IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received March 21, 2020, accepted April 5, 2020, date of publication April 10, 2020, date of current version April 22, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2987065

A Network-Centric Analysis for the Internet of Vehicles and Simulation Tools

SHAHAB TAYEB[1], (Member, IEEE), SUMANJIT GILL[1], FLETCHER TRUEBLOOD[2], ROBERT WONG[1], AND MATIN PIROUZ[2], (Member, IEEE)
[1]Department of Electrical and Computer Engineering, California State University, Fresno, CA 93740, USA
[2]Department of Computer Science, California State University, Fresno, CA 93740, USA

Corresponding author: Shahab Tayeb (tayeb@csufresno.edu)

This work was supported by the Fresno State Transportation Institute under Grant SB-1.

ABSTRACT The Internet of Vehicles (IoV) is an emerging research framework, with network and graph theories as two of the major fields. Researchers in these topics use a variety of tools and approaches to simulate and perform experimentation on their proposed methodologies. A comprehensive study to facilitate the selection of such simulation tools is lacking from the literature. In this work, we provide a systematic review of the different simulation platforms. More precisely, the contributions of this paper are fourfold: firstly, we propose a two-tier hierarchical taxonomy based on the trends in the literature; secondly, we investigate the strengths and limitations of different simulation platforms; and thirdly, we take a network theoretic approach to identify the patterns in IoV research. To this end, we create a network of the publications and populate the edges among them. Community detection is performed using Louvian and Clauset-Newman-Moore algorithms. To the best of our knowledge, this is a novel approach to reviewing the literature which provides a more in-depth analysis of the trends in the literature. Finally, we review the common datasets for IoV experimentation.

INDEX TERMS Complex networks, community detection, OSM, SUMO, VANET, VEINS.

I. INTRODUCTION

The Internet of Vehicles (IoV) is an inter-network of autonomous and connected vehicles that interact with one another using an ensemble of wireless protocols. IoV has gained much interest among researchers and practitioners alike, and many have applied graph and network theories to the IoV for improving its design and implementation. Extensive experimentation is required to evaluate the practicality of the proposed approaches.

With the prevalence of network and graph tools, researchers have applied such tools to the IoV. Figure 1 shows the number of such papers from 2013 to 2019, as well as the number of unique authors. From 2013 to 2016, the numbers were relatively low, which doubled from 2016 to 2017. This growth is noticed until today, with over 20 published studies that apply network and graph theories to the IoV.

Many researchers lack the resources to implement a ''physical'' IoV for experimentation [1]–[3]; therefore, simulation tools play an important role in measuring the performance

FIGURE 1. The number of papers and authors who applied graph and network theories to the IoV from 2013 to 2019.

of the proposed research. The selected simulator should be capable of implementing the different features of the IoV as well as supporting performance metrics that allow the researchers to evaluate their methods. Hence, selecting ''the best'' simulator for one's research is a decision that needs to be made with much thought. The environment the simulations are performed, what its capabilities are, its effectiveness in

The associate editor coordinating the review of this manuscript and approving it for publication was Ruofei Ma.

providing meaningful results, and its relevancy to the project are all aspects to be taken into account when deciding on a simulation tool.

This paper provides an evaluation of both simulation tools and experimentation methods across different applications of network and graph theories in the IoV. Based on a systematic review of the literature, we propose a taxonomy of papers that incorporate network and graph theories in addition to classifying them based on the particular IoV application. Collectively, these serve as a starting point for researchers intending to research the IoV; and these provide insights on the research gaps and how to properly conduct meaningful and effective experimentation.

To perform this research, we review the papers with keywords 'network theory,' 'graph theory,' 'IoV,' 'Intelligent Transportation Systems,' and 'VANET' and identified the most relevant ones. Our search yielded 1,329 papers published between 2015 and 2019. We considered a total of 59 papers out of them as the most relevant to the proposed taxonomy. Many publications applies machine learning, deep learning [4]–[7], and/or game theory [8], which would skew our findings so we separated them and processed the remaining for further analysis. The proposed taxonomy was iteratively optimized and 29 additional papers were considered. A summary of the papers reviewed and classified in the first round are presented in Table 1. The proposed taxonomy is aligned with the framework developed by [9].

TABLE 1. The number of papers per category.

