task. In other words, replies for the same task are grouped as one generation, and the encoding and decoding are performed on packets of the same generation. Moreover, nodes transmit the random linear combinations based on a dynamic forwarding factor shown in Algorithm 6A in [38]. We have set $l = 3$ so that the probability of not being able to decode tends to 0. The detailed proof analysis can be found in [38].

Now, we prove that the NCBC-DSDP is a correct distributed self-diagnosis protocol because it guarantees that by the end of a diagnosis session each faultless node identifies all other faultless nodes correctly and identifies with high probability the faulty nodes in a system. This is shown in Theorem 1, followed by its proof.

Theorem 1. *Given a mobile network with a communication graph $G_t = (V, L_t)$, the NCBC-DSDP is correct with high probability of completeness. That is, the diagnosis output of the NCBC-DSDP at each faultless node, u is $\mathcal{FF} \subseteq FF_u$ and $F_u \subseteq \mathcal{F}$. In other words, u correctly identifies other faultless nodes in the system and identifies the faulty nodes with high probability.*

Proof. The proof of this Theorem follows trivially from Lemmas 1 and 2. That is, Lemma 1 ensures that each node is tested multiple times, executing several tasks. The nodes in the network correctly receive each task's result, and Lemma 2 ensures this. Once a faultless node, u, collects the results of a task, j, then node u updates the diagnosis syndrome as described in Section 4. Node u, next, diagnoses the network taking into consideration the assumptions investigated earlier. Faultless nodes execute the proposed protocol correctly; hence, they generate the same updated syndrome. The updated syndrome includes nodes that consistently generate matched results, and these are diagnosed as faultless nodes. Among these faultless nodes, there might be nodes that are experiencing intermittent faults; hence, a complete diagnosis is not guaranteed. The nodes that consistently generate mismatched results are diagnosed as permanent faults, whereas the nodes that generate mismatched results from time to time are diagnosed as intermittent faults. Accordingly, faultless nodes can correctly identify all other faultless nodes, and they can identify faulty nodes with high probability taking into account the number of testing rounds and the probability of faults. Faulty nodes are expected not to follow the proposed protocol, and their diagnosis is untrustworthy, and, in fact, that has no impact on other nodes decisions since the diagnosis process is fully distributed.

4.2.2 Complexity Analysis

Here, we analyse the complexities of our proposed protocol considering four main metrics. The first is communication complexity, which refers to the number of diagnosis messages exchanged during the diagnosis session. The second is time complexity, which refers to the duration of a diagnosis session, including all the test rounds. The third is computational decoding complexity, which reflects the necessary computations to decode the decoding matrix at each node. And the fourth is the energy consumption, which refers to the amount of energy consumed as a result of the diagnosis process. These are the main overheads caused by employing our proposed protocol, and the analysis is as follows.

The communication complexity of the NCBC-DSDP is $O(2 \times n + n \times n \times d)$; where n is the number of nodes, and d is the forwarding factor. In order to drive this communication complexity, the maximum number of diagnosis messages, which is transmitted during a diagnosis session, is calculated as follows. For each task, there is at most n messages to broadcast the task and n test response messages because every node should reply one time to each task. In addition, $n \times n \times d$ coding messages are required to exchange nodes' responses. Therefore, the total number of diagnosis messages exchanged is $2 \times n + n \times n \times d$ and hence the communication complexity is $O(2 \times n + n \times n \times d)$.

The time complexity of the NCBC-DSDP is $O(\sum_{i=1}^{r} \tau_i)$; where τ_i is the diagnosis duration for a specific round, i, and r is the number of rounds that is required for a given detection accuracy,c. In particular, we need to analyse the lower bound of the number of tasks that a faulty node should perform in order to be diagnosed correctly. Let p be the probability of a faulty node, u behaves as a faultless node at specific task. Since we consider intermittent faults, a faulty node,u may perform many tasks correctly and this may happen with probability, p^r. However, a faulty node is correctly diagnosed as faulty if it performs a task, j incorrectly and this may happen with probability $1 - p^r$. In other words, the detection accuracy, $c = 1 - p^r$. By solving this equation, we can calculate the number of rounds, r, which is required to achieve a certain detection accuracy. That is, $r = \left\lceil \frac{ln\,(1-c)}{ln\,(p)} \right\rceil$ and hence the time complexity is $O(\sum_{i=1}^{r} \tau_i)$.

The computational decoding complexity of the NCBC-DSDP is $O(r \cdot n^3)$. For each round, a decoding matrix of $n \times n$ size is required to store the received coding vectors at each node. Gaussian elimination can be performed on this matrix to solve the system with complexity bounded as $O(n^3)$. Because there is one such matrix per round, the decoding complexity is $O(r \cdot n^3)$. However, this is the worst case where all tasks are executed simultaneously. Various techniques could be used to reduce the decoding complexity. However, this is out of the scope of this paper, and interested readers can consult [38].

The energy complexity of the NCBC-DSDP is $O(\sum_{i=1}^{n} EC_i)$; where EC_i is the energy consumed by a node, i. Here, we study the energy consumed by nodes to send and receive diagnosis messages during the diagnosis session. It is of interest to mention that transmissions are the main source of energy consumption in mobile networks [44-46]. We assumed that the total energy consumed by a node u for diagnosing purpose is as follows. First, the transmission energy, which is required to transmit a message, is $E_{tx} = \beta * msgSize * E_0$, where β is a transmission energy factor, $msgSize$ is the length of the message, and E_0 is the energy required by a unit of message length. Second, the receiving energy, which is required to receive a message is $E_{rx} = \mu * msgSize * E_0$, where μ is a receiving energy factor. Hence the total energy consumed by a node u is $EC_u = msgSent \times E_{tx} + msgRec \times E_{rx}$, where $msgSent$ and $msgRec$ are, respectively, the number of messages sent and the number of messages received by the node u. Hence, the overall energy consumption is $\sum_{i=1}^{n} EC_i$ and hence the energy complexity is $O(\sum_{i=1}^{n} EC_i)$.

5. Performance Evaluation and Analysis

This section presents the performance results of the proposed protocol. An extensive set of simulations has been conducted using the OMNeT++ simulator [47, 48]. This simulator has been widely used because of its availability and credibility [49-52].

5.1 Performance Metrics

We consider three primary metrics to evaluate the performance of the proposed protocol. First is the communication overhead, which represents the total number of diagnostic messages exchanged during the diagnosis session. Second is the diagnosis time that determines the diagnosis session duration. The fewer the diagnostic messages and the shorter the diagnosis time are, the more efficient the protocol is. Third is the detection accuracy, which refers to the percentage of faulty nodes that are correctly diagnosed as faulty. We use this metric to evaluate the completeness property because intermittent faults are considered, and the complete diagnosis is not ensured. The higher the detection accuracy, the more efficient the protocol is.

5.2 Description of Scenarios

The following scenarios are used to evaluate the protocol performance under various circumstances. Table 2 summarises the configurations of each scenario.

- **Scenario 1** evaluates the efficiency of the proposed protocol to diagnose permanent faults in fixed-topology networks. Here, nodes do not move in the network. The network size varies from $10 - 100$ nodes with a random deployment where 10% of nodes are faulty. The performance of our proposed protocol, the NCBC-DSDP, is compared with other existing protocols, the Static-DSDP [22], Mobile-DSDP [23], and RLNC-DSDP [24].

- **Scenario 2** evaluates the efficiency of the proposed protocol to diagnose permanent faults in dynamic topology networks where some nodes keep moving in the network. The number of nodes is fixed at 50, where the number of faulty nodes varies between 1 and 5. The performance of our proposed protocol, the NCBC-DSDP is compared with the Mobile-DSDP and RLNC-DSDP, which consider mobile networks. For this study, the random-waypoint mobility model is used, with no waiting time and with a fixed speed at 20m/s. This mobility model allows the nodes to select a random destination in the simulated area, and then they start moving until arriving at the destination and so on.

