IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 30 May 2022, accepted 16 June 2022, date of publication 1 July 2022, date of current version 26 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3187698

Cauchy Density-Based Algorithm for VANETs Clustering in 3D Road Environments

AHMED SALIH AL-OBAIDI[1], (Member, IEEE), MOHAMMED AHMED JUBAIR[2], IZZATDIN ABDUL AZIZ[3], MOHD RIDUAN AHMAD[1], (Member, IEEE), SALAMA A. MOSTAFA[4], (Member, IEEE), HAIRULNIZAM MAHDIN[4], (Member, IEEE), ABDULLAH TALAAT AL-TICKRITI[1], AND MUSTAFA HAMID HASSAN[2]

[1]Centre for Telecommunication and Research Innovation (CeTRI), Broadband and Networking (BBNET) Research Group, Fakulti Kejuruteraan Elektronik dan Kejuruteraan Komputer (FKEKK), Universiti Teknikal Malaysia Melaka (UTeM), Durian Tunggal, Malacca 76100, Malaysia
[2]Department of Computer Technical Engineering, College of Information Technology, Imam Ja'afar Al-Sadiq University, Al-Muthanna 66002, Iraq
[3]Center for Research in Data Science (CeRDaS), Computer and Information Science Department (CISD), Universiti Teknologi PETRONAS (UTP), Seri Iskandar, Perak Darul Ridzuan 32610, Malaysia
[4]Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Parit Raja, Johor 86400, Malaysia

Corresponding authors: Mohammed Ahmed Jubair (mohammed.a@sadiq.edu.iq), Mohd Riduan Ahmad (riduan@utem.edu.my), and Izzatdin Abdul Aziz (izzatdin@utp.edu.my)

This work was supported by the Yayasan Universiti Teknologi PETRONAS (UTP) Grant with the title Predicting Missing Values in Big Upstream Oil and Gas Industrial Dataset using Enhanced Evolved Bat Algorithm and Support Vector Regression, through the Center for Research in Data Science (CerDaS), UTP, Malaysia, under Grant 015LC0-353.

ABSTRACT Vehicular ad hoc networks (VANETs) are emerging to serve various types of applications for serving smart cities and intelligent transportation systems. There are several challenging factors for ensuring reliable and stable VANETs communications. VANETs clustering is essential functionality to serve routing protocols and enable reliable VANETs. Clustering algorithms for VANETs operate in a decentralized mode, which requires incorporating additional stages before deciding the clustering decisions and might create sub-optimality due to the local nature of the decentralized approach. In addition, the challenging architecture of the road environment can cause confusing clustering decisions. This problem becomes more challenging due to the evolving nature of clusters in VANETs in general and in 3D VANETs in particular. This paper attempts to solve the problem of VANETs in 3D road environments using a centralized clustering technique to develop a Cauchy density model. The model has been simulated by considering several simulation parameters including traffic, mobility, driving behavior, and road curvature. The simulator also includes an adjacency list that defines the road's points and straight-line segments. The clustering technique of the Cauchy density model determines the mobility vector to enable adding vehicles to their respective clusters. The simulator has been implemented in MATLAB to perform complex scenarios in three locations of 3D road environments. A comparison with selected benchmarks shows the superiority of our model over the benchmarks models in which our model achieves an improvement percentage of 1%, 10%, and 3% for average cluster head duration, average cluster member duration, and clustering efficiency, respectively.

INDEX TERMS VANETs, 3D road environment, clustering, Cauchy density.

I. INTRODUCTION

The outstanding progress in the various technology has created systems capable of solving numerous technical problems before enabling an intelligent transportation system. One of these systems is the intelligent transportation system that enables intelligent services in the roadway environment. Most of the services are dependent on stable communication systems between the vehicles on the road and the infrastructure. Using vehicle to vehicle (V2V) and vehicle to infrastructure (V2I), communication is established within the vehicular ad hoc network (VANETs) [1].

VANETs clustering decomposes the networks into groups (clusters) based on maximum intra-connectivity and

The associate editor coordinating the review of this manuscript and approving it for publication was Mohamed M. A. Moustafa.

minimum inter-connectivity. It is used for enabling basic infrastructure for routing and other purposes and is useful in wireless networks such as mobile ad hoc networks (MANETs) [2] and VANETs. Clustering is regarded as a useful approach for guaranteeing stable communications of VANETs under various challenges. Some challenges are the high-speed moving vehicles and the time-constrained communication or requirements [3]. Because of the fast speed of vehicles,

Topological changes often occur, causing network instability and impediments to reliable information sharing. Hence, a cluster-based structure is regarded as an optimum solution to provide better network stability in VANET because it collects the data from its neighbor and sends them to the base station or infrastructure.

Most existing proposals have considered modeling VANET in a 2D scenario, although some have considered modeling it in 3D road environment scenarios. In a recent review of [4], all the reviewed protocols considered 2D modeling in the road environment, which is generally acceptable in traditional cities. However, in recently developed cities, the 3D infrastructure is more widely used [5]. Furthermore, vehicles are expected to travel within tunnels and multi-levels of bridges and interchanges, which causes a degradation in the communication performance when 3D modeling is ignored.

The nature of VANETs clustering is similar to stream data clustering to a certain degree. First, the points are tuples of combined multi-dimensional data in data clustering and in VANETs clustering. A tuple of multi-dimensional features also represents the vehicles. Second, the role of stream data clustering is to group similar points into groups, and the role of VANETs clustering is to group vehicles into their clusters. Third, there is an evolving nature of the clusters in both stream clustering and VANETs clustering. Fourth, density is an essential criterion in deciding instances (points or vehicles) assigned to their respective clusters. One of the recently developed algorithms for stream data clustering that takes into consideration points clustering and is applicable for evolving clusters is stream Cauchy clustering [6].