Category	Quantity
V2V Communications	30
Resources Optimization	11
Predictions of Vehicle Behavior	10
Offloading	2
Congestion Control	4
Pattern Mining	2

To the best of our knowledge, this work is the first attempt in applying complex network theory to surveying the literature. A network or graph is defined as a collection of nodes and edges [10]. In the context of the IoV, nodes are either vehicles or Road-Side Units (RSU) and the edges are the communication links between them. IoV nodes are mobile, resulting in a dynamic networks, with temporal and spatial features and a changing network topology over time [10]. For the purpose of this paper, keywords are considered nodes. Edges are the connection between keywords used within the same paper. This will allow for a network of keywords to analyze further.

A. SUMMARY OF CONTRIBUTIONS

This paper has four major contributions to the body of knowledge:

- Inspired by recent trends in literature, we iteratively categorize the research in terms of application and broader focus. The findings form a novel, two-tier hierarchical taxonomy that categorizes the research at the intersection of the IoV and network theory.
- We conduct a systematic review on IoV simulators. We evaluate the strengths and limitations of current simulations and examine the trends in using specific simulators for particular research applications.
- We apply complex network theory to our findings and identify the different communities based on the keywords, which demonstrates the research trends.
- We also conduct a survey on the most commonly used datasets for IoV research.

TABLE 2. Summary of the acronyms.

Acronym	Meaning
IoV	Internet of Vehicles
VANET	Vehicular Ad-Hoc Network
V2V	Vehicle-to-Vehicle
RSU	Road-Side Unit
SUMO	Simulation of Urban Mobility
OMNeT++	Objective Modular Network Testbed in C++
VEINS	Vehicles in Network Simulation
OSM	OpenStreetMap
NS	Network Simulator

B. ORGANIZATION

The rest of this manuscript is organized as follows: Section II provides an overview on the fundamental concepts of network theory, explaining the concepts used in section IV. Section II also introduces the IoV and its underlying architecture. The proposed taxonomy is presented in Section III followed by a summary of the literature for each classification. Section IV applies network theory on the literature to visually represent the research trends. Section V provides a thorough analysis of the simulation tools used in the body of literature, followed by a discussion on research gaps in Section VI. Section VII summarizes the findings of this manuscript.

II. PRELIMINARIES

A. FUNDAMENTALS OF THE INTERNET OF VEHICLES

The IoV marks the next frontier of a new digital revolution in Intelligent Transportation Systems. They enable the transportation industry to increase productivity, the city services to converge, vehicles to become (semi-) autonomous, and communities to become smarter. There has been much research on the design, evaluation, testing, and verification of (semi-autonomous) vehicles and there are various communications technologies that are being deployed as the IoV backbone. With the proliferation of support for autonomous and connected vehicles in private and public sectors, many IoV of different types, sizes, and sensitivity levels exist. Figure 2 presents a high-level logical topology of the IoV backbone, where Autonomous Vehicles (AV), Road-Side Units (RSU), Aggregator Stations (AGT), are connected to one another.

The integration of vehicular networks and social networks gives birth to the vehicular social networks. Wang *et al.* [11] provided a review for the privacy challenges in this context.

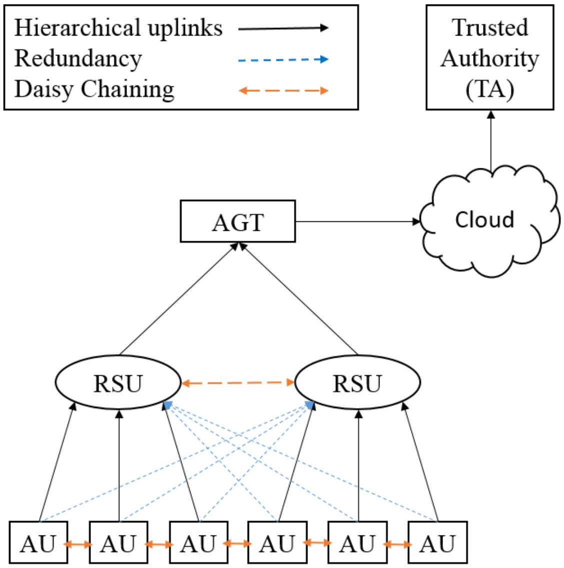


FIGURE 2. IoV network backbone.

Ning *et al.* [12], on the other hand, presented a quality-aware model for vehicular social networks. Xiong *et al.* [13] model IoV as a cyber-physical-social system and analyze its control and management.

