

Received December 26, 2021, accepted January 19, 2022, date of publication January 28, 2022, date of current version February 9, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3147498

Data Dissemination in VANETs Using Clustering and Probabilistic Forwarding Based on Adaptive Jumping Multi-Objective Firefly Optimization

MUSTAFA MAAD HAMDI^{ID}, LUKMAN AUDAH^{ID}, AND SAMI ABDULJABBAR RASHID

Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, Johor 86400, Malaysia

Corresponding author: Lukman Audah (hanif@uthm.edu.my)

Communication of this research is made possible through monetary assistance by Universiti Tun Hussein Onn Malaysia and the UTHM Publisher's Office via Publication Fund E15216.

ABSTRACT Data dissemination in a VANETs network requires a meticulous process to ensure a high quality of service and eliminate hazardous conditions due to congestion or a broadcast storm. Considering multi-metric approaches and their implicit conflicting nature, it is necessary to handle this through effective multi-objective optimization algorithms. An effective optimization can be handled using a meta-heuristic approach with a high level of solution interactions. For this purpose, firefly was selected, which is a type of meta-heuristic search algorithm. Several developments of the firefly optimization were added to increase its capability to find more dominating solutions, namely, objective decomposition, archive management, and controlled mutation for exploration and exploitation balance. This developed multi-objective optimization was designated as adaptive jumping multi-objective firefly algorithm (AJ-MOFA). Afterwards, AJ-MOFA was integrated with a clustering and forwarding mechanism (CFM). This mechanism includes three main components. The first is clustering, which uses arbitration based on the cluster head score; the second is a forwarding mechanism that uses probabilistic forwarding and the third is AJ-MOFA. The solution space design in CFM combined two variables: the first is the probability of forwarding and the second is the maximum number of nodes within one cluster. The metrics to be incorporated in the multi-objective optimizations are the packet delivery ratio (PDR), the end-to-end delay (E2E-delay) and the number of dropped packets. Comparing both AJ-MOFA and CFM with benchmarks using multi-objective optimization and networking metrics reveals the superiority in most evaluation measures, which makes them promising algorithms for data dissemination in VANETs. The results showed an accomplished PDR of 60% and an E2E delay of 6.6 seconds, while the number of dropped packets was almost nine for the entire running time of the experiment, comparing a similar or lower performance of the benchmarks for these metrics.

INDEX TERMS VANETs, data dissemination, clustering, probabilistic forwarding, multi-objective optimization, firefly.

I. INTRODUCTION

Recently, the emergence of various technologies, such as the communication revolution, artificial intelligence, hardware speed, the cloud and the fifth generation (5G) has made it feasible to build reliable inter-connected ad hoc networks for vehicles using road environments [1]. Such networks are called vehicular ad-hoc networks, or VANETs [2]. The application of VANETs is part of the concept of building intelligent

transportation systems [3], enabling smart city functionalities [4] and delivering various services [5]. Important services enabled by VANETs are emergency handling services and hazard control systems [6]. Such situations require the management of data dissemination towards a certain node or subset of nodes in the network to enable the necessary action [7]. One issue with such data dissemination is its implied risk of bottlenecks or broadcast storms [8]. Another issue is the risk of delay, that is, not meeting the real-time constraints involved in handling an emergency [9]. Therefore, the logic of data dissemination must be meticulous and ensure

The associate editor coordinating the review of this manuscript and approving it for publication was Jad Nasreddine^{ID}.

all the requirements are met. These include minimising the incidence of collisions, avoiding the risk of bottlenecks and enabling fast delivery under real-time constraints.

Data dissemination requires the assurance of optimal routes for delivering messages to their destinations within a certain time interval. This is associated with more than one layer in the network, the most important of which are the MAC layer and the routing layer [10]. The role of the MAC layer is to assure the minimum possible number of collisions between messages, while the role of the routing layer is to ensure the shortest and least-loaded route to the destination in the shortest possible time [11]. The various aspects of optimization in the data dissemination, its relationship to the dynamics of both the message generation and vehicle mobility, as well as the fact that it causes fast changes in the network are considered the most challenging aspects of this problem. The authors add its strict requirements to these issues as data dissemination is used in an emergency context and is needed to ensure a safe situation [12]. Examples of this are the risk of accidents, the need to call an ambulance quickly and the need to ensure the fast delivery of medical support. Therefore, solving this problem requires the integration of various optimization models.

Optimization as a solution to various engineering problems has been used in various mathematical and computational approaches. Meta-heuristic optimization is a key computational approach in optimization. Its exploitation of the computational power of hardware and the incorporation of a heuristic search for solutions gives it the strength and effectiveness to obtain optimal solutions. The authors add its capability of solving non-linear and NP-hard optimization problems. Some examples of these are genetic optimization [13], particle swarm optimization [14], the ant bee colony algorithm [15], firefly [16], harmony search optimization [17] and Heterogenous Computing Task Scheduling Using Improved Harmony Search optimization [18]. However, all these challenging aspects are only one part of the problem. The other element that can be added to the problem is the multi-objective, which means that each solution is evaluated from more than one fitness function. Hence, it is insufficient to rank solutions with respect to only one objective and the vector of objective values must also be used to rank the solutions. The issue with enabling a good search function is finding the balance between exploration to enable diversity and exploitation to enable dominance.

The remainder of the article is organised as follows. In section II, the contributions are presented. Next, the literature survey is provided in section III. Afterwards, the methodology is outlined in section IV, while the experimental results and evaluation are presented in section V. Lastly, the summary and conclusion are provided in section VI.

II. CONTRIBUTION

This article provides the following contributions:

- 1) Novel mathematical formulation: This study provided a novel mathematical formulation of the problem of

data dissemination as a multi-objective optimization problem by considering the settings of the forwarding mechanism and the clustering parameters as the solution space, while the packet delivery ratio and E2E delay are considered the objective space. This formulation enables the problem to be solved based on a new approach that provides not only one optimal solution to the problem but a set of non-dominated solutions. Such solutions are useful for the decision-maker to determine which one is suitable for operating the network.

- 2) Novel framework: This study provides a novel data dissemination framework based on a clustering and forwarding mechanism to enable multi-objective optimization. The framework is designated as clustering and priority-based data dissemination (CPDD). The novel aspect of the framework is the multi-objective perspective, which enables the selection of one from a set of Pareto front solutions to operate the system.
- 3) The study developed a novel multi-objective optimization algorithm based on the framework of firefly by presenting the concepts of objective decomposition, archive management and controlled mutation to obtain a balance of exploration and exploitation as new modification features of the original multi-objective firefly. The developed algorithm is named adaptive jumping multi-objective firefly optimization, or AJ-MOFA.
- 4) The evaluation of AJ-MOFA is based on nine benchmarking mathematical functions, each with a multi-objective nature, and it was compared with other state-of-the-art approaches using multi objective optimization (MOO) metrics.
- 5) This study incorporated AJ-MOFA into CPDD and evaluated its performance in terms of both MOO and networking metrics. It was also compared with other state-of-the-art approaches.

Overall, from a practical perspective, the problem of data dissemination will be handled by the new algorithm, which provides a set of non-dominated solutions to decision-makers. Hence, they can select the operating point (optimal solution) that best suits their preferences. Consequently, more flexibility is provided to satisfy the priorities of the networking metrics.

III. LITERATURE SURVEY

This section provides the literature survey. It is decomposed of two sub-sections: The first one is data dissemination as MOO problem presented in A. The second one reviews of the various approaches developed under multi-objective firefly optimization in B.

A. DATA DISSEMINATION AS MOO PROBLEM

The literature contains numerous methods under Data Dissemination problem in VANETs. In the work of [19], a cluster-based MAC protocol based on IEEE 802.11 distributed coordinated function is proposed to support both safety and non-safety messages. The clustering concept is

based on grouping nearby vehicles in the same direction. In addition, control packets formats are changed to support cluster-based communications. Furthermore, the request to send (RTS) clear to send (CTS) mechanism are used for non-safety message delivery. The analysis was based on Markov chain model and a constraint on the delay of 100 [ms] for safety messages. This is due to the strict delay requirement of QoS [20]–[22]. Similarly, Markov chain analysis was used to evaluate the work developed in [23] orthogonal frequency division multiple access (OFDMA) is proposed. The protocol supports both the transmission mode and the optimal relay. Data transmission mode is divided into two categories—direct transmission and cooperative transmission. The usage of various artificial intelligence approaches was observed in recent VANETs studies. In the work of [24], collaborative learning-based routing scheme for multi-access vehicular edge computing environment has been proposed. Reinforcement learning based on end-edge-cloud collaboration to find routes in a proactive manner with a low communication overhead has been proposed. The learned information is used to change the route pre-emptively an integration with reactive approach was done. In the work of [25], two techniques were proposed. They are decentralized moving edge and multi-tier multi-access edge clustering. The former is used to meet the throughput and latency performance requirements and the latter was used for more efficient integration of different types of access technologies. Fuzzy logic was used for jointly considering multiple inherently contradictory metrics and Q-learning was used for achieving a self-evolving capability. According to the survey of [26] the development of meta-heuristic based approaches for handling the issues of VANETs in terms of routing is the most convenient to overcome them. The survey has presented plenty of approaches for routing in VANET with using one of the famous meta-heuristic approaches such as genetic optimization, particle swarm optimization, and ant bee colony. Meta-heuristic approaches were applied in VANETs in general and message dissemination in particular. For the former, we give an example for meta-heuristic based routing the work of [27] where multi-objective non dominated sorting genetic algorithm NSGA-II was proposed for optimizing OLSR protocol. Their algorithm was made to counter the dynamical changes in the environment as well as the vehicle density changes. The framework of using meta-heuristic optimization algorithm for OLSR routing is similar to the one proposed by [28]. The modification of their framework is just by reducing the solution parameters to Hello Interval, topology control (TC) Interval, and REFRESH Interval and replacing the single optimization algorithm that was presented in the original framework with multi-objective optimization. The algorithm proposed for fitness evaluation two evaluation measures: packet loss ratio and E2E delay. The handling of the multi candidate solutions proposed by the algorithm is suggested to be done by the user according to the nature of the application.

For data dissemination, [29] the quality of service (QoS) constraint has been resolved using firefly optimization. Levy

distribution was incorporated so the algorithm is called Levy distribution. The authors have proposed reducing the running time of the algorithm by identification of search patterns in the solutions. Many researchers present the multi-objective problem of VANET dissemination in single objective way which makes the solutions non-optimal due to the non-convexity of the problem. We give an example the work of [30] discrete particle swarm optimization was proposed for selecting optimal routes for messages that increases link stability and decreases obstacle occurrence. The optimization has proposed a single objective formulation that fuses both the link stability factor and the obstacle occurrence factor which makes the solution non-suitable for non-convex solution surface. Optimization for data dissemination was also developed for MAC layer such as the work of [24] where lion optimization algorithm where an adaptive selection scheme based on distance from nearby nodes and network density is proposed. The goal of the optimization is to identify the partition that relay the emergency message. Another goal of the work is to solve the problem of hidden terminal problem.

B. MULTI-OBJECTIVE FIREFLY

Developing a multi-objective firefly algorithm for various types of applications is a research interest. This has been observed in robotics path planning [31], VANET routing [32], sustainable energy [33], manufacturing [34], power systems optimization [35], etc. Some researchers have dealt with firefly as multi-objective algorithm using a model of weighted average for the various objectives such as the work of [32] where OLSR protocol was optimized for VANET routing. This has an issue of causing local minima because of the non-convexity of the optimization surface.