This article aims to fill this gap by developing a 3D clustering Cauchy density model and a 3D traffic simulator. The article presents the following contributions:

- Proposing clustering Cauchy density model for a 3D V2I communication of VANET.
- Modeling and development of a 3D traffic simulator to simulate real road topology with a multi-level stack interchange. The simulator converts the environment map into a set of segments represented by an adjacency list that defines the road points based on straight-line segments. The adjacency list represents a graph that describes the road environment. The simulator considers deploying different parameters, including the number of vehicles, vehicle movement, and driver behavior.
- Evaluating the performance of the developed model based on standard evaluation metrics using various 3D scenarios. Then comparing the performance of the model with the state-of-the-art models.

The remainder of the article is organized as follows: the literature review is presented in Section II. It includes a background of the study and the related work. The methodology is presented in Section III. While the experimental results and analysis are presented in Section IV. Finally, the conclusion and future works are presented in Section V.

II. LITERATURE REVIEW

A. BACKGROUND

One of the recently developed algorithms for stream clustering is the work of [6]. This algorithm considers the model of density given in Equation (1).

$$\gamma_k^j = \frac{1}{1 + \frac{1}{\sigma_l^2} \frac{\sum_{i=1}^{M^j} d_{ki}^j}{M^j}} j = 1, \ldots, m \tag{1}$$

where

γ_k^j denotes the density of adding the point j to cluster k

σ_l denotes a normalization constant that spreads or narrows the membership functions

d_{ki}^j denotes the square of the Euclidean distance between the current data sample k and the i sample from cluster j

M^j denotes the number of clusters j

It adds the point to the cluster, generating a higher density for the sample. However, the algorithm requires that the density is higher than the predefined threshold. Otherwise, it uses the point for creating a new cluster. The algorithm in [6] has derived a recursive model for updating the cluster when a point is added to it.

$$j = \arg \max_{j \epsilon [1,c]} \gamma_k^j \tag{2}$$

$$M^j \leftarrow M^j + 1$$

$$e^j(k) \leftarrow z(k) - \mu^j \tag{3}$$

$$\mu^j \leftarrow \mu^j + \frac{1}{M^j} e^j(k) \tag{4}$$

$$S^j \leftarrow S^j + e^j(k) \left(z(k) - \mu^j\right)^T$$

$$\Sigma^j \leftarrow \frac{1}{M^j} S^j \tag{5}$$

B. LITERATURE SURVEY

The literature contains numerous approaches, protocols, and algorithms for VANETs clustering. In the work of [7], an enhanced weight-based clustering algorithm (EWCA) was developed to address smart cities and intelligent transportation systems challenges. This approach has considered any vehicle moving on the same road segment with the same road ID and within the transmission range of its neighbor to be suitable for the cluster formation process. This issue was attributed to the fact that all safety messages are expected to be shared among adjacent vehicles (irrespective of their relative speed) to avoid any hazardous situation. It identified some metrics based on vehicle mobility information to elect a CH. Each vehicle was associated with a predefined weight value based on its relevance. A vehicle with the highest weight value was elected as the primary cluster head (PCH). It also introduced a secondary cluster head (SeCH) as a backup to the PCH to improve the cluster stability. SeCH took over the leadership whenever the PCH was not suitable for continuing with the leadership.

Some researchers have focused on enhancing clustering using optimization approaches. In the work of [8], a grasshoppers' optimization-based clustering based on various parameters such as the number of clusters, network area, node density, and transmission range was conducted. Other researchers have tackled clustering using game theory-based optimization.

In the work of [9], a clustering method based on coalitional game theory with the purpose to improve the average V2V signal-to-noise ratio (SNR) and channel capacity while maintaining the stability of the cluster. Each vehicle attempts to form a cluster with other vehicles according to coalition value in their work. The value of coalition is formulated based on the V2V SNR, connection lifetime, and speed difference between vehicles to attain the required clustering. A higher average of SNR can be achieved in a fast-changing network topology, but the cluster's stability becomes hard to maintain.

Other approaches have focused on extending existing protocols in the work of [10] cluster-based routing protocol for Urban Vehicular Ad Hoc Networks (VANETs) using Optimized Link State Routing (OLSR). In OLSR, MultiPoint Relays (MPRs) used for routing are selected at each node using neighbors' reachability resulting in a high percentage of MPRs.

Quality-of-Service (QoS) was introduced to improve MPRs' quality in VANETs, while clustering was introduced to reduce MPRs' percentage in dense areas. Relay selection in urban VANETs routing protocols incorporates vehicle mobility metrics, velocity, and position, whose significance is lowered by abrupt topology change. Other researchers have included the driver's social aspect in the clustering. In the work of [11], clustering for vehicle networks was proposed based on drivers' social relationships combined with the instantaneous position and speed of the vehicle node.

From the previous literature, we observe that clustering in VANET has focused on different aspects related to density, driver social aspect, optimization of protocols, or extending existing routing protocols to support clustering. However, the research on VANET clustering is poor in the 3D environment. According to the survey of [4]. It is stated that the clustering in 3D structure is a research gap for VANET clustering, and it has not yet received adequate attention from researchers. The impact of 3D structure interprets this: the communication characteristics change from intra-level to enter-level communication. For example, the transmission range of the intra-level is higher than what it is in the inter-level [12]. This implies that the cluster in the 3D structure does not have spherical nature.

In an experiment conducted by [12], using both intra-level and enter-level communication between sender and receiver, the results have shown a decrease in both the ranges where the packets can be received successfully and the range where the node can be connected without normal communication. The decrease happens from the 2D to 3D with a distance of 6m, a percentage of $\frac{89.75}{144}$ under frequency band 2.4 G and $\frac{51}{139.5}$ under frequency band 5.9 G. These ratios are changed to $\frac{16.25}{89.75}$ and $\frac{5}{51.5}$ in the 3D representation with a distance of 10 m. Furthermore, generalizing various VANETs models from the 2D environment to 3D is still challenging.

In the work of [13], it is stated that Many geographical routing protocols based on greedy and face routing approaches have been designed for 2D networks, but these protocols may not be suitable in a 3D environment like hill areas, airborne networks, underground networks, underwater networks and so forth.

In the work of [14], a cluster-based dynamic V2V channel model is proposed for obstructed line-of-sight (OLOS) scenarios. The influences of vehicle obstructions on path loss, delay, and angle dispersion are embodied as changes in the statistical distribution of MPCs clusters.