IoV benefits from the advances in data analysis and big data techniques. Xu *et al.* [14] studied the mutual impact of IoV and Big Data on one another. Wang *et al.* [15] proposed a software-defined IoV to separate the control and data planes and thereby, reducing the number of rules without jeopardizing the real-time transmission performance. Ma *et al.* [16] reviewed the artificial intelligence as applied to the IoV by analyzing three key applications of perception, localization and mapping; and decision making. Their review included other emerging technologies of big data, augmented/virtual reality, and 5G communication. Kebria *et al.* [17] analyzed the impact of Convolutional Neural Network layers, filters, and filter sizes to the image processing in the IoV. They identified proper filter allocation requirements for this task and demonstrated that the performance stays largely unaffected by the varying number of filters.

In terms of routing, there have been numerous studies analyzing the most efficient routing scheme for a given topology. Cheng *et al.* [18] reviewed the problem of IoV routing and proposed four classifications of topology, position, map, and path-based routing approaches.

Other noteworthy studies in the IoV realm are as follows: Guo *et al.* [19] analyzed the problem of vehicle dynamic state estimation, including velocity, sideslip angle, yaw rate, and roll angle. Cheng *et al.* [20] developed an urban-road connectivity model using possibility, data forwarding time, link forwarding capability, and packet error rate. Cheng *et al.* [21] proposed a position prediction framework for IoV to be used by medical units. Their framework considered vehicle's attributes, road conditions, and driving load.

B. NETWORK PRELIMINARIES

As opposed to full-mesh networks [10], vehicular networks consist of nodes that may or may not be connected to other network components. In the case of our keyword network, a paper not sharing any keywords with the rest of the papers will result in a component. In networks with two or more components, a component is considered to be strongly connected if there is a directed path between every possible pair of nodes belonging to the component. If an edge in the IoV is considered to be directional, a strongly connected component would be able to disseminate messages to every vehicle in the component.

Finding the shortest paths in the IoV is an important task [22], [23]. Shortest paths can be used to optimize message dissemination [24]. They can also facilitate the identification of the best travel routes [25]. With the diameter of a network defined as the longest shortest path between two nodes within the network, it is not uncommon for the diameter of a network to be small even when the network is large [10]. The average shortest path in a network measures the efficiency of information [10]. In the case of the keyword network, a larger average shortest path can indicate there is a larger variance in types of keywords used. A smaller average shortest path can indicate the keywords used are more closely related. For the analysis of the keyword network a large diameter indicates less correlation between keywords. A smaller diameter means there is more correlation among keywords.

The degree of a single node, K, is the number of ingress and egress links [10]. The degree of node K is calculated using Equation 1

$$K_i = \sum_{j=1}^{n} A_{ij} \tag{1}$$

where A is an adjacency matrix representing a graph and K_i is the degree of node i. If there is a connection from node i to node j, $A_{ij} = 1$, otherwise $A_{ij} = 0$.

In the keyword network, a node with a higher degree means it has been used more in the papers covered. A smaller degree means the keyword has not been used often.

Density is a measure of how connected or sparse a network is. A higher density makes a network more connected. Density is ratio between the number of edges a network has over the maximum number of edges a network can have (full-mesh) [10]. Density is calculated using:

$$\rho = \frac{2m}{n(n-1)} \tag{2}$$

where m is the number of edges and n number of nodes.

A higher density in the keyword network correlates to keywords being paired together often in the papers covered. It is worth noting that real world networks have a low density and considered sparse [10].

Closeness centrality is a measure of a node's average distance to other nodes in the network. Closeness centrality is mathematically defined as:

$$Cl_i = \frac{n}{\sum_j d_{ij}} \tag{3}$$

where n is the number of nodes in the network and d_{ij} is the distance from node i to all other nodes in the network [10]. Closeness centrality is used to find the shortest time ordered paths in a temporal network [26].

Betweenness centrality plays a role in determining a node's importance in a network. Betweenness centrality captures how many times a node lies on the shortest path from/to other nodes in the network [10]. Betweenness centrality is defined as:

$$B_i = \sum_{st} n_{st}^i \tag{4}$$

If node i lies on a shortest path from s to t, $n_{st}^i = 1$. If s does not lie on a shortest path from s to t, $n_{st}^i = 0$ [10]. In the keyword network a keyword with a high betweenness centrality means it lays on more shortest paths. Being on more shortest paths makes the keyword more important as it has been used often in the papers covered [27].