The literature contains several works in the development of multi-objective firefly optimization. In the work of [36] a new variant of firefly optimization was proposed where compensation learning and elite learning were included. In the elite learning, the non-dominated solutions were stored in an external archive and random set of them was selected as elites. This extends the movement of the solutions toward elites selected from the set of non-dominated solutions and stored in an external archive. In the compensation factor, the model of moving the fireflies toward each other is modified to include a compensation factor that accelerates the firefly toward the optimal Pareto by overpassing the mate firefly in the formula. The algorithm suffers from lacking of awareness of crowding and directions of solutions in the objective space which makes it with limited diversity. This concept of elite learning was also used in the work of [37] with applying it to robotics path planning. In another application of multi-objective firefly algorithm [38], an adaptive weight firefly formula has been derived where fireflies moves toward their mates with weighted average equation between the current position and the mate position. The weight is adaptive according to the iteration in order to enable convergence.

Another work that has developed a less recent variant of firefly optimization with multi-objective aspect is the work

of [39] where non-dominated comparison has been used and fireflies moves toward the solution that optimizes the weighted average of the objectives in case of non-domination. This has an issue of risk falling in local optima in case of non-convex optimization surface.

Another more recent development of multi-objective optimization is the work of [40] where a hybrid variant of the algorithm was proposed by incorporating the crossover of the differential evolution and an adaptive formula or the random factor.

The approach is effective for balancing diversity and exploitation; however, there was no special treating of the non-domination aspect of the multi-objective optimization. [41] have developed another multi-objective variant of firefly optimization with including the concept of decomposition of objectives, mutation, neighbor sampling and virtual force.

The approach was used for configuring radio frequency identification cards RFID readers. Hence, it is considered as planning deployment for RFID in the environment. The objectives were three; namely, level of interference, coverage and cost. Their finding is an outperformance of this variant of multi-objective firefly over non-dominated sorting genetic optimization and multi-objective particle swarm optimization. This approach suffers from the issue of lacking diversity awareness while generating new fireflies. In the work of [32], reputation based weighted clustering protocol for VANETs has been proposed using multi-objective optimization. The parameters of the clustering algorithm were used as solution in the searching algorithm. The objectives of the optimization are the cluster lifetime, improved packet delivery ratio and reduced cluster overhead. Their finding is that multi-objective firefly optimization algorithm outperforms both multi-objective particle swarm optimization and comprehensive multi-objective particle swarm optimization. In the work of [42], multi-objective firefly optimization was used for the goal of big data optimization. The algorithm was equipped with crossover strategy borrowed from different evolution.

IV. METHODOLOGY

This section outlines the developed methodology for the new adaptive jumping multi-objective firefly-based data dissemination. All the symbols and abbreviations are depicted in Tables 1 and 2.

A. MOFA AND CF-MOFA

This article uses two benchmarking algorithms based on the firefly algorithm, namely the multi-objective firefly algorithm (MOFA) [39] and the compensation factor firefly algorithm (CF-MOFA) [36]. Fundamentally, firefly is a meta-heuristics searching optimization algorithm. It was inspired by nature, more specifically, by the behaviour of fireflies. The algorithm is based on several assumptions. First, all fireflies are unisex, which means that they cannot be distinguished by gender, so they are attracted to each other. Second, the brightness factor is behind attractiveness, which means that a

TABLE 1. List of used parameters.

<i>leader</i>	selected solution from archive.
<i>f</i>	Objective function
<i>g*</i>	Is the optimal obtained by transforming multi-objective functions into single objective functions in the form of random weighted sum.
<i>t_{max}</i>	Maximum number of iterations
<i>β</i>	attractiveness
<i>γ</i>	the absorption coefficient
Max Gen	Maximum number of generation
<i>m</i>	the compensation factor
<i>θ(t)</i>	Exploration vs. exploitation control parameter
<i>d_f</i> and <i>d_l</i>	Distance between the extreme solutions and boundary solutions
\bar{d}	is the average of all the consecutive distances
<i>GD</i>	Generational distance
<i>P_s</i>	Pareto set
<i>P_T</i>	True pareto

TABLE 2. List of used abbreviations.

VANETs	Vehicular ad hoc networks
AJMOFA	adaptive jumping multi-objective firefly algorithm
CFM	clustering and forwarding mechanism
PDR	packet delivery ratio
E2E-delay	end-to-end delay
5G	fifth generation
MAC	Medium access control
MOO	multi-objective optimization
CPDD	clustering and probabilistic-based data dissemination
QoS	quality of service
OFDMA	orthogonal frequency division multiple access
OLSR	Optimized link state routing protocol
TC	topology control
RFID	Radio frequency identification
CHA	cluster head announcement
MaxCS	maximum allowed cluster size
Tmax	Maximum threshold

less bright firefly will move towards a brighter one. Third, the fitness function is the factor behind brightness, which means that more fitness value implies more brightness, considering that the problem is a maximisation problem. The original firefly algorithm was developed based on a single objective function. This is provided as Algorithm 1. As observed in the pseudo-code, the concept of the algorithm is to associate the fitness function of each firefly with the light intensity, which is considered the attractiveness factor of the firefly and causes other fireflies to move towards it if they themselves have less light intensity. The mobility equation is provided as Eq. (1) in [36].

$$x_i = x_i + m\beta(x_j - x_i) + \alpha(\varepsilon - \frac{1}{2}) \quad (1)$$

where

$$\beta = \beta_0 e^{-\gamma r_{ij}^2} \quad \text{denotes the attractiveness.}$$

$$\gamma \quad \text{denotes the absorption coefficient.}$$

$$r \quad \text{denotes the distance between } i \text{ and } j.$$

$$r = \sqrt{(x_{i,1} - x_{j,1})^2 + (x_{i,2} - x_{j,2})^2 + \dots + (x_{i,d} - x_{j,d})^2} \quad (2)$$

- d denotes the dimension of the objective space
- ε denotes a random number generated uniformly between 0 and 1
- β_0 denotes a constant which takes the value of 1
- α denotes a constant that takes a value in the range between 0 and 1
- m denotes the compensation factor

The algorithm of [36] uses the mobility model given in Eq. (3).

$$x_i(t + 1) = c_1 g^* + c_2 leader + \alpha \varepsilon_i \quad (3)$$

where

- $leader$ denotes randomly selected solution from the external archive
- g^* denotes an optimal solution obtained by transforming multi-objective functions into single objective function

It was observed from the algorithm models that g^* is used for elites learning, which is not effective when the optimization is non-convex, while the leader is selected without an objective decomposition aspect at the beginning, which makes the exploration performance poor.

B. AJ-MOFA

This section provides the development of adaptive multi-objective firefly optimization. A general flowchart is depicted

Algorithm 1 Pseudocode of Firefly Algorithm

Input

f //Objective function

Output

x //optimal solution

Start

- 1: initiate the fireflies population
 - 2: evaluate the fireflies population according to f
 - 3: calculate the light intensity I for each firefly
 - 4: for $t = 1$ until $maxGeneration$
 - 5: for each two pairs of fireflies i and j
 - 6: calculate distance between i and j
 - 7: calculate attractiveness
 - 8: if ($I(i) > I(j)$)
 - 9: move j toward i based on attractiveness in Eq (1)
 - 10: else
 - 13: move i toward j based on attractiveness Eq (1)
 - 14: end
 - 15: update light intensity I
 - 16: end
 - 17: end
 - 18: **End**
-

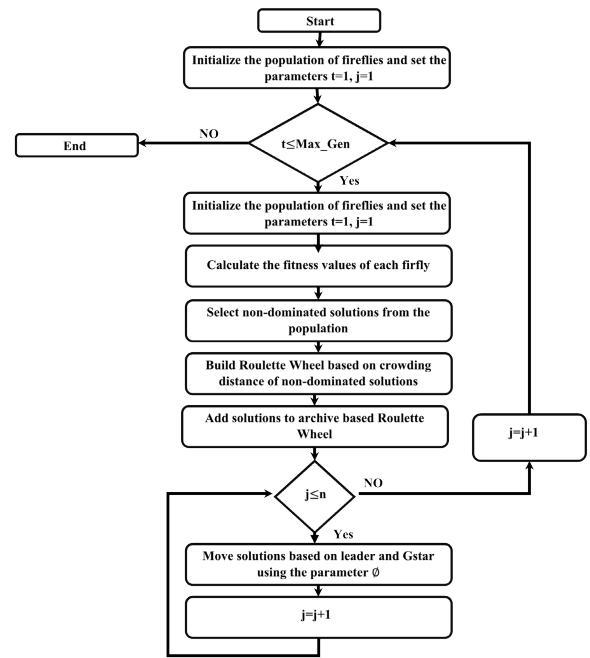


FIGURE 1. General flowchart of adaptive jumping multi-objective firefly algorithm (AJ-MOFA).

in Figure 1. The algorithm starts by initialising the population of fireflies and the parameters of the algorithm. The algorithm iterates until it reaches the maximum number of generations Max_Gen . At each iteration, the algorithm calculates the fitness values of each firefly. Next, it selects the non-dominated solutions from the population. Then, it builds the roulette wheel, based on the crowding distance of non-dominated solutions. The roulette wheel is used for selecting solutions that will be removed from the archive. Next, the algorithm goes through the solutions sequentially and moves each one according to the mobility model of the solutions, which uses the ε -greedy objective decomposition to find a balance between exploration and exploitation. In addition, the mutation is applied based on the adaptive mutation rate.

The algorithm focuses on three aspects of development. Firstly, the ε -greedy objective decomposition is presented in A. Next, archive management is outlined in B. Lastly, the exploration-exploitation trade-off is described in C.

1) OBJECTIVE DECOMPSITION

The algorithm includes a mechanism for enabling objective decomposition based on the parameter θ . The model of the parameter is given in Eq. (4).

$$\theta(t) = 1 - \frac{t}{t_{max}} \quad (4)$$

where:

- t denotes the index of iteration
- t_{max} denotes the maximum number of iterations

At the beginning of the search, t is low, which gives a θ value close to 1. This enables a move to branch 1 in the

Algorithm 2 Pseudocode Control Parameter of the Jumping Behavior

```

1: if rand <  $\theta(t)$ 
2: Sol = Sol + jumpingCF1 * rand (size (Sol)) *
   (Gstar-Sol)
3: else
4: Sol = Sol + jumpingCF2 * rand(size(Sol)) *
   (Leader-Sol)
5: End
    
```

jumping behaviour, which is given in Algorithm 2. Hence, the algorithm moves the solutions towards Gstar, which represents the best solution to a randomly selected objective and considers the uniform distribution of the probability density function. Consequently, the algorithm will be able to explore the boundary of the solution space. Meanwhile, at the end of the search, Algorithm 2 will increase and become closer to t_{max} , which will cause the value of θ to decrease.

2) ARCHIVE MANAGEMENT

The management of the archive has the role of removing the solutions that have less potential and preserving the solutions that have high potential. The algorithm generates the solutions that are to be removed, based on a roulette wheel in Eq. (5)

$$IdxsToRemove = rouletteWheel \left(\frac{1}{crowdDist} \right) \quad (5)$$

where:

$crowdDist$ denotes the crowding distance between the adjacent solution.

The crowding distance [43] can be written in Eq. (6)

$$crowdDist(i) = \sum_{j=1}^m \frac{I(next(i, m), m) - I(prev(i), m)}{f_{max}^j - f_{min}^j} \quad (6)$$

where

$next(i, m)$ denotes the next solution to i with respect to objective j

$prev(i, m)$ denotes the previous solution to i with respect to objective j

f_{max}^j denote the maximum value of objective j

Algorithm 3 Controlling the Tradeoff Between Exploration and Exploitation

```

1: if rand < exp-( iter -1/(iter-maxIteration))
2: sol = unifrnd(sol)
3: if sol is not improved
4: sol = normrnd (sol,(ub-lb)*0.05)
5: end
5: end
    
```

f_{min}^j denote the minimum value of objective j
 m denotes the number of objectives

$$crowdDist(1) = crowdDist(l) = \infty$$

where 1 and l denoted the boundary solutions.