III. METHODOLOGY

The developed methodology is presented in this section. It consists of a traffic generator that is given in sub-section A. Next, the mobility model is presented in sub-section A.1. Afterward, the model of road curvature is presented in sub-section A.2. Next, the clustering algorithm is provided in sub-section A.3. Lastly, the evaluation metrics are presented in subsection E.

A. SIMULATOR DESIGN

This section presents the developed 3D simulator. It consists of two sub-sections, namely, the traffic generator, which is presented in sub-section A. Next, is the mobility model presented in sub-section B. Lastly, the model of road curvature is given in sub-section C.

1) TRAFFIC GENERATOR

The vehicles are generated in the environment based on two probability density functions: a normal distribution for generating the vehicles on the highway as batches and an exponential distribution for generating the time interval between each batch.

$$pdf(N) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(N-\mu)^2}{2\sigma^2}} \tag{6}$$

where

μ denotes the expected size of one batch

σ denotes the standard deviation of the batch size

$pdf(N)$ denotes the probability of generating a batch with a certain size N

2) MOBILITY MODEL

In this model, the vehicles' motilities are based on the generated accelerations and their integrations. Each vehicle is assigned an acceleration that is generated using

Equation (9-11) and then integrated using Equation (7) to obtain the velocity, which is integrated using Equation (8) to obtain the distance.

The approach to generating the acceleration is based on two random variables, U_1 and U_2. The goal of the random variable U_2 is to give the vehicle a random value of the acceleration within $[0, A_{max}]$ or the deceleration $[-D_{max}, 0]$, while the goal of the random variable U_1 is to give the vehicle one of three decision (acceleration, deceleration, or neither of them). The acceleration decision is controlled by acc_i and p_r and the deceleration decision is controlled by $dacc_i$ and p_r, and both of them are controlled by AGG. AGG aims to give the model a higher probability of acceleration than deceleration with a percentage of 70%. This is based on statistical studies of driver behavior on highways. Equations (9–11) give the model of generating the acceleration. The vehicle velocity after integration in Equation (7) has to be clipped according to the maximum and minimum velocities V_{max} and V_{min}, as shown in Equations (12) and (13).

$$v_i\left(t+\Delta t\right)=v_i\left(t\right)+a_i\left(t\right).\Delta t \tag{7}$$

$$x_i\left(t+\Delta t\right)=x_i\left(t\right)+v_i(t)\Delta t \tag{8}$$

$$a_i\left(t\right)=\begin{cases} U_2.A_{max}, & if\ U_1<acc_i+p_r \\ U_2\left(-1\right).D_{max}, & \\ \quad if\ acc_i+pr\leq U_1<acc_i+dacc+2p_r & \\ 0, & otherwise \end{cases} \tag{9}$$

$$acc_i=\begin{cases} U_4\left(1-2p_r\right), & if\ U_3<\frac{3AGG}{4} \\ 0, & otherwise \end{cases} \tag{10}$$

$$dacc_i=\begin{cases} U_4\left(1-2p_r\right), & if\ \frac{3AGG}{4}<U_3<AGG \\ 0, & otherwise \end{cases} \tag{11}$$

$$if\ v_i\left(t\right)>V_{max},\ then\ v_i\left(t\right)=V_{max} \tag{12}$$

$$if\ v_i\left(t\right)>V_{min},\ then\ v_i\left(t\right)=V_{min} \tag{13}$$

3) MODEL OF ROAD CURVATURE

The road curvature will be generated based on an adjacency list that defines the points of the road that defines a straight-line segment. The adjacency list represents a graph describing the tested road environment. An example of the road structure and its equivalent graph representation is shown in Figure 1. The adjacency list is given in Table 1. As it is shown in the graph, the road structure is converted to straight line-based graphs using added points that allow the separation of any curved road into small straight-line segments that have their endpoints defined in 3D coordinates. The entire model of road curvature is described by a set of segments with their detailed 3D coordinates assuming that the road is translated into segments $\{s_1, s_2, \ldots s_n\}$, where:

$$\begin{aligned} s_i &= \left[start\left(s_i\right) end\left(s_i\right)\right]; \\ start\left(s_i\right) &= \left[x_{start}\left(s_i\right) y_{start}\left(s_i\right) z_{start}\left(s_i\right)\right]; \\ end\left(s_i\right) &= \left[x_{end}\left(s_i\right) y_{end}\left(s_i\right) z_{end}\left(s_i\right)\right]; \end{aligned}$$

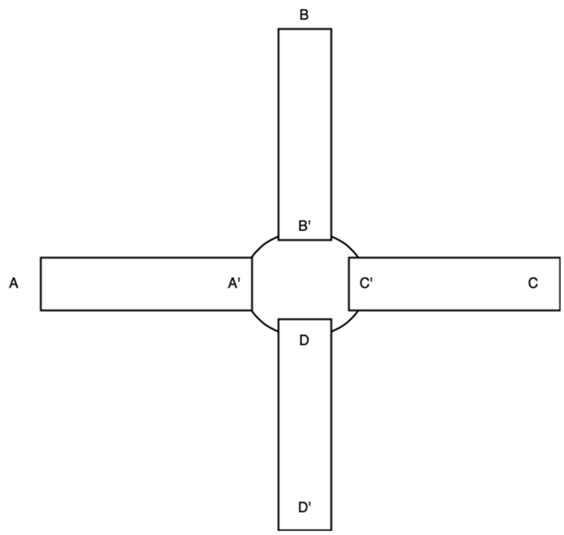


FIGURE 1. Roadway-type environment combined of the roundabout and four connected roads.

TABLE 1. The adjacency list that represents the roundabout and its four connected roads.

Points	Links	Number of lanes	Direction
A	AA’	2	1,0
A’	A’B’	1	1
B	A’D’	1	1
B’	B’C’	1	1
C	B’B	2	1,0
C’	D’C’	1	1
D	D’D	2	1,0
D’	C’C	2	1,0

Figure 1 shows a roundabout with four connected roads. Each road contains two directions, one toward the roundabout and the other in the opposite direction. While the roundabout has only one direction. The number of links and sampled points to represent this road infrastructure is 8, respectively.