Katz centrality measures the influence a node has on its immediate neighbors. Katz centrality ensures all nodes will have some importance in the network. This is done by giving all nodes some base influence level [10]. In the keyword network, a keyword with a Katz centrality will have more influence on the network, which makes it more important.

III. THE PROPOSED TAXONOMY

This section discusses the categorization performed on the papers in terms of their experimentation method and application.

We use the following labels for experimentation methods: Vehicle Simulation, Dataset, Mapping, Routes/Trajectory, and Theoretical. Papers under the label Vehicle Simulation conduct a simulation that has vehicles (represented by a simple node network) interacting with one another; their datasets are direct results of the vehicular simulation. Papers that fall under the category of Dataset include datasets based on real-life conditions. Papers that are labeled under the category of Mapping conduct simulations that are based on real-world maps/routes of cities. Papers categorized under Routes/Trajectory focus on using common routes and where the trajectories lie based on available real-world information. Papers that are categorized as Theoretical are the ones that contain a synthetic generated datasets for use in analysis.

The labels for application to the IoV are as follows: V2V Interaction, Congestion Control, Optimization, Predictions, Pattern Mining, and Offloading. Papers under the label of V2V Interaction constitute papers with vehicles interacting with one another (whether it be privacy preservation, communication ranges, etc.). Papers under the label of Congestion Control provide insight in how to alleviate congestion control. Papers that are labeled under Optimization are centered around finding optimal paths, how to optimize resource usage, etc. Papers that are under the label of Predictions focus on how to predict flow of traffic or speeds of other vehicles in the network. Papers that are under the category of Pattern Mining discuss how to reconstruct data correctly and reliably to use in network communications (e.g., how to recognize pedestrians). There are too few pattern mining papers in our taxonomy because we found that most of these papers use machine learning instead of network or graph theory. Papers that are categorized as Offloading focus on how to distribute vehicles or paths to alleviate any potential for overuse of public vehicles, roads, etc. Figure 3 shows the different research applications used each year from 2013 to 2019. Figure 4 summarizes how each of the papers analyzed are categorized in the application categories. Each method, the right hand side, is linked to its specific application on the left hand size. The nodes sizes change depending on their degree. The edge widths vary depending on the amount of time a method and specific application appear together.

Table 3 summarizes the categorization of the surveyed literature.

A. VEHICLE TO VEHICLE INTERACTIONS

One of the main features of a VANET is the spatial-temporal aspect, creating a dynamic graph [26], [28], [37], [41], [81]. Graph theory and complex network theory are used to handle this distinguishing feature of the IoV [29], [38], [42], [81], [82]. Using network density, diameter, and connectivity, Loulloudes *et al.* showed vehicles traveling on freeways tend to have longer connections with one another as compared to urban routes [28]. Path ordering, temporal closeness centrality, and strongly connected components have been used to determine the reachability of a VANET [26]. Temporal and spatial features were used in finding the time evolution of important nodes, used in Intelligent Driver Model with Lane Changes developed by Feng *et al.* [37]. Through the use of temporal approximation, Feng *et al.* demonstrated the positive impact of vehicle density on network invulnerability [35]. Convolutional neural networks and spatial temporal features were coupled to aid in the detection of anomalies in traffic flow [38]. Clustering methods were used by Avcil *et al.* to track network topology changes and to manage network stability [29]. Combining the cooperative sensing spectrum and spatial temporal features of a VANET, a reduction in user selection frequency was achieved [42]. Store-and-forward relays have been used to restore connectivity by making use of the spatial temporal aspect of VANETs [39], [40].