The algorithm gives a greater chance of preserving solutions with a higher crowding distance because it reduces the value of the removal probability when the crowding distance is higher. Hence, this approach to archive management features better exploration.

Algorithm 4 The VANET Clustering

Input:

cluster head index

Output:

Boolean variable that takes 1 if the node was successfully assigned to cluster, 0 otherwise

Start

```

1: node.timer = initiate time(cluster head index)
2: node.flag = 1;
3: while not (receive CHA or timer == 0)
4:     wait
5: end
6: if(receive CHA)
7:     node.flag = 0;
8:     send CMR message to cluster head
9:     initiate timer 2
10:    while not (timer 2 is zero or received
        confirmation)
11:        wait
12:    end
13:    if (received confirmation)
14:        change state to CM
15:        return 1;
16:    else
17:        return 0
18:    end
19: else
20:    broad CHA
21:    initiate timer 3
22:    while not (timer 3 is zero or received
        confirmation)
23:        wait
24:    end
25:    if (received confirmation)
26:        reply with CMR
27:        return 1;
28:    else
29:        return 0
30:    end
31: End
    
```

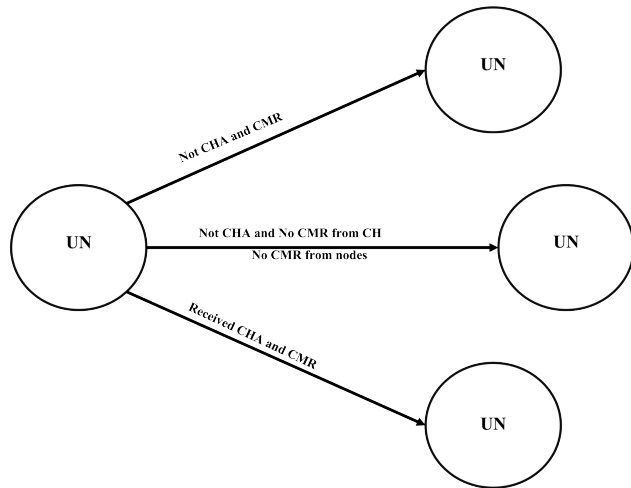


FIGURE 2. State transition diagram of the clustering process. Cluster Head (CH), undefined node (UN), cluster member (CM).

3) EXPLORATION AND EXPLOITATION TRADEOFF BASED ON MUTATION

The mutation was performed at a higher rate at the beginning of the search in order to achieve better exploration; however, at the end of the search, there was a tendency to reduce the probability of exploration using mutation in order to enable the convergence of Algorithm 3.

Algorithm 5 The Message Forwarding Process

```

Input:
Environment
Output:
Action of message of forwarding
Start
1: initialize the network by clustering scheme: (1-neighbor
   discovery 2-cluster head election 3- cluster formation)
2: for each node n
3:   if n received a Message
4:     if n is Cluster Head
5:       if (n in the message transmission direction)
6:         while (the message hadn't been received
   from the other direction)
7:           Broadcast the message
8:         end while
9:       end if
10:    elseif (n in the message transmission direction)
11:      if n's CH can't receive it
12:        Send the packet to the CH based on
   Probabilistic forwarding algorithm
13:      end if
14:    end if
15:  end if
16: end for
17: End
    
```

C. CLUSTERING AND PROBABILISTIC-BASED DATA DISSEMINATION (CPDD)

The forwarding mechanism developed by the authors is based on integrating multi-objective optimization with a clustering and forwarding mechanism. Hence, the algorithm contains three main components. The first is clustering, which uses arbitration based on the cluster head score. The second is the forwarding mechanism, which uses probabilistic forwarding, while the third is multi-objective optimization, which is based on adaptive jumping firefly optimization. Each of the three components is presented in separate sub-sections, as follows:

1) CLUSTERING

The clustering is based on arbitration [7]. The pseudocode of the clustering algorithm is given as Algorithm 4. The input is the cluster head index. The output is the clustering results, defined by assigning to each node its cluster head. The operation of the code is as follows. It starts by initiating a timer for each node according to the value of its cluster head index. Next, the node initiates a flag with the value of 1 to indicate that the node is in the process of arbitration. The flag will only return to 0 when the node receives a cluster head announcement (CHA); otherwise, the flag stays at 1 to indicate that the node has not received any cluster head announcement message, which would mean that the node is eligible to broadcast its own CHA to other nodes. When it receives the CHA from another cluster, the node performs the

Algorithm 6 The Forwarding Mechanism

```

Input
M // the message to be forwarded
Tmax //the threshold used for stopping counting
T Interval
PH // the probability of forwarding the messages that are
   received for one time only
Output
P
Start
1: C = 0
2: P = 1
3: Tc = T
4: While (Tc > 0)
5:   If (the cluster member receives message M)
6:     C = C+1
7:   end if
8: end While
9: If (C > Tmax)
10:  P = 0;
11: elseif (C == 1)
12:  P = PH
13: else
14:  P =  $\frac{2}{C^2}$ 
15: end if
16: End
    
```

TABLE 3. The solution space that is used for the optimization.

MaxCS	maximum allowed cluster size
Tmax	Maximum threshold
PH	the probability value for forwarding the packet that received only once

TABLE 4. The optimization objectives.

PDR	Packet Delivery ratio
E2E Delay	Average End to End Delay
AvrNDRP	Average number of doubled received packets reception times

joining process to the cluster, with which it is combined in two steps. Firstly, it sends a confirmation message; secondly, it triggers another time to wait for a final confirmation from the cluster head. The goal of the last timer is to prevent the maximum number of nodes allowed in the cluster from being exceeded. A state transition diagram for the clustering process is presented in Figure 2.

2) MESSAGE FORWARDING PROCESS

The message forwarding process is depicted in the pseudocode in Algorithm 5. As shown, after initiating the network with the clustering scheme and in two cases, each message receipt is distinguished before it is re-forwarded. The first case is when the receiver is the cluster head, and it then broadcasts on the condition that the message was sent in the same node direction and it had not been received previously from another direction. The second case is when the receiver is not the cluster of the node but only forwards it based on the probabilistic forwarding mechanism when the node cluster head cannot receive it.

3) FORWARDING MECHANISM

The algorithm used for the forwarding mechanism is depicted in Algorithm 6. The inputs are the message M; Tmax, which represents the threshold for stopping the count; and the time interval that is used for counting. The output is the forwarding probability that the node uses for forwarding. As provided in Algorithm 6, the algorithm starts by initiating the counter C with zero, the probability of forwarding with 1 and the timer with T. Next, the algorithm iterates for each count of the timer Tc, checks the receipt of the message M and increments the counter by one if the message is received again. The counter is then used to update the probability, based on giving a value of 1 for C = 1 and a value of 0 when C is higher than the threshold Tmax. Otherwise, the probability will take the value of $\frac{2}{C^2}$ which indicates a decreasing function with respect to the count.

4) MULTI-OBJECTIVE OPTIMIZATION

The optimization conducted was based on its solution space and objective space. For the solution space, the three decision variables are presented in Table 3. The optimization objectives are presented in Table 4. $x = (MaxCS, Tmax, PL)$

and the objective space $f = (PDR, E2Edelay, AvrNDRP)$. Mathematically, $x \in R^3$ and $y = R^3$.

The interaction between AJ-MOFA and the other parts of the data dissemination is presented in Figure 3. It consists of two paths, namely a forward path that is responsible for providing the decision variables and a backward path that is responsible for providing the objective functions.

5) MULTI-OBJECTIVE EVALUATION

This section provides an explanation of the MOO evaluation measures [44]. Some of these are related to the optimality of the solutions, while others are related to the spread of solutions.

- 1) Set coverage: This section indicates to the domination of one Pareto generated from one approach on the other. Assuming that p_{s1}, p_{s2} are two Pareto, then the set coverage or c measure of p_{s2} is the percentage of solutions of p_{s2} that are dominated by solutions from p_{s1} over the total number of solutions in p_{s2} . The goal is to minimize this measure in Eq. (7) in [44].

$$c(p_{s1}, p_{s2}) = \frac{|\{y \in p_{s2} \mid \exists x \in p_{s1}: x > y\}|}{|p_{s2}|} \quad (7)$$

- 2) Hypervolume: This measure indicates the volume of the dominated portion of the objective space over the worst solutions, which are regarded as reference solutions for calculating the hypervolume. Hence, this is calculated as the union of a hypercube whose diagonal is the segment that connects the worst solution and a point in the Pareto front. The goal is to maximise the hypervolume, which indicates the spread of solutions in the objective space. This is given in Eq. (8) in [44].

$$HV = volume \left(\bigcup_{x \in p_s} Hyper\ Cuber(x) \right) \quad (8)$$

- 3) Number of non-dominated solutions: This measure indicates to how much rich is the results of the optimization in terms of the non-dominated solutions. It is measured by the cardinal of the Pareto front. It is given in Eq. (9) in [44].

$$NDS(N) = |P_S| \quad (9)$$

- 4) Delta measure: The delta metric is given in Eq. (10). Its role is to measure how far the solutions are equally distributed in the Pareto front, and it is preferable that it is minimised in [44].

$$\Delta = \frac{d_f + d_l + \sum_{i=1}^{N-1} |d_i - \bar{d}|}{d_f + d_l + (N - 1)\bar{d}} \quad (10)$$

where N is the number of solutions.

d_f, d_l denote the distance between the extreme solution and the boundary solution

d_i denotes the distance between two consecutive solution

\bar{d} is the average of all the consecutive distances d_i for $i = 1, 2, \dots, N - 1$

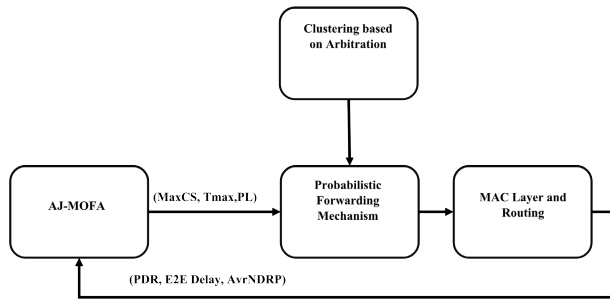


FIGURE 3. Flowchart of the interaction between AJ-MOFA, the clustering and the probabilistic forwarding mechanism.

A graphical representation of the calculation of the delta metric is shown in Figure (4).

- 5) Generational distance: The generational distance assumes the knowledge of the true Pareto front. This enables the calculation of the deviation of the found results from the true Pareto front. It is calculated based on the Eq. (11) in [44].

$$GD(P_s, P_T) = \frac{\sqrt{\sum_{i=1}^{|P_s|} d_i^2}}{|P_s|} \quad (11)$$

d_i^2 denotes the distance between the solution and the nearest solution in the true Pareto front

V. EXPERIMENTAL RESULTS AND EVALUATION

This section provides the experimental results and evaluation. For the evaluation, MATLAB 2019b was used to implement the developed algorithm and simulate the testing scenarios.

The evaluation is composed of two sub-sections: in sub-section A, the results of multi-objective optimization based on mathematical functions are presented, while in sub-section B, the evaluation of multi-objective-based data dissemination is provided.

A. RESULTS OF MULTI-OBJECTIVE OPTIMIZATION BASED ON MATHEMATICAL FUNCTIONS

The first part of the evaluation was to compare the work.

The valuation was performed on nine mathematical problems used for benchmarking [44]. We present their mathematical formulas in Table 5.