- **Scenario 3** studies the effect of increasing the number of intermittent faults on the performance of our proposed protocol**.** The total number of nodes in this scenario is 80, where the number of faulty nodes varies between 2 and 30, with a high fault probability of 70%. Since this scenario focuses on diagnosing intermittent faults, three testing rounds are used.

- **Scenario 4** investigates the performance of our proposed protocol with the number of testing rounds, varying between 3 to 15. The total number of nodes is 80, where 20 nodes are experiencing intermittent faults with a low fault probability, 10%.

- **Scenario 5** studies the effect of node mobility on the performance of our proposed protocol. The total number of nodes is 50, where $2 - 18$ nodes are experiencing intermittent faults with 70% fault probability. This scenario uses the random-waypoint mobility model with no waiting time and with random node speeds between 2m/s and 20m/s. They are compared with no node mobility case. Three testing rounds are used in this scenario because intermittent faults are considered.

- **Scenario 6** studies the performance of our proposed protocol with various fault probabilities, varying between 0.1 to 1.0. In this scenario, the total number of nodes is 80, which includes 20 faulty nodes, experiencing intermittent and permanent faults. In this scenario, two testing rounds are used.

Table 2. Simulation Scenarios.

Notation	**Scenario 1**	**Scenario 2**	**Scenario 3**	**Scenario 4**	**Scenario 5**	**Scenario 6**
Protocols	Static-DSDP Mobile-DSDP RLNC-DSDP NCBC-DSDP	Mobile-DSDP RLNC-DSDP NCBC-DSDP	NCBC-DSDP	NCBC-DSDP	NCBC-DSDP	NCBC-DSDP
# of Nodes, n	10 - 100	50	80	80	50	80
Network Topology	Fixed	Dynamic	Fixed	Fixed	Dynamic	Fixed
Node Speed, s	0	20 m/s	0	0	2 – 20 m/s	0
Fault Types	Permanent	Permanent	Intermittent	Intermittent	Intermittent	Permanent & Intermittent
# of Faulty Nodes, n_f	1 – 10	1 – 5	2 – 30	20	2 – 18	20
# of Test Rounds, r	1	1	3	3 – 15	3	2
Fault Probability, P_f	100%	100%	70%	10%	70%	10% - 100%
Area Size	600m×600m	300m×300m	600m×600m	600m×600m	300m×300m	600m×600m
Transmission Range	150	150	150	150	150	150

In Scenarios 3, 4, 5 and 6, we focused on studying the performance of our proposed protocol and evaluating its ability to diagnose intermittent faults under various settings. However, we exclude other protocols since they do not diagnose intermittent faults.

5.3 Simulation Results and Analysis

This subsection presents the simulation results obtained for the scenarios described in the previous subsection. The simulations are carried out and repeated 10 times with different random seed numbers to provide a confidence interval of 95% ± 10% margin of error. Hence, the simulation results represent the average value of the 10 runs for every point.

5.3.1 Performance Comparison of Various Diagnose Protocols for Permanent Faults in Fixed Networks (Scenario 1)

Fig. 3 compares the communication overhead of our proposed protocol, the NCBC-DSDP, and the other existing protocols, the Static-DSDP, Mobile-DSDP and RLNC-DSDP with a different number, n, of nodes. It can be seen that the communication overhead of all considered protocols increases with n in the network. The reason is that these protocols perform the diagnosis process in a distributed fashion where each node participates in the diagnosis process and transmits diagnosis messages. Therefore, the larger the n, the more diagnosis messages the nodes transmit. However, NCBC-DSDP shows much better performance. For example, at $n = 100$, the NCBC-DSDP sends about $85 - 80\%$ fewer diagnosis messages than the Static-DSDP and Mobile-DSDP protocols, and 50% fewer diagnosis messages than the RLNC-DSDP. There are two main reasons for these results. The first reason is that the NCBC-DSDP requires two kinds of diagnosis messages to be exchanged during the diagnosis session, whereas the other considered protocols need three kinds of messages to be exchanged by each node in the network. In other words, the NCBC-DSDP, in essence, demands fewer diagnosis messages. The second reason is that the NCBC-DSDP employs a probabilistic flooding and RLNC technique to propagate the diagnosis messages and that reduces the required number of diagnosis messages significantly. The Static-DSDP and Mobile-DSDP employ a simple flooding mechanism to propagate the local views of nodes, and that causes too much communication overhead, as shown in Fig. 3. Unlike the Static-DSDP, which needs nodes to reply to each task they receive, the Mobile-DSDP and RLNC-DSDP require nodes to reply to $\sigma \leq n$; therefore, nodes send a fewer number of test response messages. This reduction in the number of test responses makes Mobile-DSDP slightly better than the Static-DSDP. The RLNC-DSDP shows better performance than the Static-DSDP and Mobile-DSDP because it uses an RLNC technique instead of simple flooding to propagate the local view of nodes.

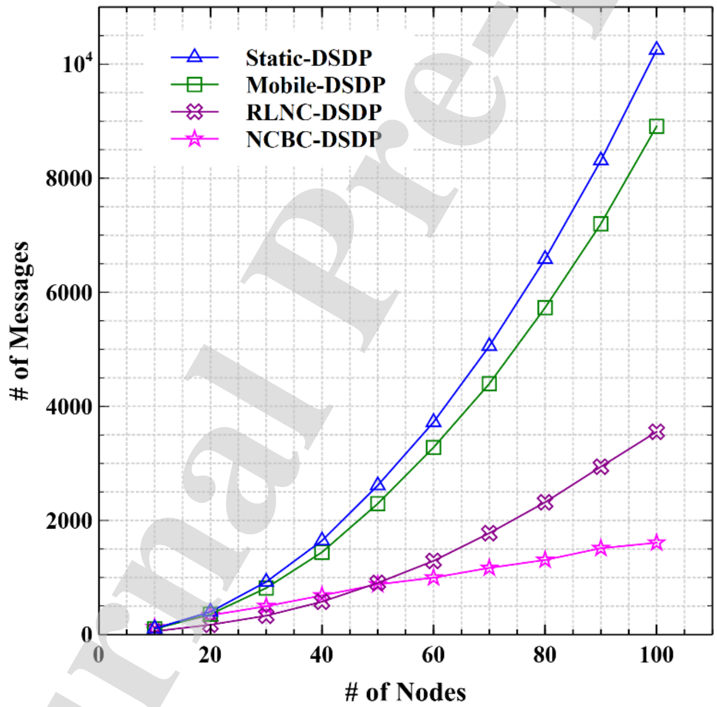

Figure 3. The communication overhead vs number, n, of nodes in Scenario 1.

Fig. 4 presents the diagnosis time of the NCBC-DSDP, Static-DSDP, Mobile-DSDP and RLNC-DSDP with different n. Clearly, NCBC-DSDP shows less diagnosis time than other protocols for the same reasons described above: the two-step diagnosis process that the NCBC-DSDP uses and the RLNC technique that it employs. Since we only consider permanent faults in this scenario, all the protocols achieve 100% detection accuracy within a single testing round.

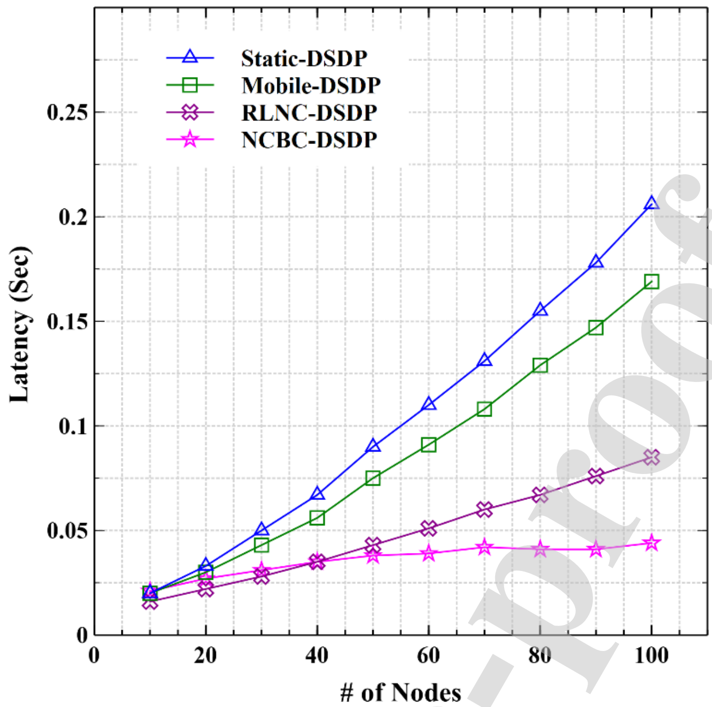

Figure 4. The diagnosis time vs number, n, of nodes in Scenario 1.