B. CLUSTERING

Our model of VANET clustering is presented in Table 2. As observed in the pseudocode, the method accepts the cluster set C as an input, including 1) the set of IDs of vehicles as *ID*, 2) the set of IDs of roads as *RoadID*, 3) the direction of the vehicle in the road *direction*, 4) the cartesian coordinate of the vehicle *position*, which is represented as (x, y, z) coordinates, 5) the magnitude of the velocity vector *velocity*, 6) the magnitude of the acceleration , 7) *acceleration*, and 8) the transmission range of the vehicle. The output of the clustering vehicle is C, which represents the updated cluster. The algorithm starts by going through the vehicles one by one and updating the mobility vector, which is a combination of position, velocity, and acceleration. In addition, an update of the meta-data, which contains *RoadID* and *the direction* of each vehicle, is done with respect to the road.

Subsequently, the algorithm distinguishes between two cases: the first one is the cold start case where the input cluster set is empty, so the code creates a new cluster with the corresponding vehicle. The second case is where there are already established clusters in the input cluster set, so the algorithm calculates the density of the vehicle for the existing clusters using the method *CalculateDensity*(). It selects the cluster where the corresponding vehicle provides the maximum Gamma value as a candidate cluster for adding the vehicle. The vehicle is added to the candidate cluster only when three conditions are met. The first one is that the value of Gamma is higher than the current cluster Gamma. The second one is when the vehicle is on the same road. The third one is when the vehicle has the same direction as the current cluster. Otherwise, the vehicle is used to create a new cluster.

TABLE 2. Clustering algorithm.

```
Input:
(1) rate: the working rate of the algorithm.
(2) ID: the id of the vehicle on the network.
(3) RoadID: the ID of the road the vehicle is currently on.
(4) Direction: the direction of the vehicle on the road [0: from the start
point
to the endpoint, 1 the opposite direction ].
(5) Position: the vehicle's Cartesian coordinates [x, y, ] and z.
(6) Velocity: the magnitude of the velocity vector.
(7) Acceleration: the magnitude of the acceleration vector.
(8) transmission Range: the transmission range of the vehicle.
Output:
C: the updated Clusters set.
Start algorithm
C ⟵ emptyClustersSet()
streamCounter ⟵ 1
while receiving vehicles data do
  for each vehicle i do
    vehicleMobility ←
[Position(i); Position(i); V elocity(i) ; Acceleration(i)]
    metaData ← [RoadID(i); Direction(i)]
    if numel(C) == 0 then
      AddNewCluster(C, vehicleMobility, metaData, ID(i))
    else
      C ⟵ CalculateDensity(C; vehicleMobility)
      bestCluster ⟵ MaxGamma(C; vehicleMobility)
      [cRoadID; Center; cDir]  bestCluster:metaData
      if ((bestCluster.Gama ≪ C.Gama) ∨ (cRoadID ≠
pRoadID) ∨ (pDir ≠ cDir)) then
        AddNewCluster(C, vehicleMobility, metaData, ID(i))
      else
        C = UpdateCluster(C, vehicleMobility)
      end if
    end if
  end for
  deletEmptClusters(C)
  if rem(streamCounter, rate) == 0 then
  rest(C)
  end if
  streamCounter ++
end while
End algorithm
```

When a vehicle is added to a new cluster, two operations happen simultaneously: the cluster counter is incremented by one while the vehicle's information is stored in its corresponding input point in the cluster, namely, the mobility vector, the meta-data, and the ID. The pseudocode is presented in Table 3.

TABLE 3. An algorithm for AddNewCluster.

```
Input:
(1) C: the Clusters set.
(2) vehicleMobility: the mobility data of the vehicle.
(3) metaData: [roadID, dir].
(4) ID: the vehicle ID.
Output:
C: the updated Clusters set.
start algorithm
C.c ⟵ C.c + 1
C.clusters(C.c): Mj ← 1
C.clusters(C.c).miuj ← vehicleMobility
C.clusters(C.c).CHID ← ID
C.clusters(C.c).metaData ← metaData
End algorithm
```

The third algorithm is presented in Table 4. The algorithm's input is cluster set C before the update, the vehicle mobility vector named *vehicleMobility*. The output is cluster C after updating the values of Gamma when the node is added.

TABLE 4. An algorithm for CalculateDensity.

```
Inputs: C: clusters set vehicleMobility: the vehicle mobility data
Output: C: updated Clusters set
for each cluster j in C; Euclidian distance^2; End
Input:
(1) C: Clusters set.
(2) vehicleMobility: the mobility data of the vehicle.
Output:
C: the updated Clusters set.
Start algorithm
for each cluster j in C do
  d ← (vehicleMobility − C.clusters(j).miuj)^2
  C.clusters(j).gammaj ← 1/d
end for
End algorithm
```