The baseline approaches are non-dominated sorting genetic algorithm (NSGA-II) [43] MOFA [39], and CF-MOFA [36].

For each problem, the set coverage, t-test of the set coverage, delta-metric, generational distance, hypervolume and number of non-dominated solutions were generated. Furthermore, the Pareto front was also illustrated. Each metric was generated in a statistical manner using ten runs and is represented as a boxplot diagram.

1) PARETO FRONT

The Pareto front provides the set of non-dominated solutions after the search. Figure 5 shows the Pareto front generated

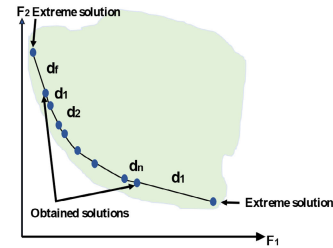


FIGURE 4. Calculation of delta measure.

by the different algorithms. Considering that the problem is a minimization problem, it is preferred to have a Pareto front with lower values of Objective 1 and Objective 2. We find in Figure 5 -a- that the four algorithms have provided approximately the same pareto front. However, this is not applicable to other sub-figures. In sub-figure 5 -b-, AJ-MOFA was more capable of providing lower values of objectives compared with other benchmarks. This is also monitored in sub-figure -e- and -f- where both AJ-MOFA and NSGA-II have reached more optimal Pareto front compared with other benchmarks. In other sub-figures, we find that NSGA-II has provided more optimal Pareto front than other benchmarks such as sub-figure -g- and -h-. We notice that both AJ-MOFA and NSGA-II were capable of converging towards the true Pareto, compared with the other two benchmarks. This was also observed in a more challenging function with discontinuity, such as KUR. Similarly, in some functions, NSGA-II and AJ-MOFA performed similarly, demonstrating superiority over CF-MOFA and MOFA, such as KUR, ZDT1, ZDT2 and ZDT3. In FON and in POL, all the algorithms converged towards the same Pareto front. This can be explained by the role of the jumping behaviour in AJ-MOFA on one side and the importance of objective decomposition when selecting Gstar.

2) SET COVERAGE

Previous Pareto front graphs show the general performance of the convergence of the optimization algorithm towards the true Pareto front. However, they do not provide quantitative inter-comparison in terms of the dominance percentage. Hence, the authors utilised the set coverage, which compared the approaches interchangeably in terms of the domination percentage of the results. The first part of the evaluation includes a comparison of the domination aspect of the original MOFA, the compensated factor CF-MOFA, AJ-MOFA and NSGA-II. The aim of developing the AJ-MOFA algorithm was for it to be compared with the other three approaches. Hence, six values of the set coverage are presented, each of which represents either C (AJ-MOFA, method) or C (method, AJ-MOFA). As observed in Figure 6, AJ-MOFA accomplished higher domination over the three methods in the boxes from 1 to 6. The value of AJ-MOFA had the value 0.25 of the set coverage over MOFA, compared with the near-zero value of the set coverage for MOFA over AJ-MOFA. Similarly, the set coverage of AJ-MOFA over

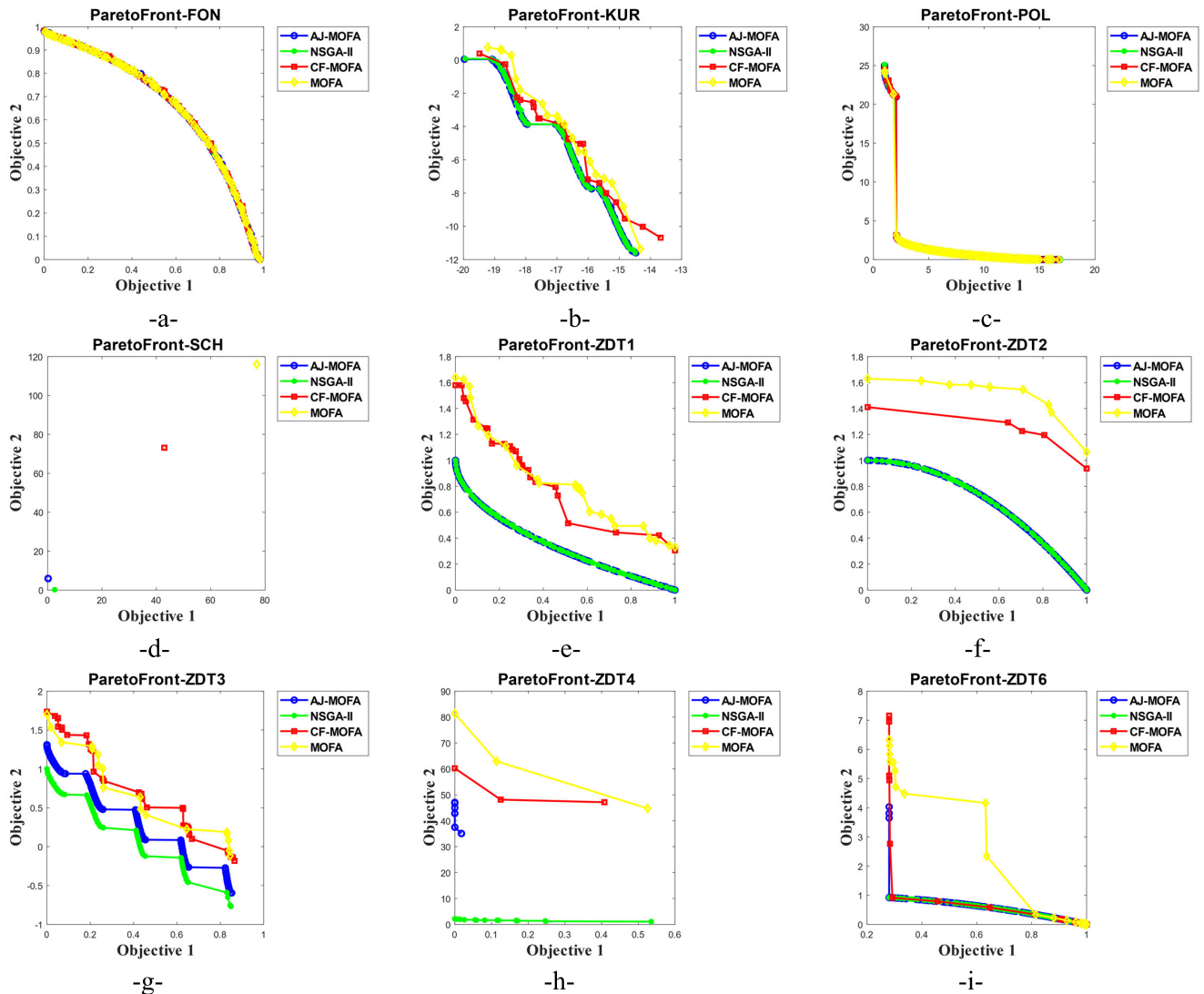


FIGURE 5. Pareto front of the 9 mathematical problem for each of the developed algorithms a- FON, -b- KUR, -c- POL, -d- SCH, -e- ZDT1, -f- ZDT2, -g- ZDT3, -h-ZDT4,-i- ZDT6.

CF-MOFA was also near 0.25, compared with the near 0.1 value of the set coverage of CF-MOFA over AJ-MOFA. In boxes 7 and 8, it can be observed that both CF-MOFA and MOFA displayed similar performances, with MOFA having marginal superiority. However, it was observed that NSGA-II outperformed both MOFA and CF-MOFA in terms of domination from boxes 9 to 12.

The second function that was evaluated was KUR. The set coverage, as presented in Figure 7, shows a similar pattern of performance when compared with FON. However, it was noted that the percentage of the superiority of AJ over both MOFA and CF-MOFA was higher than its corresponding values of FON (it reached a percentage of 95%). Similarly, it was observed that the dominance percentage of AJ over NSGA-II, with respect to KUR, was higher than its corresponding value of FON.

For more elaboration, we present the numerical values of the descriptive statistics (max, min, first and third quartile,

and median) that are derived from the boxplot figures of set coverage for both FON and KUR problems in Tables 6 and 7 respectively. The observation is clear superiority in domination of AJ-MOFA over each of CF-MOFA and MOFA and slightly better domination of NSGA-II over AJ-MOFA. This interpreted by the evolutionary nature of searching of NSGA-II compared with the swarm nature of AJ-MOFA. However, AJ-MOFA was better in attaining more dominant solutions over the other algorithms form its family, namely, MOFA and CF-MOFA.

3) SET COVERAGE T-TEST METRIC

The t-test was used to test the superiority of AJ-MOFA over the other benchmarks. The rejection of the null hypothesis h_0 was tested. When it was rejected, the alternative hypothesis h_a was then confirmed. This was achieved using t-test values, which had to be lower than 0.05. The results in Table 8 demonstrate that the comparison between the set coverage

TABLE 5. Descriptions of test problems.

Problem	n	Variable bounds	Objective functions	Optimal solutions	Comments
FON	3	$[-4,4]$	$f_1(x) = 1 - \exp\left(-\sum_{i=1}^3\left(x_i - \frac{1}{\sqrt{3}}\right)^2\right)$ $f_2(x) = 1 - \exp\left(-\sum_{i=1}^3\left(x_i - \frac{1}{\sqrt{3}}\right)^2\right)$	$x_1 = x_2 = x_3$	Non-convex
KUR	3	$[-5,5]$	$f_1(x) = \sum_{i=1}^{n-1}(-10 \exp(-0.2\sqrt{x_i^2 + x_{i+1}^2}))$ $f_2(x) = \sum_{i=1}^n(x_i ^{0.8} + 5 \sin x_i^2)$	Refer to [45]	Non-convex
ZDT1	30	$[0,1]$	$f_1(x) = x_1$ $f_2(x) = g(x) \left[1 - \sqrt{\frac{x_1}{g(x)}}\right]$ $g(x) = 1 + [9(\sum_{i=2}^n x_i) / (n - 1)]^{0.25}$	$x_1 \in [0,1]$ $x_i = 0$ $i = 2,3, \dots, n$	Convex
ZDT2	30	$[0,1]$	$f_1(x) = x_1$ $f_2(x) = g(x) \left[1 - \left(\frac{f_1(x)}{g(x)}\right)^2\right]$ $g(x) = 1 + [9(\sum_{i=2}^n x_i) / (n - 1)]^{0.25}$	$x_1 \in [0,1]$ $x_i = 0$ $i = 2,3, \dots, n$	Non-convex
ZDT3	30	$[0,1]$	$f_1(x) = x_1$ $f_2(x) = g(x) \left[1 - \sqrt{\frac{x_1}{g(x)}} - \frac{x_1}{g(x)} \sin(10\pi x_1)\right]$ $g(x) = 1 + 9(\sum_{i=2}^n x_i) / (n - 1)$	$x_1 \in [0,1]$ $x_i = 0$ $i = 2,3, \dots, n$	Convex, disconnected
ZDT6	10	$[0,1]$	$f_1(x) = 1 - \exp(-4x_1) \sin^6(10\pi x_1)$ $f_2(x) = g(x) \left[1 - \left(\frac{f_1(x)}{g(x)}\right)^2\right]$ $g(x) = 1 + [9(\sum_{i=2}^n x_i) / (n - 1)]^{0.25}$	$x_1 \in [0,1]$ $x_i = 0$ $i = 2,3, \dots, n$	Convex, Non-uniformly spaced