5.3.2 Performance Comparison of Various Diagnose Protocols for Permanent Faults in Mobile Networks (Scenario 2)

Fig. 5 compares the communication overhead of the Mobile-DSDP, RLNC-DSDP and NCBC-DSDP protocols in a mobile network under Scenario 2. In this scenario, the movement of nodes leads to topology changes. Here, we excluded the Static-DSDP from this comparison because it tolerates no topology changes. As seen in Fig. 5, the communication overhead of the NCBC-DSDP is nearly constant regardless of topology changes. That is, the NCBC-DSDP shows robust behaviour against topology changes. Moreover, it is less than both the Mobile-DSDP and RLNC-DSDP. For example, when the number of faulty nodes is 5, the Mobile-DSDP sends 4 times more diagnosis messages than the NCBC-DSDP. In addition, the RLNC-DSDP sends about 1.5 times more diagnosis messages than the NCBC-DSDP. The figure shows that the RLNC-DSDP is more efficient than the Mobile-DSDP, imposing less communication overhead. Again, the credit goes to the RLNC technique employed by the RLNC-DSDP.

Fig. 6 compares the diagnosis time of the protocols under Scenario 2. All the considered protocols show robust results with the increase in the number of faulty mobiles. Clearly, the diagnosis time of the Mobile-DSDP is about 4 times longer than the NCBC-DSDP's diagnosis time. In addition, the diagnosis time of the RLNC-DSDP is approximately double that of the NCBC-DSDP. Again, all the considered protocols achieve 100% detection accuracy within single testing round because only permanent faults are considered. That is, every fault-free node, either mobile or fixed, identified both the fixed and the faulty mobile nodes.

5.3.3 Performance of the NCBC-DSDP for Different Number of Intermittent Faults in a Fixed Network (Scenario 3)

Fig. 7 shows the number of messages required to detect various numbers, n_f, of faults after $r = 3$ testing rounds under Scenario 3. The figure shows that the number of messages is independent of n_f. Similarly, Fig. 8 shows that the diagnosis time is more or less constant with n_f. Both these results show the robustness of the NCBC-DSP regarding the increasing number of faults in the network. Fig. 9 illustrates the detection accuracy of the NCBC-DSDP under Scenario 3. It can be seen that the detection accuracy is above 95% due to using a high fault probability, P_f = 70% and $r = 3$. This result is consistent with the expected detection accuracy shown in Fig. 2.

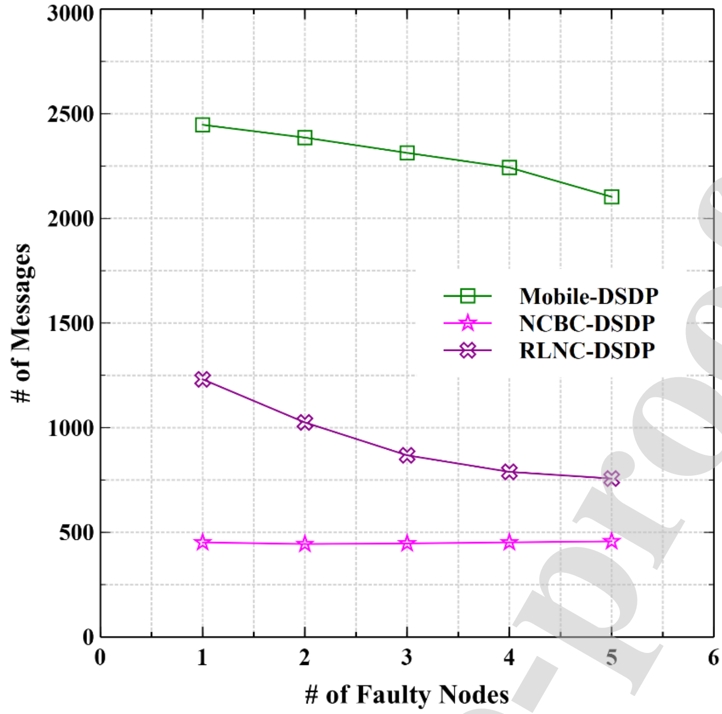

Figure 5. The communication overhead vs the number of faulty mobiles in Scenario 2.

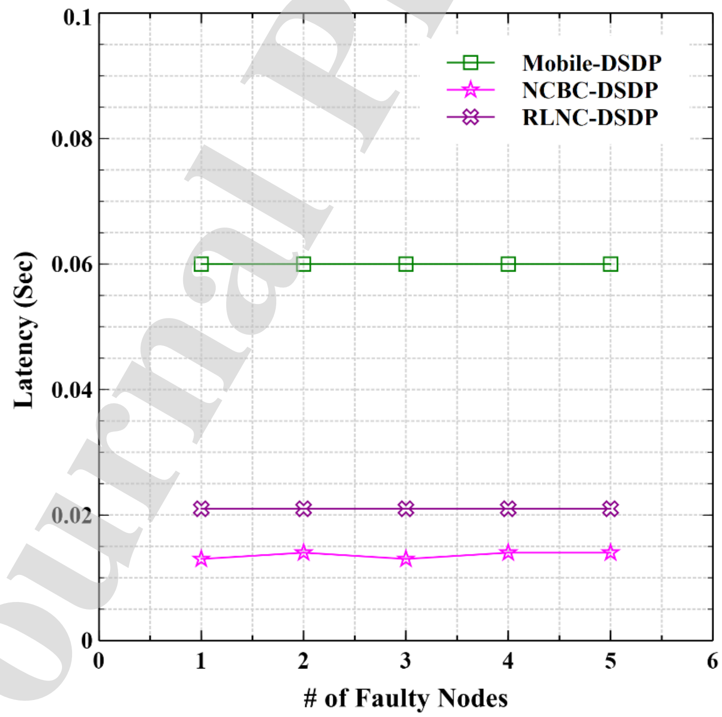

Figure 6. The diagnosis time vs the number of faulty mobiles in Scenario 2.

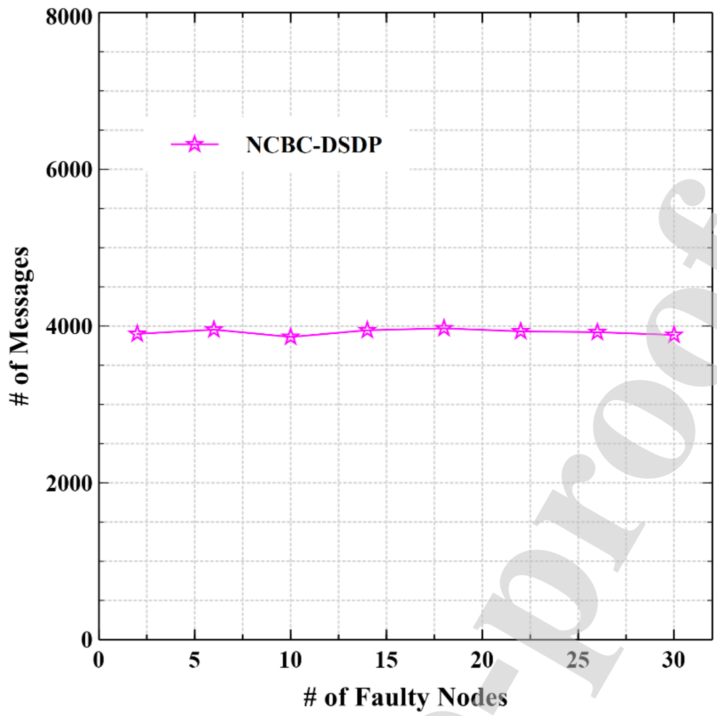

Figure 7. The communication overhead vs the number, n_f, of faulty nodes in Scenario 3.