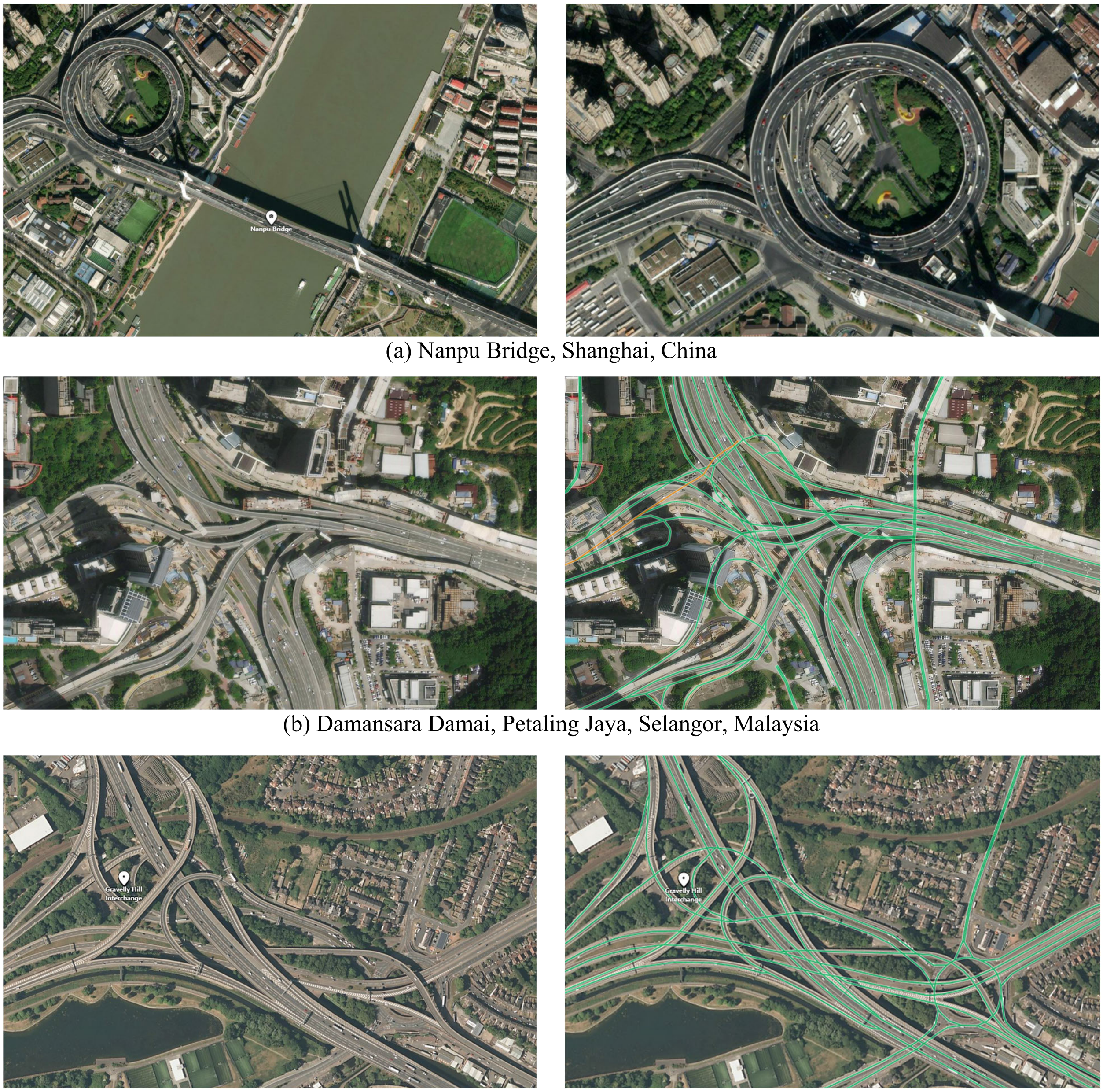
C. EVALUATION METRICS

The evaluation of clustering approaches in the VANETs concentrates on the stability of the generated clusters. More specifically, the longer state of the vehicle in terms of its role as cluster head or cluster member can be used as the clustering performance metrics, which have been called the cluster head duration and cluster member duration, respectively.

Another aspect of the performance is the clustering efficiency, which indicates the percentage of vehicles participating in the clusters. This metric shows a view of the effectiveness of the clustering approach. Linking this measure to the other aspects of the performance, such as the number of clusters that need to be minimized, can provide more stability and effectiveness.

1) CLUSTERING EFFICIENCY

The clustering efficiency is defined as the percentage of vehicles participating in a clustering procedure during the simulation. It is calculated by dividing the number of vehicles that were part of the clusters (cluster member or cluster head) over the total number of vehicles.



(a) Nanpu Bridge, Shanghai, China

(b) Damansara Damai, Petaling Jaya, Selangor, Malaysia

(c) Spaghetti Junction, Gravelly Hill Interchange, Birmingham, UK

FIGURE 2. The three locations that are used for the evaluation scenario [18].

2) AVERAGE CLUSTER HEAD DURATION

The average of the cluster head duration indicates the average time of being in the state of the cluster head before changing the cluster head. As it has been mentioned earlier, the longer CH duration is an indicator of more stability of the clustering approach. The approach calculating the average cluster head duration is by dividing the total cluster head period over the number of changes to cluster head from any other state.

3) AVERAGE CLUSTER MEMBER DURATION

This measure is an indicator of the stability of the clustering approach. It refers to the average period of being in the state of a cluster member. Thus, for each conversion to a cluster member, we calculate the time and divide it by the total number of changes to the state of the cluster member. Our goal is to maximize the cluster member duration.

4) NUMBER OF CLUSTERS

Our goal is to minimize the number of clusters. The number of clusters defines how many clusters result from the clustering algorithm during the whole lifetime of the network.

IV. EXPERIMENTAL RESULTS

This section presents an analysis of the generated evaluation's measures for the VANETs clustering using various simulation scenarios and comparing each with three benchmarks. Figure 2 The three locations that are used for the simulation

TABLE 5. Parameters setting our developed approach and the benchmarks.

Parameters	Our	CBSC	K-Means	ECE-GP
Transmission range	300 m	300 m	300 m	300 m
gammaC : cluster density initiating threshold	1/300^2	-	-	-
Blob Size	-	3000 pixel	-	-
Structuring element	-	Disk with R=61	-	-
Alpha: position coefficient for CH selection	-	1/3	1/3	-
Beta: velocity Coefficient for CH selection	-	1/3	1/3	-
Gamma: Acceleration coefficient for CH selection	-	1/3	1/3	-
Max intensity of the clusters	-	-	5	-
CMC radius	-	-	-	0.1
Grid Granularity	-	-	-	0.25
CMC Min Threshold	-	-	-	1
CMC Decay Ratio	-	-	-	2
Decay	-	-	-	1
Merging Time Interval	-	-	-	5
Minimum Number Of Merging Links	-	-	-	1

scenarios. The benchmarks of this study are an evolving data clustering algorithm (EDCA) [14], a center-based stable clustering [15], a center-based stable evolving clustering with grid partitioning (CEC-GP) [16], and a mutated k-means algorithm [17].