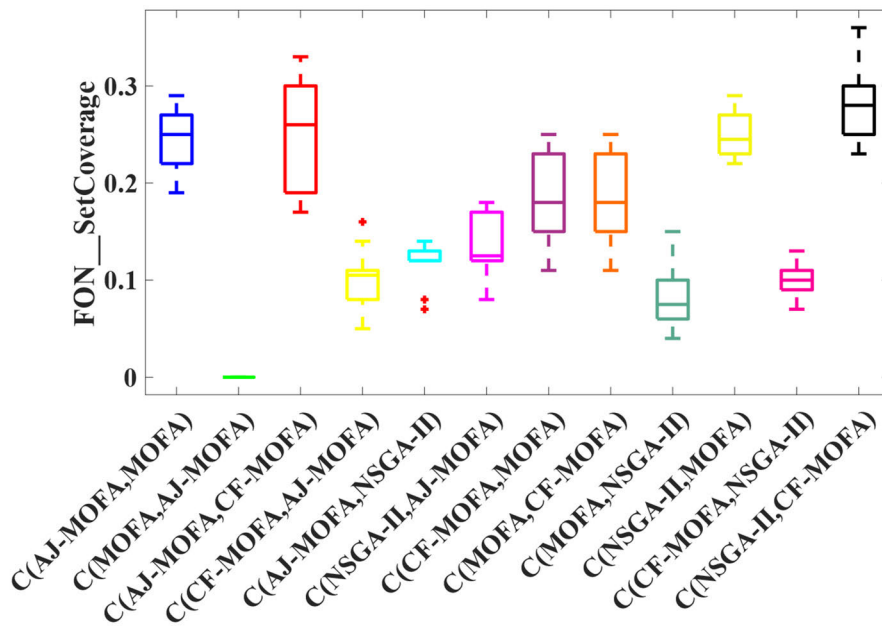


FIGURE 6. Set coverage comparison of FON between AJ-MOFA and benchmarks.

of AJ-MOFA and those of the three benchmarks (MOFA, CF-MOFA and NSGA-II) for the FON functions shows that h_0 was rejected for MOFA and CF-MOFA. This illustrates the statistical significance of the superiority over them because the t-test value was lower than 0.05. However, for the FON function, there was no statistical significance of the superiority of the domination of CF-MOFA over NSGA-II.

In addition to FON, the t-test values of the set coverage of AJ-MOFA compared with the benchmarks for KUR

are shown in Table 8. Table 9 shows that AJ-MOFA accomplished statistical significance over the three benchmarks MOFA, CF-MOFA and NSGA-II because the t-test values were lower than 0.05 for all the set coverage tests.

4) GENERATIONAL DISTANCE

The second metric used for comparison between the two approaches was the generational distance, which indicates the

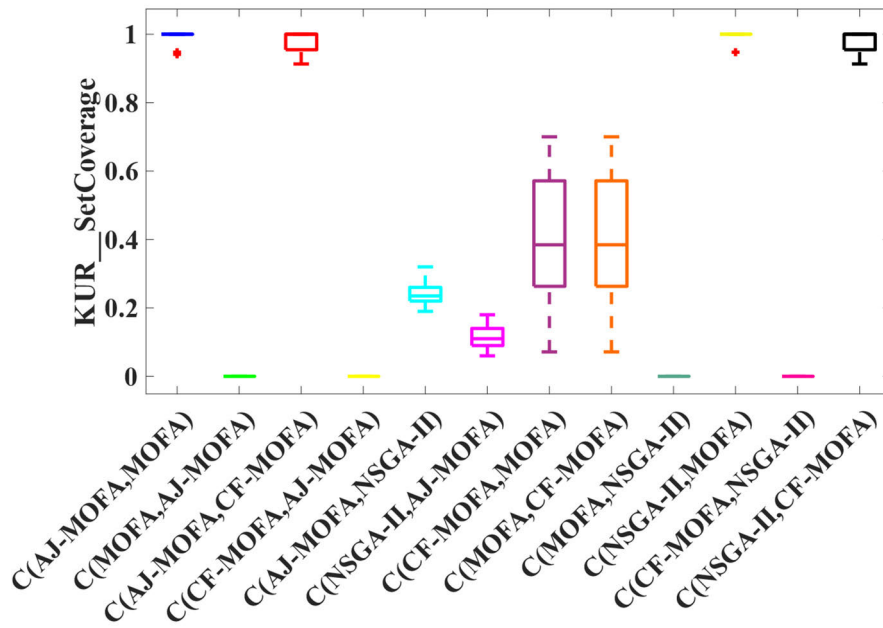


FIGURE 7. Set coverage comparison of KUR between AJ-MOFA and benchmarks.

TABLE 6. Descriptive statistics of set coverage of AJ-MOFA set coverage compared with benchmarks for FON.

Set Coverage	C(AJ-MOFA,MOFA)	C(MOFA,AJ-MOFA)	C(AJ-MOFA,CF-MOFA)	C(CF-MOFA,AJ-MOFA)	C(AJ-MOFA,NSGA-II)	C(NSGA-II,AJ-MOFA)
Min	0.19	0	0.17	0.05	0.07	0.08
Max	0.29	0	0.33	0.16	0.14	0.18
1st Quartile	0.22	0	0.19	0.08	0.12	0.12
Median	0.25	0	0.26	0.105	0.12	0.125
3th Quartile	0.27	0	0.3	0.11	0.13	0.17

TABLE 7. Descriptive statistics of set coverage of AJ-MOFA set coverage compared with benchmarks for KUR.

Set Coverage	C(AJ-MOFA,MOFA)	C(MOFA,AJ-MOFA)	C(AJ-MOFA,CF-MOFA)	C(CF-MOFA,AJ-MOFA)	C(AJ-MOFA,NSGA-II)	C(NSGA-II,AJ-MOFA)
Min	1	0	0.91304	0	0.06	0.071429
Max	1	0	1	0	0.18	0.7
1st Quartile	1	0	0.95455	0	0.09	0.26316
Median	1	0	1	0	0.11	0.38445
3th Quartile	1	0	1	0	0.14	0.57143

TABLE 8. t-test comparison between our set coverage of our developed AJ-MOFA with the benchmarks for FON function.

T-test	C(AJ-MOFA, MOFA), (MOFA, AJ-MOFA)	C(AJ-MOFA, CF-MOFA), (CF-MOFA, AJ-MOFA)	C(AJ-MOFA, NSGA-II), (NSGA-II, AJ-MOFA)
FON	2.36E-09	0.0001614	0.2267

TABLE 9. t-test comparison between our set coverage of our developed AJ-MOFA with the benchmarks for KUR function.

T-test	C(AJ-MOFA, MOFA), (MOFA, AJ-MOFA)	C(AJ-MOFA, CF-MOFA), (CF-MOFA, AJ-MOFA)	C(AJ-MOFA, NSGA-II), (NSGA-II, AJ-MOFA)
KUR	3.948EE-16	5.9822E-15	0.0001651

distance between the found Pareto by the approach and the true Pareto.

The generational distances of AJ-MOFA, CF-MOFA, NSGA-II and MOFA for FON function are depicted

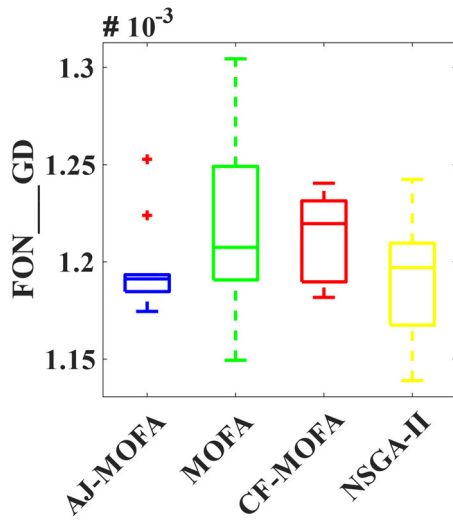


FIGURE 8. Generational distance comparison of FON between AJ-MOFA and benchmarks.

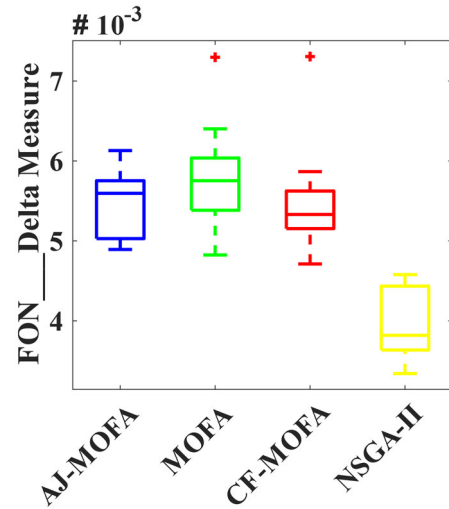


FIGURE 10. Delta metrics comparison of FON between AJ-MOFA and benchmarks.

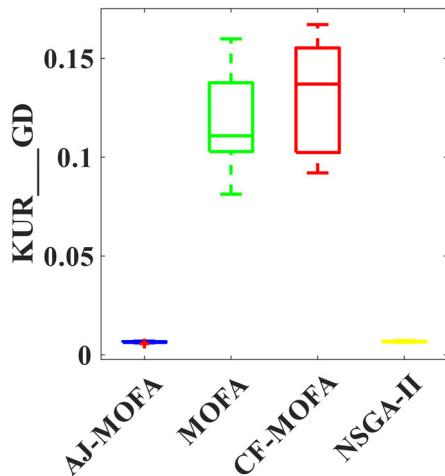


FIGURE 9. Generational distance comparison of KUR between AJ-MOFA and benchmarks.

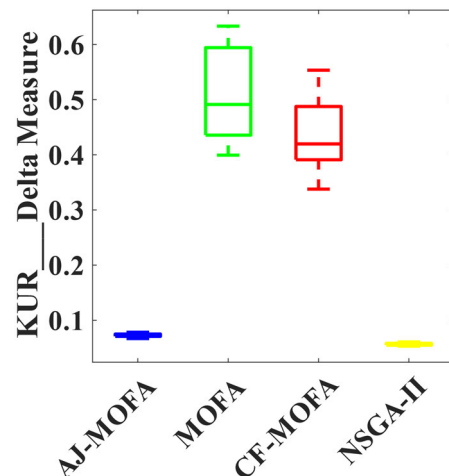


FIGURE 11. Delta metrics comparison of KUR between AJ-MOFA and benchmarks.

in Figure 8. AJ-MOFA has achieved lower maximum value, 75th percentile and whisker than the other benchmarks. The lower values of these elements in GD are indicator to superiority. On the other side, we find that NSGA-II could achieve lower minimum values compared with other algorithm which was the only aspect of superiority over AJ-MOFA. As can be observed, the generational distance of AJ-MOFA is lower than its corresponding values for the benchmarks, which reveals the superiority of the former over the latter. This can also be observed in Figure 9, for KUR function for the two approaches MOFA and CF-MOFA, while NSGA-II provided a low generational distance for KUR function, which was similar to AJ-MOFA in that regard. As shown in Figure 8, the generation distance was between 1.18 and 1.2 for AJ-MOFA, which is lower than all benchmarks. Similarly, the generational distance of less than 0.01 was provided for AJ-MOFA, compared with the benchmarks' values that were higher than 0.1, as shown in Figure 9. Observing the differences between

the different algorithms in terms of generational distance, we find that both AJ-MOFA and NSGA-II have generated almost identical values of minimum, maximum, median, 25th and 75th percentile which is an indicator to more stability and better performance compared with other algorithms.

5) DELTA METRIC

The role of the delta metric is to verify the capability of searching for and finding equally distributed non-dominated solutions for its Pareto front. A lower value of delta is equivalent to better diversity and equal distributed functions.