Another core feature of the IoV is inter-vehicle communication, making message dissemination within the IoV one of the core issues [27], [30], [36]. Message dissemination causes overhead that must be considered to keep a VANET from being overloaded [27]. Using neighbors within one and two hops of each vehicle, a complex networks for data dissemination protocol among vehicles was used to find the best relay nodes to retransmit data. The relay nodes are selected using degree centrality and betweenness centrality [27]. Through the use of bipartite graphs

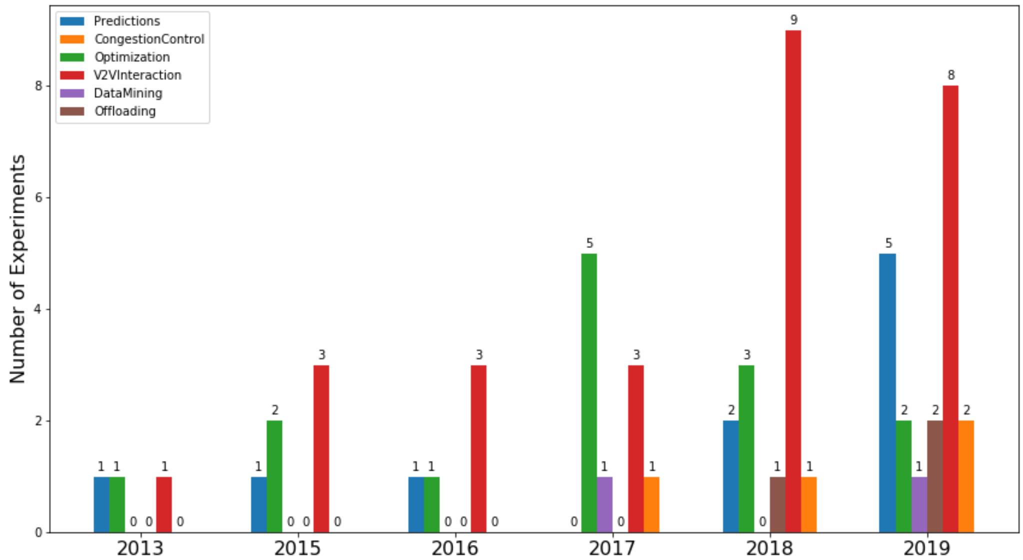


FIGURE 3. Different IoV methods from 2013 to 2019.

and the Kuhn-Munkres algorithm, Chai *et al.* [57] evaluated the transmission performance of relay vehicles to minimize overhead and maximize coverage. The Maximum Broadcast Efficiency Relaying (MBER) Algorithm is a graph theory based optimization algorithm, developed to lessen the vehicle as an obstacle effect within VANETs. MBER outperforms distance-based routing algorithms [58]. The use of an undirected weighted graph to represent the IoV creates a scale free network property, where only a few nodes have enough connections that can be considered as network hubs. With a scale free network and the small world effect, Wang *et al.* [36] developed a hub selection scheme for collecting data. The data is then used for optimizing message dissemination [83].

Vehicles belonging to a VANET are highly mobile, thus making the topology of a VANET dynamic [25]. However, the dynamic topology does have some structure to it which comes from the underlying road structure. This structure has some inherent patterns, i.e. stop signs and traffic signals, which can lead to some patterned mobility [32], [33]. By applying graph theory to capture the dynamic nature of VANETs, Eiza *et al.* [25] makes use of the evolving nature of the network to find reliable routes for highway travel using Dijkstra's algorithm. A comparison between Adhoc On-demand Distance Vector (AODV) and Link State Geographic Routing protocol(LSGR) revealed that LSGR outperforms AODV in packet drop, throughput, and average end-to-end delay [33].

B. OFFLOADING

An issue with VANETs and other mobile networks is that devices in the network have limited resources for performing operations [45], [80]. To solve this problem, offloading calculations using fog computing, edge computing, or cloud computing are often used to solve the issue of limited resources [45], [80]. Luo *et al.* propose a method using a graph theory based algorithm to improve data sharing and cooperation in edge computing assisted 5G network [45]. Another approach is the quality of experience for the user, which Zhang *et al.* improve by introducing mobile edge computing using an approach based on Skyline Graph Model and Directed Acyclic Graph theory to store and update the network [80].

C. CONGESTION CONTROL

Traffic congestion is a common problem due to the dynamic environment, rush hours, and deadlocks [61], [75], [76]. VANETs have the potential to solve this problem using a routing technique to direct vehicles in a way that reduces congestion in traffic [61], [63], [71], [75], [76]. Xia *et al.* propose a greedy traffic light and queue aware routing protocol (GTLQR) using network theory to reduce traffic congestion [75]. Perronnet *et al.* also focus on routing at intersections, but also includes road reservation using a graph theory method to prevent deadlocks in traffic [76]. Other papers focus on an opportunistic routing protocol like the Dempster-Shafer evidence theory based method proposed by Kashani *et al.* or propose temporal graph methods to solve the routing problem due to the high mobility of a VANET, such as in Zhao *et al.* [61], [63]. Khan *et al.* propose a different approach using oriented evolving graph and clustering to improve reliable routing requests in the network [71].