A. EXPERIMENTAL DESIGN

This study found that EDCA is originally used for data clustering. However, it considers the data as a stream that makes it applicable to the VANETs clustering if it has been considered a data point representing the feature associated with the vehicle. Table 5 shows the settings of the main simulation parameters.

B. EVALUATION RESULTS

This section provides the experimental results conducted in this article. The experimental results show that we have conducted scenarios in three locations with 3D architecture. Also, as presented in the evaluation metrics, we are interested in cluster head duration, cluster member duration, and clustering efficiency. In addition, we find the clustering change rate as an overall evaluation metric.

The time series of the cluster head duration of our developed approach and its comparison with the benchmark is presented in Figure 3. They are visualized for the batch number, as shown in the figure. The time series shows that our approach has provided the highest values of clustering duration for the cluster membership and head duration, which is indicative of the algorithm's stability.

The second algorithm in terms of the cluster head duration is center-based stable clustering (CBSC). However, it suffers from some sudden drop in the duration in some time intervals. This issue results from the slow response of the algorithm to the dynamics of cluster change. It has been improved in our approach. In addition, we observe similar behavior of our algorithm in terms of cluster member duration for having a higher value and a lower drop in the time series compared with CBSC, which has generated the second superior performance. On the other side, we find that ECE-GP has generated

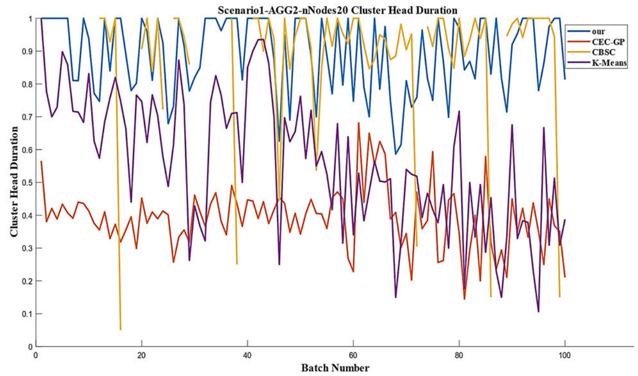


FIGURE 3. Time series of the cluster head duration for the batch number.

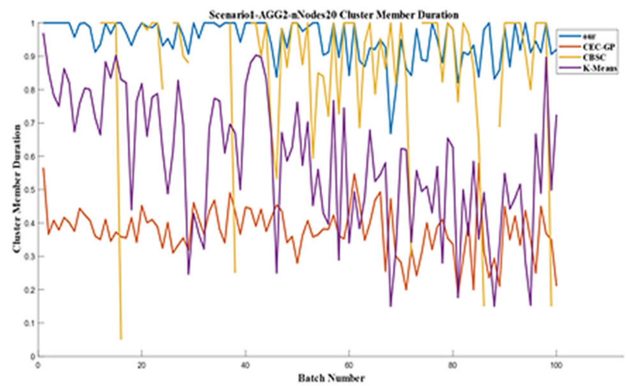


FIGURE 4. Time series of the cluster member duration for the batch number.

the least performance in terms of cluster member duration after K-means.

The third metric that was generated is the clustering efficiency depicted in Figure 5. We find that our algorithm has generated a high clustering efficiency compared with ECE-GP and K-means. In addition, we find that CBSC could accomplish higher values of clustering efficiency in some

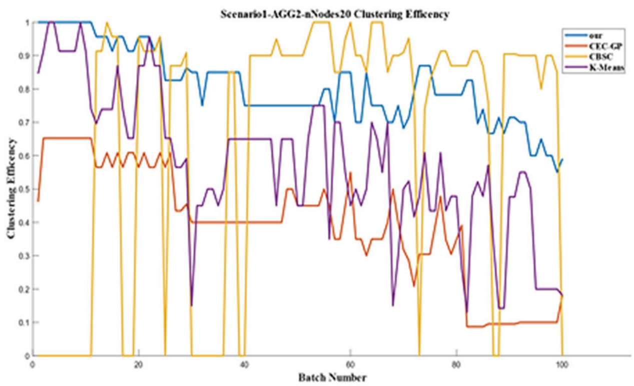


FIGURE 5. Time series of the clustering efficiency for the batch number.

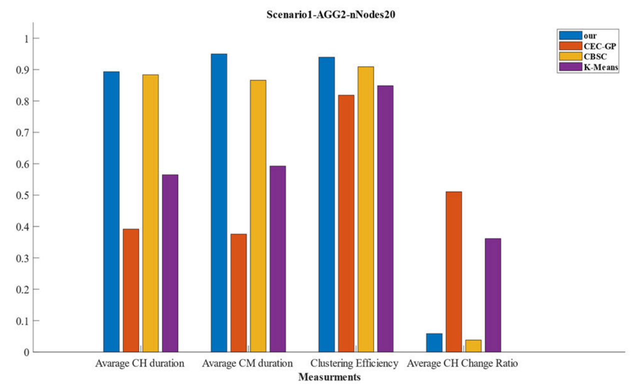


FIGURE 6. Overall evaluation metrics for our developed clustering and the benchmarks.

TABLE 6. Overall comparison results between our developed clustering and the benchmarks.

AGG	avrCHD	avrCMD	avrClutsEffic	CHCR
ours	**0.8933**	**0.9497**	**0.9394**	0.0585
CEC-GP	0.3915	0.3756	0.8182	0.5105
CBSC	0.8833	0.8659	0.9091	**0.038**
K_Mean	0.5648	0.5926	0.8485	0.3615
Improvement percentage	1%	10%	3%	-

time intervals. However, there was a high decline in clustering efficiency of the algorithm in some time intervals.

To summarize, we present the overall values of the metrics in Figure 6 for our model and the benchmarks. Our model has accomplished a higher average cluster head duration, cluster member duration, and clustering efficiency than the benchmarks. On the other hand, we find that the cluster head change rate of our model and CBSC were the least with significant improvement over the two benchmarks, namely, ECE-GP and K-means.