In Figure 10, it can be observed that AJ-MOFA generated similar values of delta to MOFA and CF-MOFA in terms of the minimum values. However, AJ-MOFA has accomplished lower median value than MOFA. On the other side, we find that NSGA-II has provided the least values of minimum, maximum, median, 25th and 75th percentile. Figure 11 shows, NSGA-II and AJ-MOFA generated the same values of delta,

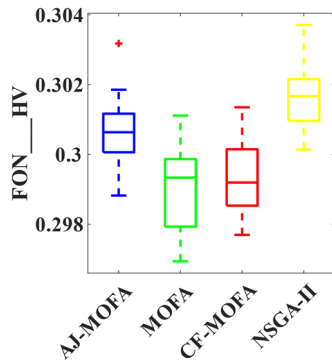


FIGURE 12. Hypervolume comparison of FON between AJ-MOFA and benchmarks.

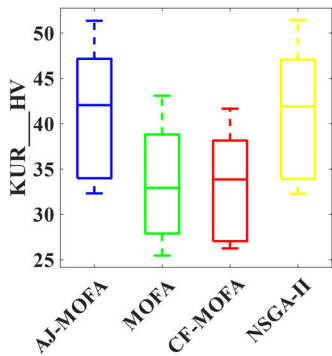


FIGURE 13. Hypervolume comparison of KUR between AJ-MOFA and benchmarks.

which were lower than those of the two benchmarks. Furthermore, they provide almost identical values of minimum, maximum, median, 25th and 75th percentile which is an indicator to more stability in addition to superiority.

6) HYPERVOLUME

The other metric that is generated is the hypervolume that indicates to the diversity of the solutions with respect to the objective space. The higher the values of the hypervolume are equivalent to better performance. This measure is secondary after the domination measures. We observe in Figure 12, that the highest hypervolume that has been accomplished is for

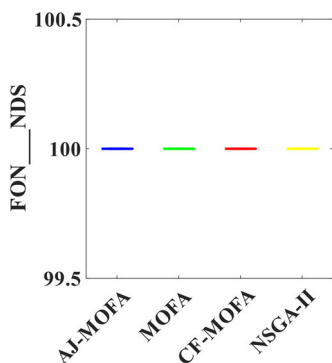


FIGURE 14. Number of non-dominated solutions comparison of FON between AJ-MOFA and benchmarks.

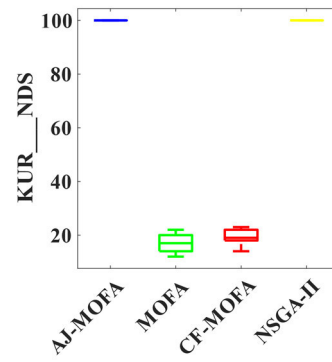


FIGURE 15. Number of non-dominated solutions comparison of KUR between AJ-MOFA and benchmarks.

both AJ-MOFA and NSGA-II which indicates to the diversity of the searching in FON function. In addition, we find that AJ-MOFA has accomplished minimum value of hypervolume higher than the minimum values of MOFA and CF-MOFA. Also, the median value of hypervolume of AJ-MOFA is higher than the median value of MOFA and CF-MOFA. For KUR function, we observe that AJ-MOFA has accomplished similar value of hypervolume comparing with NSGA-II and higher than the other benchmarks MOFA and CF-MOFA as show in Figure 13. Furthermore, we find that both AJ-MOFA and NSGA-II have the same minimum, maximum, median, 25th and 75th percentile which shows that they are equal in terms of the availability of choices with respect to the objectives.

7) NUMBER OF NON-DOMINATED SOLUTIONS

The last metric that was generated was the number of non-dominated solutions. In Figure 14, we present the number of non-dominated solutions of NDS for AJ-MOFA and the benchmarks for FON. It can be observed that all the approaches generated the same number of non-dominated solutions, 100 solutions (the same number as the size of the population).

In addition, we observed that for KUR function, as shown in Figure 15, both AJ-MOFA and NSGA-II accomplished the highest values of the number of non-dominated solutions, compared with the two benchmarks, MOFA and CF-MOFA. The value of NDS for AJ-MOFA was 100 in terms of minimum, maximum, median, 25th and 75th percentile, while it was lower than 20 for MOFA and CF-MOFA with different values of minimum, maximum, median, 25th and 75th percentile.

8) SUMMARY OF EVALUATION RESULTS AND DISCUSSION

Observing Table 11, it can be seen that AJ-MOFA achieved superior results over the benchmarks for the majority of the functions. The set coverage indicated the dominance of the approach over the benchmarks. The only function for which NSGA-II was superior to AJ-MOFA was ZDT4. Another observation is that NSGA-II was the second least superior after AJ-MOFA. Another aspect of the performance was the

TABLE 10. Sorting the algorithms according to their superiority with respect to the various evaluation metrics from best to worst.

ObjFun	SetCoverage	GD	HV	NDS	Delta
SCH1	AJ-MOFA	AJ-MOFA	NSGA-II	AJ-MOFA	CF-MOFA, MOFA
	NSGA-II	NSGA-II	AJ-MOFA	NSGA-II	AJ-MOFA
	CF-MOFA	CF-MOFA	CF-MOFA	CF-MOFA, MOFA	NSGA-II
	MOFA	MOFA	MOFA	MOFA	
FON	AJ-MOFA , NSGA-II	AJ-MOFA NSGA-II	NSGA-II AJ-MOFA	same	NSGA-II CF-MOFA
	MOFA	MOFA	MOFA		AJ-MOFA
	CF-MOFA	CF-MOFA	CF-MOFA		MOFA
POL	AJ-MOFA	AJ-MOFA	NSGA-II	same	NSGA-II
	NSGA-II	NSGA-II	AJ-MOFA		AJ-MOFA
	CF-MOFA	CF-MOFA	CF-MOFA		CF-MOFA
	MOFA	MOFA	MOFA		MOFA
KUR	AJ-MOFA	AJ-MOFA	AJ-MOFA	AJ-MOFA , NSGA-II	NSGA-II
	NSGA-II	NSGA-II	NSGA-II	NSGA-II	AJ-MOFA
	MOFA	MOFA	MOFA	CF-MOFA	CF-MOFA
	CF-MOFA	CF-MOFA	CF-MOFA	MOFA	MOFA
ZDT1	AJ-MOFA	AJ-MOFA	AJ-MOFA	AJ-MOFA , NSGA-II	AJ-MOFA NSGA-II
	NSGA-II	NSGA-II	NSGA-II	NSGA-II	NSGA-II
	CF-MOFA	CF-MOFA	CF-MOFA	MOFA	CF-MOFA
	MOFA	MOFA	MOFA	CF-MOFA	MOFA
ZDT2	AJ-MOFA	AJ-MOFA	AJ-MOFA	AJ-MOFA , NSGA-II	AJ-MOFA NSGA-II
	NSGA-II	NSGA-II	NSGA-II	NSGA-II	NSGA-II
	CF-MOFA	CF-MOFA	CF-MOFA	CF-MOFA	CF-MOFA
	MOFA	MOFA	MOFA	MOFA	MOFA
ZDT3	AJ-MOFA	NSGA-II	NSGA-II	AJ-MOFA , NSGA-II	AJ-MOFA NSGA-II
	NSGA-II	AJ-MOFA	AJ-MOFA	NSGA-II	NSGA-II
	CF-MOFA	CF-MOFA	CF-MOFA	CF-MOFA	CF-MOFA
	MOFA	MOFA	MOFA	MOFA	MOFA
ZDT4	NSGA-II	NSGA-II	NSGA-II	NSGA-II	AJ-MOFA
	AJ-MOFA	AJ-MOFA	AJ-MOFA	AJ-MOFA	NSGA-II
	CF-MOFA	MOFA	CF-MOFA	MOFA	CF-MOFA
	MOFA	CF-MOFA	MOFA	CF-MOFA	MOFA
ZDT6	AJ-MOFA , NSGA-II	NSGA-II AJ-MOFA	AJ-MOFA NSGA-II	NSGA-II AJ-MOFA	NSGA-II AJ-MOFA
	CF-MOFA	CF-MOFA	MOFA	MOFA	CF-MOFA
	MOFA	MOFA	CF-MOFA	CF-MOFA	MOFA

TABLE 11. The simulation parameters that are used for generating the results.

Parameter Symbol	Meaning	Value	Unit
AGG	Aggressive behavior factor	0.8	-
P_n	Probability of normal lane change	0.5	-
P_u	Probability of urgent lane change	0.9	-
$d_{rel,F}$	relative forward distance for normal lane changing	100	Meters
T_R	Transmission Range	300	Meter
$EXP(n)$	Expected number of incoming Vehicles	2	Vehicle
$Exp(T)$	Expected Time Interval	8	Seconds
T	Time Unite	0.1	Seconds

generational distance, which indicates the distance between the found Pareto and the true Pareto. It was determined that AJ-MOFA was also superior to all the benchmarks with respect to GD, except for ZDT4, ZDT6 and ZDT3. The

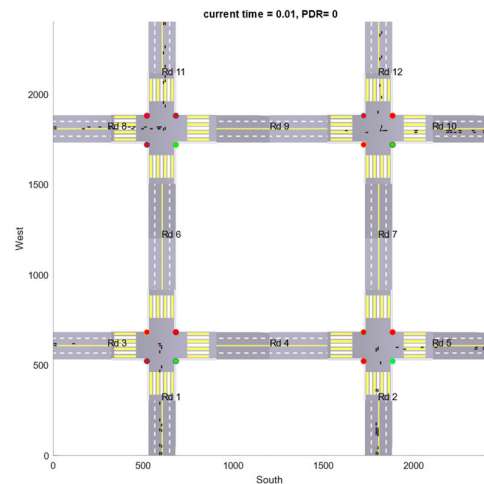


FIGURE 16. Snapshot of the simulation environment used for evaluation.

other metric that indicated the level of performance was the hypervolume, which reflects the flexibility of choices for the

decision-maker when the hypervolume is higher. It was found that NSGA-II and AJ-MOFA were the best in terms of hypervolume. Moreover, AJ-MOFA was better for KUR, ZDT1, ZDT2 and ZDT6. Another aspect of the performance was the number of non-dominated solutions. Like the previous metric, NSGA-II and AJ-MOFA performed best with respect to this metric for the majority of the functions. The last metric generated was the delta metric, which indicates the equal distribution of solutions in the Pareto front and their closeness to the true Pareto. It was found that the delta of AJ-MOFA was better for ZDT2, ZDT3 and ZDT4.

The results were generated based on the parameter depicted in Table 10. We add to this the channel fading model that was used, which is Nakagami. In addition to this, we present in Figure 16 a screenshot from the simulator while running.

As shown, the simulator consists of four junctions. Each one is connected to four roads and at the end of each road is a traffic light. The roads are two-directional and each direction has two lanes.

B. DATA DISSEMINATION

The experiments were executed using a seven-phase model as follows:

- Phase 1: All traffic lights are off
- Phase 2: The traffic lights are in junction-centred mode with the South –North directions open.
- Phase 3: The traffic lights are in junction-centred mode with the East –West directions open.
- Phase 4: The traffic lights are in road-centred mode with the South direction open.
- Phase 5: The traffic lights are in road-centred mode with the West direction open.
- Phase 6: The traffic lights are in road-centred mode with the North direction open.