D. PREDICTIONS

Predicting components of a VANET such as traffic speed and the best route for vehicles is important for being able to better control the traffic and to optimize road services [25], [69], [70]. The majority of papers classified under

TABLE 3. Taxonomy of graph and network theory papers.

Method	Application	Year	Paper
V2V Communication	temporal/spatial	2015	[28] [29]
	social vehicle interactions	2015	[30]
	data dissemination	2015	[31]
	network stability	2015	[29]
	scalability of VANET	2016	[32]
	routing	2016	[33]
	trust/privacy preservation	2016	[34]
	temporal/spatial	2017	[26] [35]
	data dissemination at RSU level	2018	[27]
	data dissemination	2018	[36]
	temporal/spatial	2018	[37] [38] [39] [40] [41]
	anomaly detection	2018	[38]
	heal disconnectedness	2018	[39]
	transmission time, store and forward	2018	[40]
	cooperative sensing	2018	[41]
	temporal/spatial,cooperative sensing prediction	2019	[42]
	resource allocation	2019	[43]
	edge computing	2019	[44] [45]
	data dissemination	2019	[31] [46] [47]
	privacy preservation	2019	[48]
	network security	2019	[49]
	temporal/spatial	2019	[50]
Resource Optimization	optimization of RSU delay reduction	2015	[51]
	data dissemination	2015	[24]
	prediction, routing	2016	[52]
	routing	2017	[53]
	multipath routing	2017	[54]
	optimize RSU placement	2017	[55]
	optimizing power consumption, network performance	2017	[56]
	dissemination	2017	[57]
	dissemination	2018	[58]
	optimization of RSU	2018	[59]
	data dissemination	2019	[60]
	routing, trust	2019	[61]
	ride sharing	2019	[62]
	temporal/spatial routing	2019	[63]
	throughput/channel allocation	2019	[64]
Traffic Predictions	routing	2013	[25]
	feasible paths based on data traffic	2015	[65]
	prediction, recommender system	2016	[66]
	prediction of packet collision	2018	[67]
	routing	2018	[68]
	real time speed estimation	2019	[69] [70]
	reliable routing	2019	[71]
	traffic flow	2019	[72]
	temporal/spatial traffic prediction	2019	[73]
Congestion Control	congestion	2017	[74]
	routing	2018	[75]
	deadlock prevention	2019	[76]
	multi hop routing	2019	[77]
Pattern Mining	IoV forensics	2017	[78]
	data mining	2019	[79]
Offloading	Edge Services	2018	[80]
	Edge Computing	2019	[45]