The overall results, as it is presented in Table 6, show that the average cluster head duration was 0.8933 compared with 0.8833 of CBSC, cluster member duration 0.9497 compared with 0.8659 of CBSC, 0.9394 compared with 0.9091 of a blob. On the other side, CBSC has accomplished the least clustering change ratio of 0.03.

TABLE 7. T-test values of our method for the benchmarks.

t-test	CEC-GP	CBSC	K_Mean
CHD	0.00E+00	6.62E-03	0.00E+00
CMD	0.00E+00	5.92E-02	0.00E+00
ClutsEffic	1.63E-06	3.13E-01	8.00E-08
CHCR	1.50E-07	4.75E-02	9.00E-08

In order to support our finding of the improvement achieved by our method, we perform a t-test analysis, as it is shown in Table 7.

Table 7 show that our model has accomplished statistical significance superiority over the benchmarks with a confidence level higher than 0.99.

V. CONCLUSION AND FUTURE WORKS

VANETs are gaining popularity for a variety of applications in smart cities and intelligent transportation systems. Various reasons make assuring dependable and stable VANET connectivity difficult. VANET clustering is a critical feature for serving routing protocols and enabling dependable VANETs. Clustering algorithms for VANETs function in a decentralized mode, which necessitates the inclusion of additional phases before making clustering judgments and may result in sub-optimality due to the local nature of the decentralized approach. Furthermore, the difficult design of the road environment can lead to a muddled clustering decision. The dynamic nature of clusters in VANETs, and 3D VANETs in general, makes this challenge more difficult. This article uses a centralized clustering technique based on a created Cauchy density model to address the challenge of 3D VANETs. A traffic generation model, a mobility model for creating driving behavior, and an algorithm for modeling road curvature based on an adjacency list that identifies the points of the road that create a straight-line segment have all been developed as part of the project. A clustering approach that defines a mobility vector and employs Cauchy-based density to allow cars to be added to their appropriate clusters is also included. For evaluation, we have conducted test scenarios in three locations with 3D architecture were considered and a comparison with benchmarks. The results revealed that our developed model outperformed the benchmarks by 1%, 10%, and 3% for average cluster head duration, average cluster member duration, and clustering efficiency, respectively. We will test this model along with the data dissemination algorithms in a 3D urban environment in future work.

ACKNOWLEDGMENT

The authors wish to acknowledge the tremendous support from the Department of Computer and Information Sciences (CISD), UTP, Malaysia. They would also like to thank the Centre for Telecommunication Research and Innovation (CeTRI), Faculty of Electronic and Computer Engineering, Universiti Teknikal Malaysia Melaka (UteM), Malaysia, for all the academic support and facilities.

REFERENCES

[1] S. Malik and P. K. Sahu, "A comparative study on routing protocols for VANETs," *Heliyon*, vol. 5, no. 8, Aug. 2019, Art. no. e02340.

[2] M. H. Hassan, S. A. Mostafa, A. Budiyono, A. Mustapha, and S. S. Gunasekaran, "A hybrid algorithm for improving the quality of service in MANET," *Int. J. Adv. Sci., Eng. Inf. Technol.*, vol. 8, no. 4, pp. 1218–1225, 2018.

[3] M. Mukhtaruzzaman and M. Atiquzzaman, "Clustering in vehicular ad hoc network: Algorithms and challenges," *Comput. Elect. Eng.*, vol. 88, Dec. 2020, Art. no. 106851.

[4] A. Katiyar, D. Singh, and R. S. Yadav, "State-of-the-art approach to clustering protocols in VANET: A survey," *Wireless Netw.*, vol. 26, no. 7, pp. 5307–5336, Oct. 2020.

[5] C. Wen, A. F. Habib, J. Li, C. K. Toth, C. Wang, and H. Fan, "Special issue on 3D sensing in intelligent transportation," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 4, pp. 1947–1949, Apr. 2021.

[6] I. Škrjanc, S. Ozawa, T. Ban, and D. Dovžan, "Large-scale cyber attacks monitoring using evolving Cauchy possibilistic clustering," *Appl. Soft Comput.*, vol. 62, pp. 592–601, Jan. 2018.

[7] A. B. Tambawal, R. M. Noor, R. Salleh, C. Chembe, and M. Oche, "Enhanced weight-based clustering algorithm to provide reliable delivery for VANET safety applications," *PLoS ONE*, vol. 14, no. 4, Apr. 2019, Art. no. e0214664.

[8] W. Ahsan, M. F. Khan, F. Aadil, M. Maqsood, S. Ashraf, Y. Nam, and S. Rho, "Optimized node clustering in VANETs by using meta-heuristic algorithms," *Electronics*, vol. 9, no. 3, p. 394, Feb. 2020.

[9] S. Sulistyo, S. Alam, and R. Adrian, "Coalitional game theoretical approach for VANET clustering to improve SNR," *J. Comput. Netw. Commun.*, vol. 2019, pp. 1–13, Jul. 2019.

[10] M. Kadadha, H. Otrok, H. Barada, M. Al-Qutayri, and Y. Al-Hammadi, "A cluster-based QoS-OLSR protocol for urban vehicular ad hoc networks," in *Proc. 14th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Jun. 2018, pp. 554–559.

[11] L. Li, W. Wang, and Z. Gao, "Driver's social relationship based clustering and transmission in vehicle ad hoc networks (VANETs)," *Electronics*, vol. 9, no. 2, p. 298, Feb. 2020.

[12] L. Zhu, C. Li, Y. Wang, Z. Luo, Z. Liu, B. Li, and X. Wang, "On stochastic analysis of greedy routing in vehicular networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 6, pp. 3353–3366, Dec. 2015.

[13] N. K. Gupta, R. S. Yadav, and R. K. Nagaria, "3D geographical routing protocols in wireless ad hoc and sensor networks: An overview," *Wireless Netw.*, vol. 26, no. 4, pp. 2549–2566, May 2020.

[14] R. Hyde, P. Angelov, and A. R. MacKenzie, "Fully online clustering of evolving data streams into arbitrarily shaped clusters," *Inf. Sci.*, vol. 382, pp. 96–114, Mar. 2017.