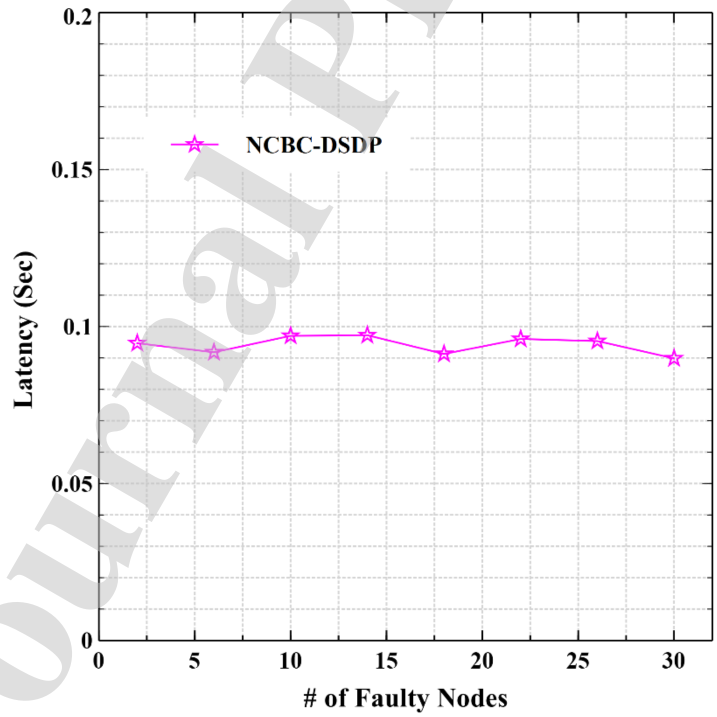

Figure 8. The diagnosis time vs the number, n_f, of faulty nodes in Scenario 3.

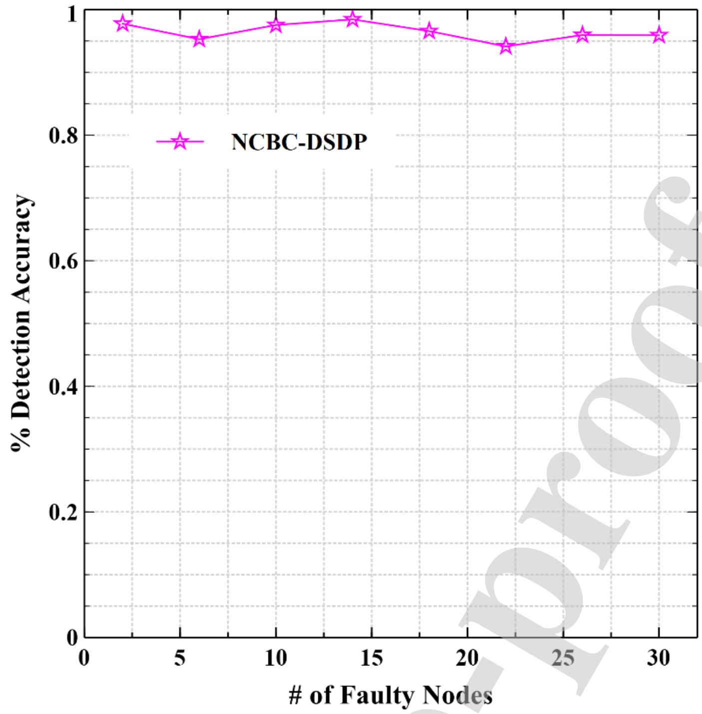

Figure 9. The detection accuracy vs the number, n_f, of faulty nodes in Scenario 3.

5.3.4 Performance of the NCBC-DSDP for Intermittent Faults with Different Number of Testing Rounds in a Fixed Network (Scenario 4)

Fig. 10 and Fig. 11 show the communication overhead and the diagnosis time of the NCBC-DSDP with the number,r of testing rounds under Scenario 4, respectively. As expected, both the number of diagnosis messages and the diagnosis time increase with r. This is because the diagnosis process goes through multiple rounds in order to diagnose nodes at different times.

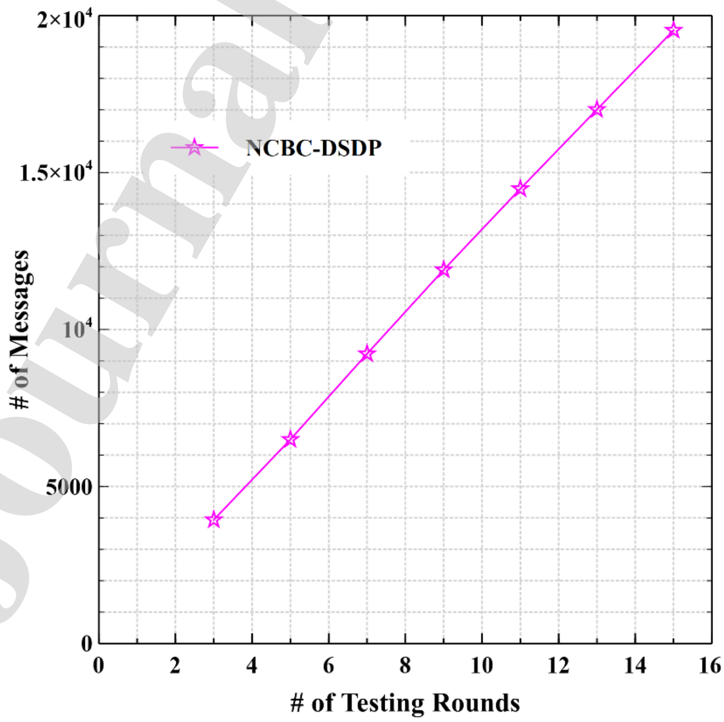

Figure 10. The communication overhead vs number, r, of testing rounds in Scenario 4.

Fig. 12 presents the detection accuracy of the NCBC-DSDP in this scenario. Clearly, the detection accuracy increases with r, as excepted, because the detection accuracy is directly proportional to the number of testing rounds. In this scenario, because we use a very low fault probability, P_f = 10%, the fault detection becomes harder. The figure shows that using 3 testing rounds, the detection accuracy of the NCBC-DSDP is about 53%. That is, fault-free nodes in the network have identified about 11 out of 20 faulty nodes. The detection accuracy rises above 80% after 7 testing rounds. It is clear that a high detection accuracy comes with a high communication overhead and a long diagnosis time.

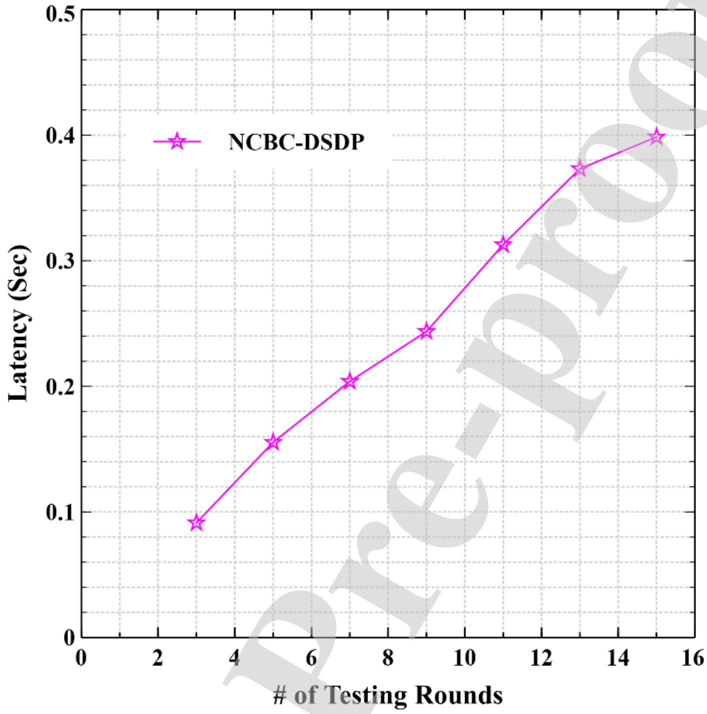

Figure 11. The diagnosis time vs number, r, of testing rounds in Scenario 4.