TABLE 12. The parameters values for the different algorithms.

parameter	AJ-MOFA	MOFA	CMOFA	NSGAI
Number of Solutions	25	25	25	25
Number of Iterations	25	25	25	25
Lower Bound	[2 1 0.5]	[2 1 0.5]	[2 1 0.5]	[2 1 0.5]
Upper Bound	[10 15 1]	[10 15 1]	[10 15 1]	[10 15 1]
Alpha	1	1	1	-
beta0	1	1	1	-
gamma	0.1	0.1	0.1	-
CF	2	-	2	-
jumping CF	2	-	-	-
Mutate Ratio	0.5	-	-	0.5
Mutation fraction	-	-	-	0.1
Crossover Ratio	-	-	-	1.2
Crossover fraction	-	-	-	0.667

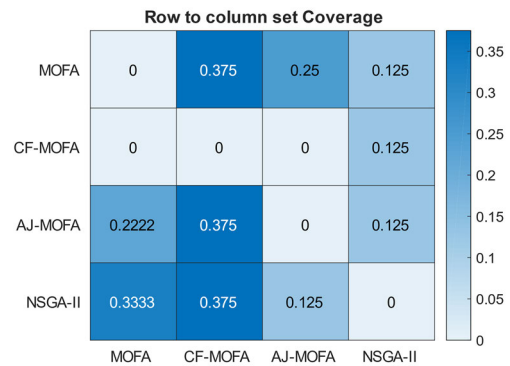


FIGURE 17. Set coverage results for the three variants MOFA, CMOFA and AJ-MOFA.

Phase 7: The traffic lights are in road-centred mode with the East direction open.

The usage of all the phases was important to simulate all the possible traffic states that occur in an urban environment.

The developed AJ-MOFA was used with MOFA, CMOFA and NSGA-II to optimise the data dissemination using probabilistic broadcasting. The parameters of the algorithm operation are given in Table 12. We selected a number of solutions and iterations equal to 25, which was adequate to obtain acceptable results without consuming time in the simulation. We also determined the boundary of the solution space as between [2 1 0.5] and [10 15 1] for the parameters = (MaxCS, Tmax, PH) because MaxCS is an integer number and this range is heuristically acceptable. For the initial probability, it was any number between 0.5 and 1. The other parameters are standard parameters for random searching based on firefly.

The set coverage indicates the domination after running each optimization algorithm, namely MOFA, CF-MOFA, AJ-MOFA and NSGA-II, on the simulation network. The results of the set coverage are shown in Figure 17. The results show that AJ-MOFA dominated MOFA, CF-MOFA and NSGA-II, by percentages of 0.222, 0.375 and 0.125 respectively.

In addition, the poorest performance was observed for CF-MOFA, which was incapable of providing any

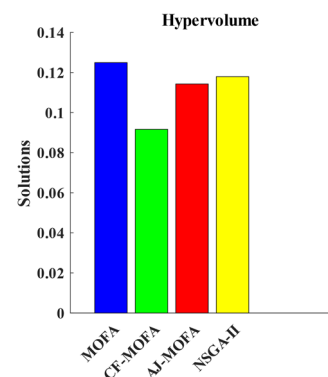


FIGURE 18. Hypervolume of the developed approach AJ-MOFA and its comparison.

domination over the other approaches. To illustrate the performance in terms of the flexibility of the decision making for each algorithm, we present the hypervolume in Figure 18. AJ-MOFA was as competitive as NSGA-II and better than CF-MOFA.

It was found that despite the high domination of the results of our approach, it generated a relatively high hypervolume, which was close to the NSGA-II hypervolume. It should be remembered that the hypervolume is a secondary metric and having less value does not mean inferiority but indicates how much allocated space the dominant solutions have. It was observed that AJ-MOFA reached a hypervolume between 0.1 and 0.12, higher than that of CF-MOFA, which had around 0.09.

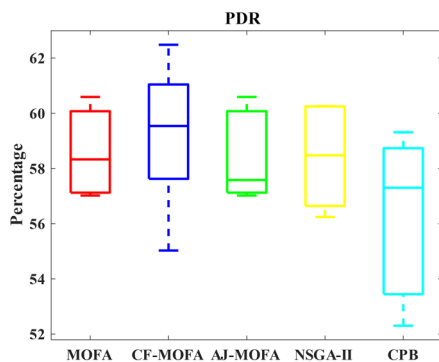


FIGURE 19. PDR for 100 selected solutions from the Pareto front of our developed CF-MOFA and its comparison with the benchmarks.

It was observed that the highest hypervolume was achieved for MOFA. Considering that MOFA was outperformed by the other approaches with respect to domination, the values of this measurement are not an indicator of superiority. A competitive performance was also observed in terms of hypervolume for NSGA-II and AJ-MOFA.

The second part of the evaluation included selecting 100 solutions from the Pareto front from each approach and providing a boxplot comparison of their metrics.

The first metric that was provided was a PDR of 100 randomly selected solutions from the Pareto front, which is shown in Figure 19. It is observed that AJ-MOFA has provided minimum and maximum PDR higher than NSGA-II and CPB. Also, it shows that a significant part of the solutions generated more PDR than the other benchmarks except for CF-MOFA, which provided a small number of solutions with higher PDR. However, this does not indicate the superiority of CF-MOFA as the set coverage showed the domination of AJ-MOFA. AJ-MOFA provided a range of PDR from 57 to higher than 60, which was much higher than CPB. This PDR is associated with less delay, as can be observed in Figure 21.

The second metric that was generated was the doubled received packets times, as depicted in Figure 20. It was observed that AJ-MOFA provided the second-fewest solutions in terms of this metric, which is an indicator of its good performance compared with CF-MOFA. The latter provided

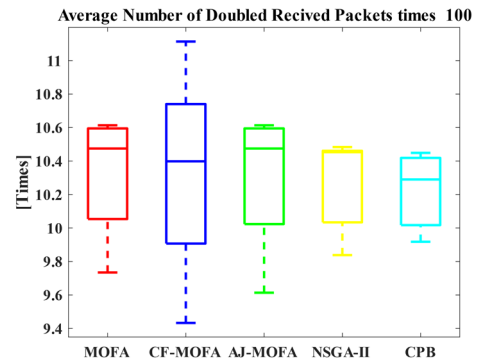


FIGURE 20. Average number of doubled received packets for 100 selected solutions from the Pareto front of our developed CF-MOFA and its comparison with the benchmarks.

less high whisker in terms of this metric as its range of solutions was far higher which indicates to higher diversity in the performance. The PDR percentage for AJ-MOFA reached more than 60% as the highest achieved value and around 57% as the minimum value. These were bigger than CPB, which reached less than 60% as the highest value and around 52% as the minimum value.

The third metric that was generated was the E2E delay, which is depicted in Figure 21. Obviously, the AJ-MOFA approach showed the least E2E delay in terms of median, minimum, maximum and whisker, compared with the benchmarks. This is also an indicator of its superiority and the effectiveness of the various aspects that were used when conducting the optimization, namely objective decomposition and the balance between exploration and exploitation. The number of doubled received packets ranged between 10.6 and 9.6 for AJ-MOFA, compared with the highest ranges of CPB, CF-MOFA, MOFA and NSGA-II. The minimum value for NSGA-II was 9.8.

To elaborate the performance further, we present the time series of one randomly selected solution. The time series, as depicted in Figure 22, shows a convergence performance of the time series in the steady state with some differences between the various approaches. It was observed that the steady state PDR of AJ-MOFA was higher than those of

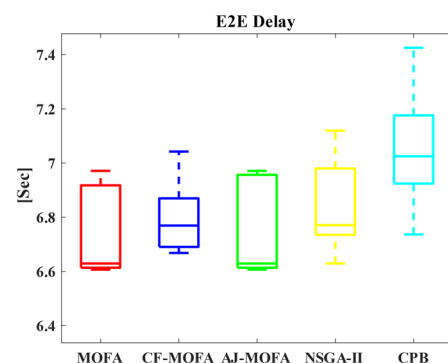


FIGURE 21. Average E2E Delay for 100 selected solutions from the Pareto front of our developed CF-MOFA and its comparison with the benchmarks.

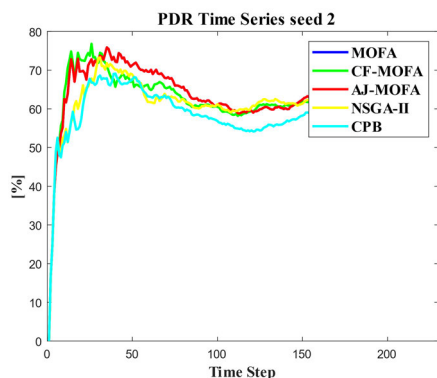


FIGURE 22. Time series of PDR for our developed approach and its comparison with the benchmarks.

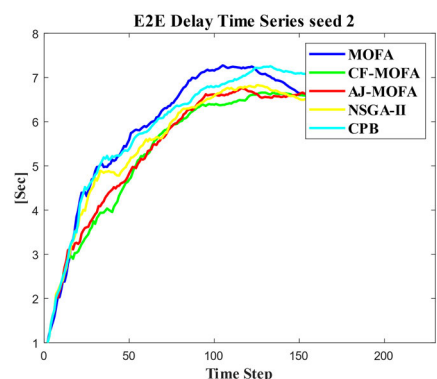


FIGURE 23. Time series of E2E Delay for our developed approach and its comparison with the benchmarks.

the other benchmarks while the steady state of CPB was the lowest. It was observed that AJ-MOFA accomplished the least E2E delay, which reached the value of 6.6 [sec]. This was much lower than the accomplished E2E delay of CPB, which was around 6.7 [sec].

The observed superiority of AJ-MOFA is also consistent with the E2E delay results, which are presented in Figure 23 as a time series. We observe that the highest E2E delay resulted for CPB, while our developed multi-objective optimization-based data dissemination converged to result in a similar performance in terms of E2E delay. The figure shows that PDR reached an accuracy higher than CPB in the steady state.

VI. SUMMARY AND CONCLUSION

This article has addressed the problem of data dissemination from the perspective of multi-objective optimization. Firstly, it proposed a novel framework for data dissemination based on clustering and probabilistic forwarding. Next, it formulated mathematically the problem of data dissemination based on three objective functions, namely PDR, E2E delay and the number of dropped packets. For the solution space design, the optimization uses three variables, namely the maximum allowed cluster size, the maximum threshold and the probability and value for forwarding a packet that is received only once. In addition, it uses a non-linear probabilistic

function to generate the probability of forwarding. In the area of optimization, the article provides three contributions as part of a newly developed multi-objective firefly algorithm, namely objective decomposition, archive management and the exploration and exploitation trade-off based on mutation. In terms of mathematical optimization, the evaluation was conducted using both benchmark mathematical functions and actual VANETs simulations. In terms of data dissemination, the evaluation has demonstrated its superiority, in terms of domination and other MOO metrics, over the benchmarks. In addition, it has shown an accomplished PDR of 60%, an E2E delay of 6.6 seconds and a number of dropped packets that was almost nine for the entire running time of the experiment. This can be compared with the similar or worse performances of the benchmarks for these metrics. Future work should extend the model to handle data dissemination in various urgent scenarios, like car accidents and toll stations. In addition, we add hyper-heuristics for the dynamic selection of the parameters of the meta-heuristic.