[15] X. Cheng and B. Huang, "A center-based secure and stable clustering algorithm for VANETs on highways," *Wireless Commun. Mobile Comput.*, vol. 2019, pp. 1–10, Jan. 2019.

[16] M. S. Talib, A. Hassan, T. Alamery, Z. A. Abas, A. A.-J. Mohammed, A. J. Ibrahim, and N. I. Abdullah, "A center-based stable evolving clustering algorithm with grid partitioning and extended mobility features for VANETs," *IEEE Access*, vol. 8, pp. 169908–169921, 2020.

[17] M. Ramalingam and R. Thangarajan, "Mutated k-means algorithm for dynamic clustering to perform effective and intelligent broadcasting in medical surveillance using selective reliable broadcast protocol in VANET," *Comput. Commun.*, vol. 150, pp. 563–568, Jan. 2020.

[18] *Bing Maps*. Accessed: Jul. 17, 2022. [Online]. Available: https://www.bing.com/maps

AHMED SALIH AL-OBAIDI (Member, IEEE) received the B.Sc. degree in software engineering from Al-Mansour University, Iraq, in 2007, and the M.Sc. degree in information technology from the UKM, in 2016. He is currently pursuing the Ph.D. degree with the Faculty of Electronic and Computer Engineering, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia. His research interests include radio frequency (RF), wireless ad hoc and vehicular networks, the IoT, and networks. He is a member of PMI.

MOHAMMED AHMED JUBAIR was born in Iraq. He received the B.Sc. degree in computer engineering from Al-Marrif University, Iraq, in 2012, the M.Sc. degree from the Computer Science Networks Department, Universiti Kebangsaan Malaysia (UKM), Malaysia, in 2017, and the Ph.D. degree from the Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia (UTHM), Malaysia, in 2021. He worked as a Teaching Assistant with Al-Marrif University, from 2012 to 2014. He is currently working as a Senior Lecturer at the Department of Computer Technical Engineering, College of Information Technology, Imam Ja'afar Al-Sadiq University, Iraq. His research interests include software engineering, artificial intelligence, communication systems, networking, and machine learning.

IZZATDIN ABDUL AZIZ received the Ph.D. degree in information technology from Deakin University, Australia, working in the domain of hydrocarbon exploration and cloud computing. He is currently heading the Centre for Research in Data Science (CeRDaS), Universiti Teknologi PETRONAS (UTP), where he focuses in solving complex upstream oil and gas (O&G) industry problems from the viewpoint of data analytics, machine learning, big data, and AI. He is working closely with O&G companies in delivering solutions for complex problems, such as offshore O&G pipeline corrosion rate prediction, O&G pipeline corrosion detection, machine failure predictions, securing data on clouds and designing and implementing metocean prediction system and bridging upstream and downstream oil and gas businesses from the viewpoints of data science, machine learning, and data analytics. He has secured more than USD five million worth of industrial project grants for numerous data analytics projects. He is also an Executive Committee Member of the IEEE Computer Society Malaysia and a member of the Malaysia Science Academy Special Interest Group in Machine Learning.

MOHD RIDUAN AHMAD (Member, IEEE) received the degree (Hons.) in computer system and communication engineering from the Universiti Putra Malaysia, in 2003, the M.Eng. degree with a specialization in cross-layer design of MAC protocols for multi-in multi-out-based wireless sensor network from the University of Wollongong, Australia, in 2008, and the Ph.D. degree with a specialization in atmospheric discharges from Uppsala University, Sweden, in 2014. From 2015 to 2016, he was with the MIT, USA, where he focused on the understanding and characterization of microwave radiation emitted by lightning flashes. He is currently the Head of the Broadband and Networking (BBNET) Research Group, Universiti Teknikal Malaysia Melaka (UTeM). He is also an Associate Professor with the Faculty of Electronics and Computer Engineering, UTeM.

SALAMA A. MOSTAFA (Member, IEEE) received the B.Sc. degree in computer science from the University of Mosul, Iraq, in 2003, and the M.Sc. and Ph.D. degrees in information and communication technology from the Universiti Tenaga Nasional (UNITEN), Malaysia, in 2011 and 2016, respectively. He is currently the Head of the Center of Intelligent and Autonomous Systems (CIAS), Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia (UTHM). He has produced more than 200 articles in journals, book chapters, conferences, and tutorials. He has completed 14 industrial projects and 23 research projects. His research interests include autonomous agents, adjustable autonomy, human–computer collaboration, machine learning, optimization, and software quality assurance.

HAIRULNIZAM MAHDIN (Member, IEEE) received the Ph.D. degree from Deakin University, Australia, in 2012. He is currently an Associate Professor with the Department of Information Security and Web, Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia. He is also working as the Deputy Dean of Research, Development, and Publication. His research interests include data management, the IoT, RFID, information security, software engineering, and web technology.

ABDULLAH TALAAT AL-TICKRITI was born in Iraq. He received the B.Sc. degree in computer engineering from Al-Marrif University, Iraq, in 2012, the M.Sc. degree from the Computer Science Networks Department, Universiti Kebangsaan Malaysia (UKM), Malaysia, in 2017, and the Ph.D. degree from the Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia (UTHM), Malaysia, in 2021. He worked as a Teaching Assistant with Al-Marrif University, from 2012 to 2014. His research interests include software engineering, artificial intelligence, communication systems, networking, and machine learning.

MUSTAFA HAMID HASSAN received the B.Sc. degree in computer engineering from Al-Marrif University, Iraq, in 2012, and the M.Sc. degree from the Computer Science Networks Department, Universiti Kebangsaan Malaysia (UKM), in 2017. He is currently pursuing the Ph.D. degree with the Faculty of Computer Science and Information Technology, Universiti Tun Hussen Onn Malaysia (UTHM). He is also working as a Senior Lecturer at the Department of Computer Technical Engineering, College of Information Technology, Imam Ja'afar Al-Sadiq University, Iraq. His research interests include communication systems, artificial intelligence, networking, and machine learning.