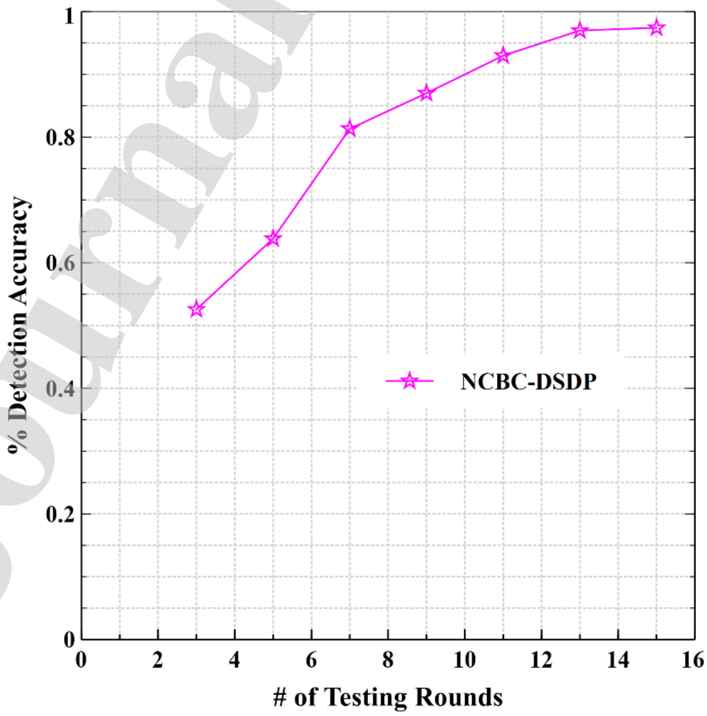

Figure 12. The detection accuracy vs number, r, of testing rounds in Scenario 4.

5.3.5 Performance of NCBC-DSDP for Intermittent Faults in Mobile Network with various Speeds (Scenario 5)

Fig. 13 shows the number of diagnosis messages required to detect a variable number of faulty nodes using the NCBC-DSDP in mobile networks where nodes are moving with various speeds. Clearly, there is no significant difference between the performances under various speeds. In other words, the communication overhead is robust against node speeds and topology changes. Fig. 14 also shows a similar observation that the latency is independent of speeds. Fig. 15 compares the detection accuracy under different speeds. Clearly, the detection accuracy is robust with topology changes.

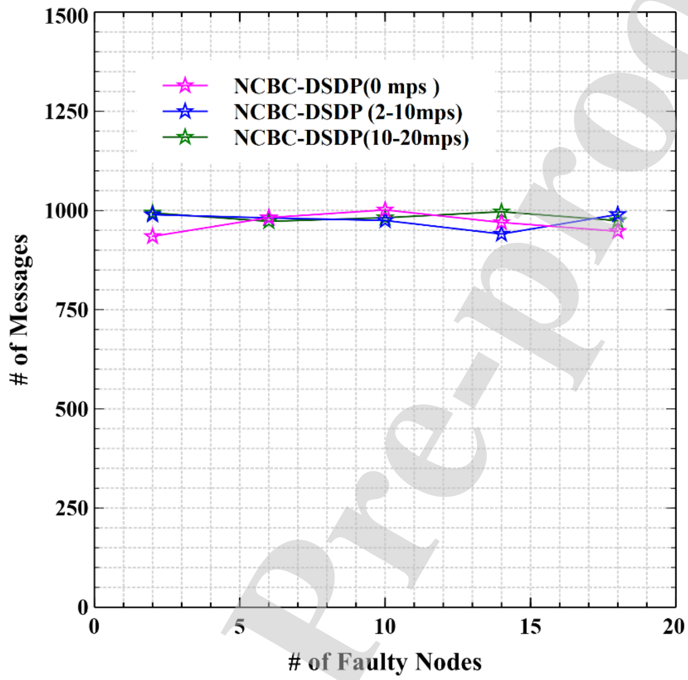

Figure 13. The communication overhead vs the number of faulty nodes under various node speeds (Scenario 5).

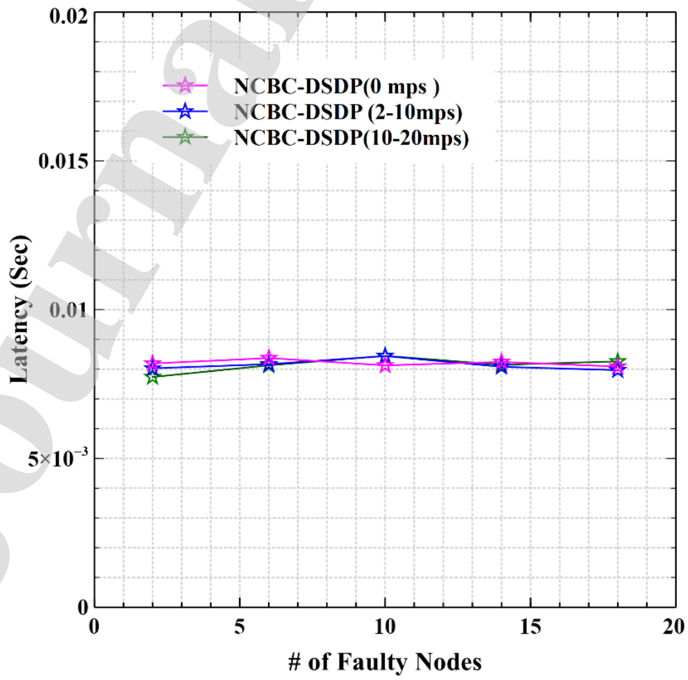

Figure 14. The diagnosis time vs the number of faulty nodes under various node speeds (Scenario 5).

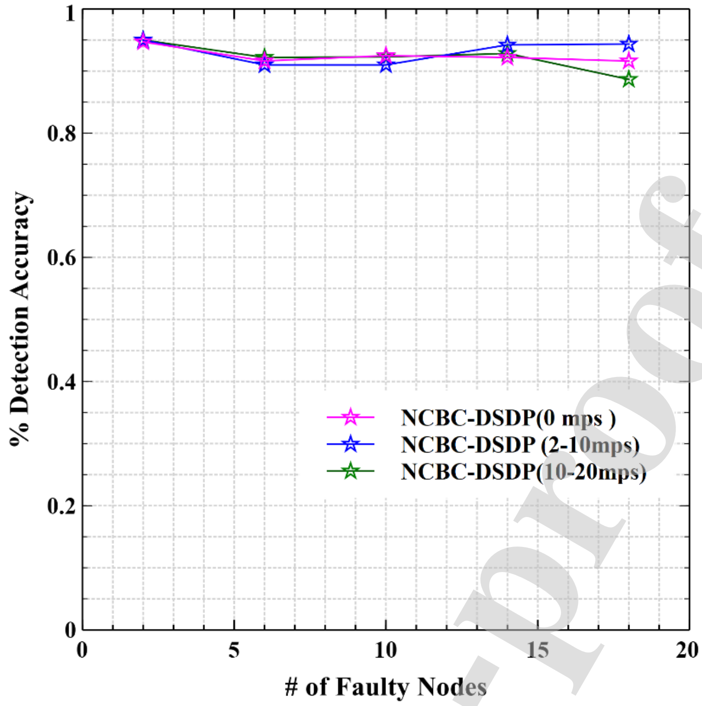

Figure 15. The detection accuracy vs the number of faulty nodes under various node speeds (Scenario 5).