REFERENCES

- [1] R. Hussain, F. Hussain, and S. Zeadally, "Integration of VANET and 5G Security: A review of design and implementation issues," *Future Gener. Comput. Syst.*, vol. 101, pp. 843–864, Dec. 2019.
- [2] E. Schoch, F. Kargl, M. Weber, and T. Leinmuller, "Communication patterns in VANETs," *IEEE Commun. Mag.*, vol. 46, no. 11, pp. 119–125, Nov. 2008.
- [3] Y. Yang and R. Bagrodia, "Evaluation of VANET-based advanced intelligent transportation systems," in *Proc. 6th ACM Int. Workshop Veh. Internetworking (VANET)*, 2009, pp. 3–12.
- [4] M. Abdelgadir, R. A. Saeed, and A. Babiker, "Mobility routing model for vehicular ad-hoc networks (VANETs), smart city scenarios," *Veh. Commun.*, vol. 9, pp. 154–161, Jul. 2017.
- [5] A. Kumar, S. Tripathi, and R. S. Raw, "Bringing healthcare to doorstep using VANETs," in *Proc. 3rd Int. Conf. Comput. Sustain. Global Develop. (INDIACom)*, 2016, pp. 2856–2859.
- [6] A. Ullah, S. Yaqoob, M. Imran, and H. Ning, "Emergency message dissemination schemes based on congestion avoidance in VANET and vehicular FoG computing," *IEEE Access*, vol. 7, pp. 1570–1585, 2019.
- [7] L. Liu, C. Chen, T. Qiu, M. Zhang, S. Li, and B. Zhou, "A data dissemination scheme based on clustering and probabilistic broadcasting in VANETs," *Veh. Commun.*, vol. 13, pp. 78–88, Jul. 2018.
- [8] E. A. Feukeu and T. Zuva, "Dynamic broadcast storm mitigation approach for VANETs," *Future Gener. Comput. Syst.*, vol. 107, pp. 1097–1104, Jun. 2020.
- [9] R. Oliveira, C. Montez, A. Boukerche, and M. S. Waghman, "Reliable data dissemination protocol for VANET traffic safety applications," *Ad Hoc Netw.*, vol. 63, pp. 30–44, Aug. 2017.
- [10] Y. Duan, V. C. S. Lee, K.-Y. Lam, W. Nie, and K. Liu, "A cross-layer design for data dissemination in vehicular ad hoc networks," *Neural Comput. Appl.*, vol. 31, no. 7, pp. 2869–2887, Jul. 2019.
- [11] M. Ghribi and A. Meddeb, "Survey and taxonomy of MAC, routing and cross layer protocols using wake-up radio," *J. Netw. Comput. Appl.*, vol. 149, Jan. 2020, Art. no. 102465.
- [12] K. N. Qureshi et al., "Nature-inspired algorithm-based secure data dissemination framework for smart city networks," *Neural Comput. Appl.*, vol. 33, pp. 10637–10656, 2021, doi: 10.1007/s00521-020-04900-z.
- [13] H. Xue, K. Kim, and H. Youn, "Dynamic load balancing of software-defined networking based on genetic-ant colony optimization," *Sensors*, vol. 19, no. 2, p. 311, Jan. 2019.
- [14] M. H. Ali, B. A. D. Al Mohammed, A. Ismail, and M. F. Zolkipli, "A new intrusion detection system based on fast learning network and particle swarm optimization," *IEEE Access*, vol. 6, pp. 20255–20261, 2018.
- [15] S. Vimal, M. Khari, R. G. Crespo, L. Kalaivani, N. Dey, and M. Kaliappan, "Energy enhancement using multiobjective ant colony optimization with double Q learning algorithm for IoT based cognitive radio networks," *Comput. Commun.*, vol. 154, pp. 481–490, Mar. 2020.

- [16] M. Krishnan, S. Yun, and Y. M. Jung, "Improved clustering with firefly-optimization-based mobile data collector for wireless sensor networks," *AEU-Int. J. Electron. Commun.*, vol. 97, pp. 242–251, Dec. 2018.
- [17] M. Agrawal, R. Bansal, A. Choudhary, and A. P. Agrawal, "Heterogeneous computing task scheduling using improved harmony search optimization," in *Proc. Int. Conf. Adv. Comput., Commun. Control Netw. (ICACCCN)*, Oct. 2018, pp. 11–15.
- [18] M. Chiang, "Nonconvex optimization for communication networks," in *Advances in Applied Mathematics and Global Optimization*. Boston, MA, USA: Springer, 2009, pp. 137–196.
- [19] A. F. M. S. Shah, H. Ilhan, and U. Tureli, "CB-MAC: A novel cluster-based MAC protocol for VANETs," *IET Intell. Transp. Syst.*, vol. 13, no. 4, pp. 587–595, Apr. 2019.
- [20] A. F. M. S. Shah, M. A. Karabulut, H. Ilhan, and U. Tureli, "Performance optimization of cluster-based MAC protocol for VANETs," *IEEE Access*, vol. 8, pp. 167731–167738, 2020.
- [21] A. F. M. S. Shah, H. Ilhan, and U. Tureli, "RECV-MAC: A novel reliable and efficient cooperative MAC protocol for VANETs," *IET Commun.*, vol. 13, no. 16, pp. 2541–2549, Oct. 2019.
- [22] A. F. M. S. Shah, H. Ilhan, and U. Tureli, "Performance and complexity analysis of MAC protocol for VANETs," in *Proc. IEEE 10th Annu. Inf. Technol., Electron. Mobile Commun. Conf. (IEMCON)*, Oct. 2019, pp. 1081–1086.
- [23] M. A. Karabulut, A. F. M. S. Shah, and H. Ilhan, "OEC-MAC: A novel OFDMA based efficient cooperative MAC protocol for VANETs," *IEEE Access*, vol. 8, pp. 94665–94677, 2020.
- [24] C. Wu, Z. Liu, F. Liu, T. Yoshinaga, Y. Ji, and J. Li, "Collaborative learning of communication routes in edge-enabled multi-access vehicular environment," *IEEE Trans. Cognit. Commun. Netw.*, vol. 6, no. 4, pp. 1155–1165, Dec. 2020.
- [25] C. Wu, Z. Liu, T. Yoshinaga, Y. Ji, and D. Zhang, "Spatial intelligence toward trustworthy vehicular IoT," *IEEE Commun. Mag.*, vol. 56, no. 10, pp. 22–27, Oct. 2018.
- [26] R. Hajlaoui, H. Guyennet, and T. Moulahi, "A survey on heuristic-based routing methods in vehicular ad-hoc network: Technical challenges and future trends," *IEEE Sensors J.*, vol. 16, no. 17, pp. 6782–6792, Sep. 2016.
- [27] N. Harrag, A. Refoufi, and A. Harrag, "New NSGA-II-based OLSR self-organized routing protocol for mobile ad hoc networks," *J. Ambient Intell. Humanized Comput.*, vol. 10, no. 4, pp. 1339–1359, Apr. 2019.
- [28] J. Toutouh, J. García-Nieto, and E. Alba, "Intelligent OLSR routing protocol optimization for VANETs," *IEEE Trans. Veh. Technol.*, vol. 61, no. 4, pp. 1884–1894, May 2012.
- [29] M. Elhoseny, "Intelligent firefly-based algorithm with levy distribution (FF-L) for multicast routing in vehicular communications," *Expert Syst. Appl.*, vol. 140, Feb. 2020, Art. no. 112889.
- [30] M. Chahal and S. Harit, "Optimal path for data dissemination in vehicular ad hoc networks using meta-heuristic," *Comput. Electr. Eng.*, vol. 76, pp. 40–55, Jun. 2019.
- [31] A. Hidalgo-Paniagua, M. A. Vega-Rodríguez, J. Ferruz, and N. Pavón, "Solving the multi-objective path planning problem in mobile robotics with a firefly-based approach," *Soft Comput.*, vol. 21, no. 4, pp. 949–964, Feb. 2017.
- [32] C. J. Joshua and V. Varadarajan, "An optimization framework for routing protocols in VANETs: A multi-objective firefly algorithm approach," *Wireless Netw.*, vol. 27, pp. 1–10, Jul. 2019.
- [33] N. Bahadormanesh, S. Rahat, and M. Yarali, "Constrained multi-objective optimization of radial expanders in organic rankine cycles by firefly algorithm," *Energy Convers. Manage.*, vol. 148, pp. 1179–1193, Sep. 2017.
- [34] S. Karthikeyan, P. Asokan, S. Nickolas, and T. Page, "A hybrid discrete firefly algorithm for solving multi-objective flexible job shop scheduling problems," *Int. J. Bio-Inspired Comput.*, vol. 7, no. 6, pp. 386–401, 2015.
- [35] G. Chen, X. Yi, Z. Zhang, and H. Wang, "Applications of multi-objective dimension-based firefly algorithm to optimize the power losses, emission, and cost in power systems," *Appl. Soft Comput.*, vol. 68, pp. 322–342, Jul. 2018.
- [36] L. Lv, J. Zhao, J. Wang, and T. Fan, "Multi-objective firefly algorithm based on compensation factor and elite learning," *Future Gener. Comput. Syst.*, vol. 91, pp. 37–47, Feb. 2019.
- [37] T. Fan, J. Wang, M. Feng, X. Zhang, J. Wang, and R. Wu, "Application of multi-objective firefly algorithm based on archive learning in robot path planning," *Int. J. Intell. Inf. Database Syst.*, vol. 12, no. 3, pp. 199–211, 2019.
- [38] S. Neelima, N. Satyanarayana, and P. K. Murthy, "A novel multi-objective firefly algorithm for optimization of association rules," in *Proc. Int. Conf. Big Data Anal. Comput. Intell. (ICBDAC)*, Mar. 2017, pp. 428–431.
- [39] X.-S. Yang, "Multiobjective firefly algorithm for continuous optimization," *Eng. Comput.*, vol. 29, no. 2, pp. 175–184, Apr. 2013.
- [40] H. Wang, W. Wang, L. Cui, H. Sun, J. Zhao, Y. Wang, and Y. Xue, "A hybrid multi-objective firefly algorithm for big data optimization," *Appl. Soft Comput.*, vol. 69, pp. 806–815, Aug. 2018.
- [41] C. X. Zhao, C. Wu, J. Chai, X. Wang, X. Yang, J.-M. Lee, and M. J. Kim, "Decomposition-based multi-objective firefly algorithm for RFID network planning with uncertainty," *Appl. Soft Comput.*, vol. 55, pp. 549–564, Jun. 2017.
- [42] T. Shu, X. Gao, S. Chen, S. Wang, K. Lai, and L. Gan, "Weighing efficiency-robustness in supply chain disruption by multi-objective firefly algorithm," *Sustainability*, vol. 8, no. 3, p. 250, Mar. 2016.
- [43] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan, "A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II," in *Proc. Parallel Problem Solving From Nature VI Conf.*, 2000, pp. 849–858.
- [44] A. Metiaf, Q. Wu, and Y. Aljeroudi, "Searching with direction awareness: Multi-objective genetic algorithm based on angle quantization and crowding distance MOGA-AQCD," *IEEE Access*, vol. 7, pp. 10196–10207, 2019.
- [45] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*, vol. 16. Hoboken, NJ, USA: Wiley, 2001.



MUSTAFA MAAD HAMDI was born in Al-Anbar, Iraq. He received the B.Eng. degree in computer engineering technology from the Al-Maarif University College, Iraq, and the M.Sc. degree in communication and computer engineering from the University Kebangsaan Malaysia (UKM), Malaysia. He is currently pursuing the Ph.D. degree with the Department of Communication Engineering, University Tun Hussein Onn Malaysia (UTHM), Malaysia. His research interests include wireless and mobile communications, VANET, MANET, and satellite communication, and cryptographic.



LUKMAN AUDAH was born in Kuala Lumpur, Malaysia. He received the B.Eng. degree in telecommunications from the Universiti Teknologi Malaysia, in 2005, and the M.Sc. degree in communication networks and software and the Ph.D. degree in electronic engineering from the University of Surrey, U.K. He is currently a Lecturer with the Communication Engineering Department, University Tun Hussein Onn Malaysia. His research interests include wireless and mobile communications, internet traffic engineering, network system management, data security, and satellite communication.



SAMI ABDULJABBAR RASHID was born in Al-Anbar, Iraq. He received the B.Eng. degree in computer engineering technology from the Al-Maarif University College, Iraq, and the M.Sc. degree in communication and computer engineering from the University Kebangsaan Malaysia (UKM), Malaysia. He is currently pursuing the Ph.D. degree with the Department of Communication Engineering, University Tun Hussein Onn Malaysia (UTHM), Malaysia. His research interests include wireless and mobile communications and VANET.

...