5.3.6 Performance of the NCBC-DSDP for Intermittent and Permanent Faults in a Fixed Network (Scenario 6)

Fig. 16 illustrates the communication overhead of the NCBC-DSDP protocol regarding various fault probabilities after a fixed number of testing rounds, r = 2 under Scenario 6. It can be seen that the number of messages exchanged is independent of P_f, and it is approximately 2,350 messages.

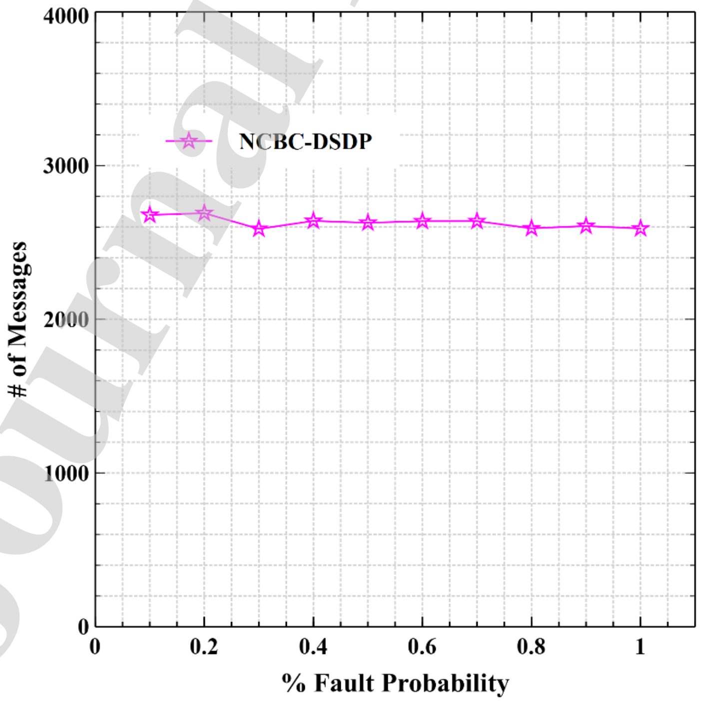

Figure 16. The communication overhead vs fault probability, P_f, in Scenario 6.

A similar observation can be seen in Fig. 17, which shows the diagnosis time of the NCBC-DSDP. In Fig. 17, it is clear that the diagnosis time of the NCBC-DSDP is more or less the same with P_f . This is because the number of testing rounds is fixed to 2. Fig. 18 shows the detection accuracy with regards to different fault probabilities after 2 testing rounds. The detection accuracy is about 25% for P_f = 10%. In other words, just 5 out of 20 faulty nodes can be detected correctly. However, for P_f = 90%, the detection accuracy increases above 95%. These results are reasonable because the lower the fault probability, the higher a node acts faultlessly; hence, the more difficult it will be to detect the fault. Thus, we can conclude the detection accuracy is inversely proportional to P_f.

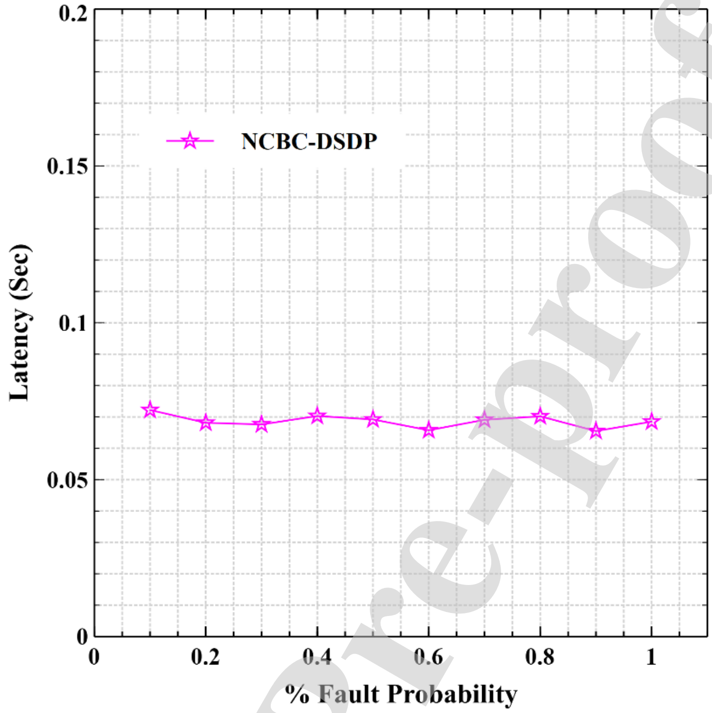

Figure 17. The diagnosis time vs. fault probability, P_f, in Scenario 6.

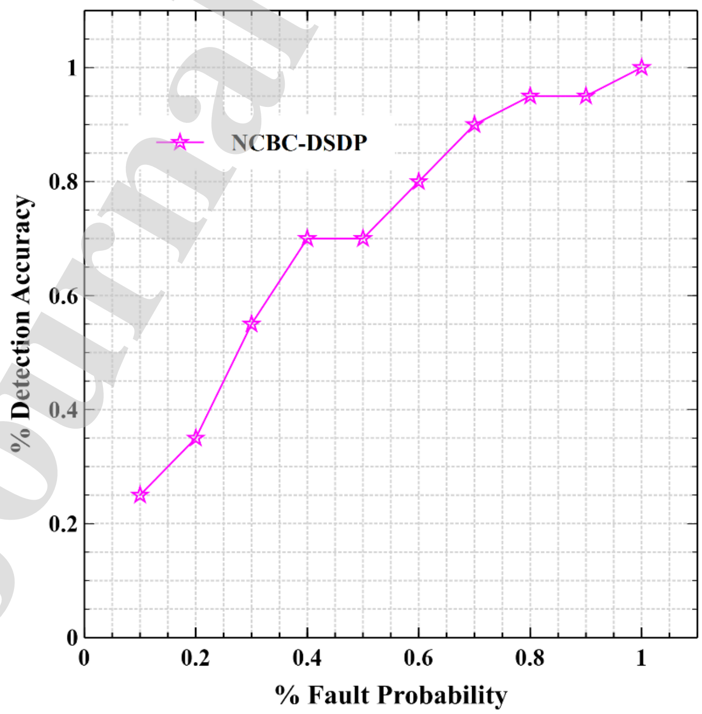

Figure 18. The detection accuracy vs. fault probability, P_f, in Scenario 6.

5.3.7 The performance of the NCBC-DSDP in terms of energy consumption

Here, we investigate the energy consumption of our proposed protocol compared to other protocols. The same settings of Scenario 1 were used in this investigation. Further, the following energy settings were considered. The energy consumption in transmission mode is 330mA and the energy consumption in receiving mode is 230mA. Both modes use a 5.0V energy supply [44, 45].

Fig. 19 compares the average energy consumption of our proposed protocol, the NCBC-DSDP, and the other existing protocols, the Static-DSDP, Mobile-DSDP and RLNC-DSDP with a different number of nodes. Overall, the energy consumption increases with the number of nodes for all considered protocols. This is because the larger the number of nodes is, the more diagnosis messages the nodes transmit and receive, and the more energy the nodes consumed. However, the NCBC-DSDP shows better performance in terms of energy consumption. For example, when the number of nodes is 100, the average energy consumption per node in the NCBC-DSDP is about 100 Joules. Using the RLNC-DSDP, the average energy consumption is about 250 Joules. The Static-DSDP and Mobile-DSDP consume 7-9 times more energy than the RLNC-DSDP. However, the results show insignificant differences when the number of nodes is less than 30. The main reason is that the number of transmissions of the NCBC-DSDP and RLNC-DSDP is very close to that of the Static-DSDP and Mobile-DSDP, as shown in Fig. 3. Particularly, this is due to the sparsity of the network.

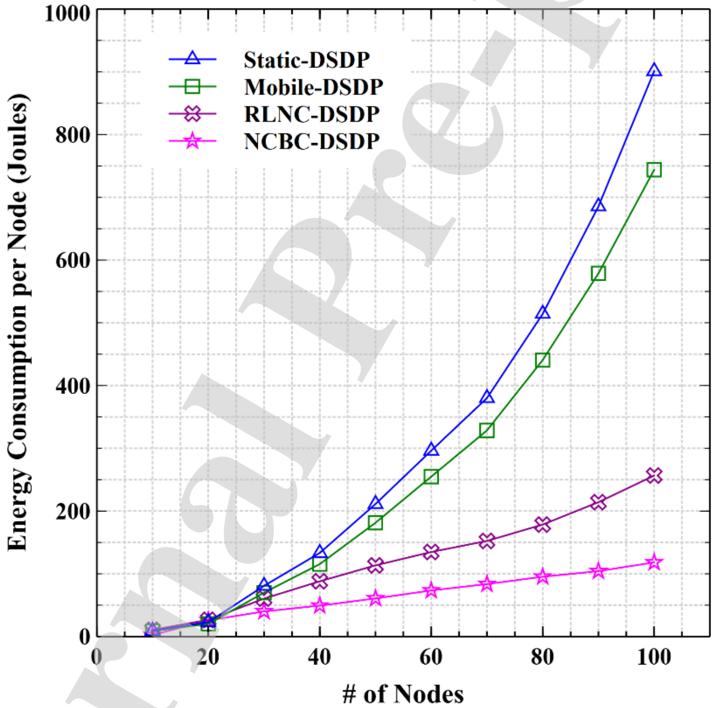

Figure 19. The average energy consumption per node vs the number of nodes in Scenario 1

6. Discussions and Conclusions

This paper investigates the problem of system-level diagnosis for hybrid faults in mobile networks. This problem aims to enhance the dependability of a network by allowing fault-free nodes to reach a consensus on the state of the network, identifying faulty nodes and avoiding their impact. This problem is of utmost interest in mobile networks where nodes perform their tasks, such as routing cooperatively; hence, being aware of the state of other nodes in the network is crucial for successful operations. While the comparison approach is one of the most practical approaches that have been proposed to deal with the system diagnosis problem, existing comparison-based diagnosis models suffer from various limitations in mobile networks. Notably, the majority of existing models diagnose fixed topology networks and permanent faults. Nonetheless, mobile networks experience dynamic topology and intermittent faults more often.

Therefore, the main contribution of this paper is the proposed probabilistic comparison-based model. The benefit of this model lies in its potential to diagnose hybrid faults in mobile networks, respecting their intrinsic characteristics. This model can provide a correct diagnosis with high probability completeness using test tasks. Identifying intermittent faults is a challenge because several testing rounds are required and that imposes too much overhead. The proposed model employs a network coding communication paradigm to exchange the diagnosis messages so that the diagnosis overhead is minimized. It is noteworthy that the proposed diagnostic model could employ regular tasks that a mobile network performs; hence, the diagnosis overhead it causes can be easily justified. This paper also advanced the NCBC-DSDP protocol as the one to implement the proposed model, demonstrating its potential to improve the dependability of mobile networks. The NCBC-DSDP imposes no restrictions on network topology or communication; hence, it is the most suitable protocol for mobile networks to the best of our knowledge.

Both the analytical and the simulation results verify the merit of the NCBC-DSDP in terms of communication and time complexity as well as detection accuracy. The simulation results show that the proposed protocol could perform the diagnosis process with very little overhead and very low latency. Reduced overhead is a benefit for energy efficiency because energy consumption depends primarily on the number of messages exchanged. The detection accuracy is high. It could also be improved by executing more testing rounds if more accuracy is required. However, multiple testing rounds should be used with caution because it increases the diagnosis overhead. It is recommended that the diagnosis process stops when a satisfactory accuracy is achieved. In other words, the diagnosis process can be adaptive with the detection accuracy rather than using a fixed number of testing rounds. These results prove that the NCBC-DSDP can provide a correct diagnosis with a high probability of completeness under both static and dynamic network topologies. In other words, the results demonstrate the robustness of our proposed protocol with topology changes.

The future directions of this research include investigating the implementations of our proposed model into Vehicular Ad Hoc Networks (VANETs) and Wireless Mesh Networks (WMNs). These networks have infrastructure that supports their operations, and that can be useful to reduce the diagnosis overhead. In this sense, nodes forming the backbone of a network can contribute to the diagnosis process more than other nodes in the network. Another future direction would be to study a more general fault model that considers malicious faults. Our proposed model and protocol are limited to benign independent faults. However, it is very common to have malicious and dependent faults in mobile networks. Therefore, there is a need to deal with such kind of faults in mobile networks

References

[1] P. Ruiz and P. Bouvry, "Survey on broadcast algorithms for mobile ad hoc networks," *ACM computing surveys (CSUR),* vol. 48, p. 8, 2015.

[2] A. Rasheed, S. Gillani, S. Ajmal, and A. Qayyum, "Vehicular Ad Hoc Network (VANET): A Survey, Challenges, and Applications," Singapore, 2017, pp. 39-51.

[3] D. Reina, M. Askalani, S. Toral, F. Barrero, E. Asimakopoulou, and N. Bessis, "A survey on multihop ad hoc networks for disaster response scenarios," *International Journal of Distributed Sensor Networks,* vol. 11, p. 647037, 2015.

[4] P. Novotny, B. J. Ko, and A. L. Wolf, "Locating Faults in MANET-Hosted Software Systems," *IEEE Transactions on Dependable and Secure Computing,* vol. 15, pp. 452-465, 2018.

[5] K. Miranda, A. Molinaro, and T. Razafindralambo, "A survey on rapidly deployable solutions for post-disaster networks," *IEEE Communications Magazine,* vol. 54, pp. 117-123, 2016.

[6] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: a survey," *Computer Networks,* vol. 47, pp. 445-487, 2005.

[7] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer Networks,* vol. 52, pp. 2292-2330, 2008.

[8] T. Qiu, N. Chen, K. Li, D. Qiao, and Z. Fu, "Heterogeneous ad hoc networks: Architectures, advances and challenges," *Ad Hoc Networks,* vol. 55, pp. 143-152, 2017.

[9] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Communications Surveys & Tutorials,* vol. 21, pp. 2224-2287, 2019.

[10] S. K. Sarkar, T. G. Basavaraju, and C. Puttamadappa, *Ad hoc mobile wireless networks: principles, protocols, and applications*: CRC Press, 2016.

[11] D. Raposo, A. Rodrigues, J. S. Silva, and F. Boavida, "A Taxonomy of Faults for Wireless Sensor Networks," *Journal of Network and Systems Management,* vol. 25, pp. 591-611, July 01 2017.

[12] Z. Zhang, A. Mehmood, L. Shu, Z. Huo, Y. Zhang, and M. Mukherjee, "A survey on fault diagnosis in wireless sensor networks," *IEEE Access,* vol. 6, pp. 11349-11364, 2018.

[13] C. Titouna, M. Aliouat, and M. Gueroui, "FDS: Fault Detection Scheme for Wireless Sensor Networks," *Wireless Personal Communications,* vol. 86, pp. 549-562, January 01 2016.

[14] A. Avizienis, J.-C. Laprie, B. Randell, and C. Landwehr, "Basic concepts and taxonomy of dependable and secure computing," *IEEE Transactions on Dependable and Secure Computing,* vol. 1, pp. 11-33, 2004.

[15] F. Foukalas, P. Pop, F. Theoleyre, C. A. Boano, and C. Buratti, "Dependable Wireless Industrial IoT Networks: Recent Advances and Open Challenges," in *2019 IEEE European Test Symposium (ETS)*, 2019, pp. 1-10.

[16] F. P. Preparata, G. Metze, and R. T. Chien, "On the Connection Assignment Problem of Diagnosable Systems," *IEEE Transactions on Electronic Computers,* vol. EC-16, pp. 848-854, 1967.

[17] A. T. Dahbura, *System-level diagnosis: A perspective for the third decade*: Plenum Press, 1988.

[18] A. K. Somani, "System level diagnosis: A review," *Technique Report, Dependable Computer Laboratory, Iowa State University,* 1997.

[19] E. P. Duarte Jr, R. P. Ziwich, and L. C. Albini, "A survey of comparison-based system-level diagnosis," *ACM Computing Surveys (CSUR),* vol. 43, p. 22, 2011.

[20] H. Jarrah, N. I. Sarkar, and J. Gutierrez, "Comparison-based system-level fault diagnosis protocols for mobile ad-hoc networks: A survey," *Journal of Network and Computer Applications,* vol. 60, pp. 68-81, 2016.

[21] D. M. Blough and H. W. Brown, "The broadcast comparison model for on-line fault diagnosis in multicomputer systems: theory and implementation," *IEEE Transactions on Computers,* vol. 48, pp. 470-493, 1999.

[22] S. Chessa and P. Santi, "Comparison-based system-level fault diagnosis in ad hoc networks," in *Proceedings. 20th IEEE Symposium on Reliable Distributed Systems, 2001.* , 2001, pp. 257-266.

[23] M. Elhadef, A. Boukerche, and H. Elkadiki, "Diagnosing mobile ad-hoc networks: two distributed comparison-based self-diagnosis protocols," in *Proceedings of the 4th ACM international workshop on Mobility management and wireless access*, 2006, pp. 18-27.

[24] H. Jarrah, P. Chong, N. I. Sarkar, and J. Gutierrez, "A Time-Free Comparison-Based System-Level Fault Diagnostic Model for Highly Dynamic Networks," in *Proceedings of the 11th International Conference on Queueing Theory and Network Applications*, 2016, p. 12.

[25] M. Malek, "A comparison connection assignment for diagnosis of multiprocessor systems," in *Proceedings of the 7th annual symposium on Computer Architecture*, 1980, pp. 31-36.

[26] K.-Y. Chwa and S. L. Hakimi, "Schemes for fault-tolerant computing: A comparison of modularly redundant and t-diagnosable systems," *Information and Control,* vol. 49, pp. 212-238, 6// 1981.

[27] A. T. Dahbura, K. K. Sabnani, and L. L. King, "The comparison approach to multiprocessor fault diagnosis," *IEEE Transactions on Computers,* pp. 373-378, 1987.

[28] A. Pelc, "Undirected graph models for system-level fault diagnosis," *IEEE Transactions on Computers,* vol. 40, pp. 1271-1276, 1991.

[29] S. Rangarajan and D. Fussell, "A probabilistic method for fault diagnosis of multiprocessor systems," in *1988 The Eighteenth International Symposium on Fault-Tolerant Computing. Digest of Papers*, 1988, pp. 278-283.

[30] M. N. Sahoo and P. M. Khilar, "Intermittent fault diagnosis in dynamic topology MANETs," *International Journal of Signal and Imaging Systems Engineering,* vol. 8, pp. 345-355, 2015.

[31] H. N. Azar and H. T. Malazi, "Decentralized detection of hybrid faults in mobile sensor nodes," *Simulation Modelling Practice and Theory,* vol. 87, pp. 210-225, 2018.

[32] L. Xu, L. Ji, and S. M. Zhou, "An efficient self-diagnosis protocol for hierarchical wireless mesh networks," *Concurrency and Computation: Practice and Experience,* vol. 25, pp. 2036-2051, 2013.

[33] N. Aljeri, M. Almulla, and A. Boukerche, "An efficient fault detection and diagnosis protocol for vehicular networks," in *Proceedings of the third ACM international symposium on Design and analysis of intelligent vehicular networks and applications*, 2013, pp. 23-30.

[34] D. Reina, S. Toral, P. Johnson, and F. Barrero, "A survey on probabilistic broadcast schemes for wireless ad hoc networks," *Ad Hoc Networks,* vol. 25, pp. 263-292, 2015.

[35] V. Bhandari and N. H. Vaidya, "Reliable broadcast in radio networks with locally bounded failures," *IEEE Transactions on Parallel and Distributed Systems,* vol. 21, pp. 801-811, 2009.

[36] Y.-C. Tseng, S.-Y. Ni, Y.-S. Chen, and J.-P. Sheu, "The broadcast storm problem in a mobile ad hoc network," *Wireless Networks,* vol. 8, pp. 153-167, 2002.

[37] R. C. Gómez, I. Gonzalez-Herrera, Y.-D. Bromberg, L. Réveillère, and E. Rivière, "Density and mobility-driven evaluation of broadcast algorithms for MANETs," in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, 2017, pp. 2308-2313.

[38] C. Fragouli, J. Widmer, and J.-Y. Le Boudec, "Efficient broadcasting using network coding," *IEEE/ACM transactions on Networking*, vol. 16, pp. 450-463, 2008.

[39] A. Mostefaoui, E. Mourgaya, and M. Raynal, "Asynchronous implementation of failure detectors," in *2013 43rd Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, 2003, pp. 351-351.

[40] T. Matsuda, T. Noguchi, and T. Takine, "Survey of network coding and its applications," *IEICE Transactions on Communications*, vol. 94, pp. 698-717, 2011.

[41] Y. Wu, "Network coding for wireless networks," in *Adaptation and Cross Layer Design in Wireless Networks*, ed: CRC Press, 2018, pp. 213-242.

[42] J. Widmer and J.-Y. Le Boudec, "Network coding for efficient communication in extreme networks," in *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, 2005, pp. 284-291.

[43] C. Fragouli, J. Widmer, and J.-Y. Le Boudec, "On the benefits of network coding for wireless applications," in *4th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, 2006*, 2006, pp. 1-6.

[44] J.-C. Cano and P. Manzoni, "A performance comparison of energy consumption for mobile ad hoc network routing protocols," in *Proceedings 8th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (Cat. No. PR00728)*, 2000, pp. 57-64.

[45] T. Wang and C. Huang, "Energy Consumption analysis under Random Mobility Model," in *Proceedings of 2011 International Conference on Engineering and Information Management (ICEIM 2011)*, 2011.

[46] N. E. Majd, N. Ho, T. Nguyen, and J. Stolmeier, "Evaluation of parameters affecting the performance of routing protocols in mobile ad hoc networks (MANETs) with a focus on energy efficiency," in *Future of Information and Communication Conference*, 2019, pp. 1210-1219.

[47] A. Varga and R. Hornig, "An overview of the OMNeT++ simulation environment," in *Proceedings of the 1st international conference on Simulation tools and techniques for communications, networks and systems & workshops*, 2008, p. 60.

[48] A. Varga, "A practical introduction to the OMNeT++ simulation framework," in *Recent Advances in Network Simulation*, ed: Springer, 2019, pp. 3-51.

[49] K. Wehrle, M. Günes, and J. Gross, *Modeling and tools for network simulation*: Springer Science & Business Media, 2010.

[50] G. Chengetanai and G. B. O'Reilly, "Survey on simulation tools for wireless mobile ad hoc networks," in *2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, 2015, pp. 1-7.

[51] A. S. Toor and A. Jain, "A survey on wireless network simulators," *Bulletin of Electrical Engineering and Informatics*, vol. 6, pp. 62-69, 2017.

[52] A. Zarrad and I. Alsmadi, "Evaluating network test scenarios for network simulators systems," *International Journal of Distributed Sensor Networks*, vol. 13, p. 1550147717738216, 2017.

29 Jan 2020

Dear Editor of Journal of Computer Communications,

Our submitted paper has 5 authors. The contributions of each author are as follows:

Hazim Jarrah: Methodology, Software, Formal analysis, Writing- Original draft preparation

Peter Chong: Conceptualization, Resources, Validation, Supervision, Writing - Review & Editing.

Chris Rapson: Visualization, Investigation.

Nurul I. Sarkar: Supervision, Funding acquisition, Writing - Review & Editing

Jairo Gutierrez: Writing - Review & Editing, Supervision

Thank you.

Best rgds,

Prof. Peter Chong

Auckland University of Technology

29 Jan 2020

Dear Editor of Journal of Computer Communications,

We like to confirm that there is no conflict of interest for all the authors in our submitted paper.

Thank you.

Best rgds,

Prof. Peter Chong

Auckland University of Technology