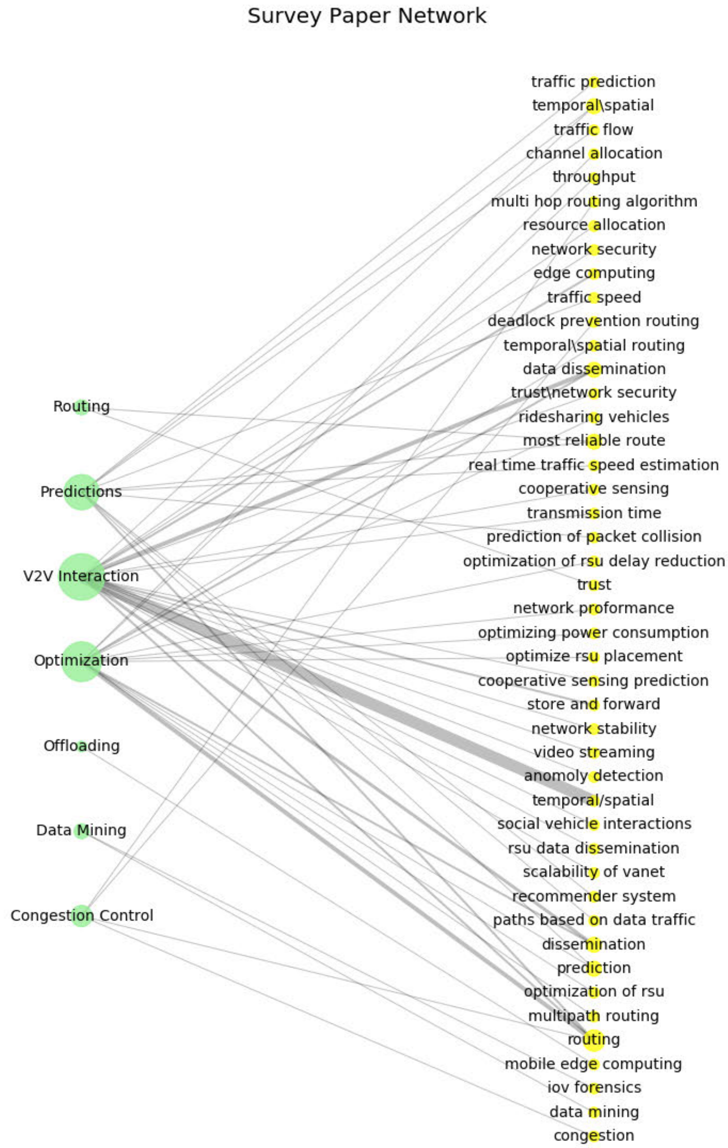


FIGURE 4. Bipartite network of methods, left hand side, and specific application, right hand side. Node sizes vary depending on the degree. Edge widths vary depending the number of times a method and specific application occur together.

predictions focus on real time traffic speed to enhance routing [69], [70], [72], [73]. Tao *et al.* suggests a Delay-based Spatial-Temporal Autoregressive Moving Average model to deal with the travel delay problem in short term traffic prediction. Other papers focus on deep learning such as in Elbery *et al.* where they propose a graph convolutional generative autoencoder which uses deep learning and graph theory to predict real-time traffic speeds [69]. A similar paper by Ge *et al.* propose a temporal graph convolutional network technique to capture the spatial/temporal component of the traffic speed prediction problem [70]. Kim *et al.* take a slightly different approach using a recurrent neural network with a spatio-temporal graph to learn and predict traffic features [73]. Other prediction papers focus on creating routes updated by an evolving graph such as in Eiza *et al.* instead of predicting traffic speeds [25]. Elbery *et al.* propose another prediction application for suggesting carpools using a weighted bipartite graph between users and places [66].

E. OPTIMIZATION

Transmission of communications between vehicles is critical for the IoV [84]. Using stochastic network calculus theory and Dijkstra's algorithm, Peng *et al.* [53] develop a methodology to minimize end-to-end transmission delay between source vehicles and destination vehicles. This is done while optimizing routing constraints. Chai *et al.* [57] use bipartite graphs, network calculus theory, and the K-M algorithm to evaluate the usage of relay vehicles in optimizing transmission performance. Communications should also be optimized to follow the shortest path between vehicles when transmitting information. Kochhar and Mandoria [24] proposed a protocol derived from ant-based routing algorithms to determine which path is optimal. In specific cases of traffic jams and accidents in the network, the communications between vehicles needs to spread rapidly, which falls prey to transmission delay. Wang *et al.* [60] create a crowd sensing-based system along with a cluster-based optimization framework to minimize average delay, delivery ratio, cost, etc. in these circumstances. Huang *et al.* [58] develop a graph theory-based optimization algorithm for VANETs, the maximum broadcast efficiency relaying (MBER) algorithm. The messages relating to road safety are relayed effectively while also meeting criteria of minimal delay and reliability.

Optimization of RSU placement pertains to the well-being of the IoV ecosystem, since RSUs are facilitating structures for network communications. Bao *et al.* [59] develop a method using space-time graphs to determine where RSUs should be placed optimally, while still minimizing end-to-end delay. Another strategy to effectively place RSUs in the IoV is discussed in the work of Kim *et al.* [55], who use graph theory and their proposed polynomial running time approximation algorithm. Mehar *et al.* [51] attempt to address RSU placement with their two-step solution, ODEL, which is based on a genetic algorithm coupled with Dijkstra's algorithm. Their solution reduces delay as well as deployment cost.

Optimal routing for vehicles in the IoV is a key concern. Zhang *et al.* [52] employ geocast spatio-temporal routing to determine the best routing paths to provide optimal quality-of-service for data delivery, while maintaining a scalable method. Quality of service is also a focus in Cui *et al.* [62], who use network calculus and bipartite graphs to optimize wait time for passengers and autonomous vehicles. They also aim to minimize energy consumption and travel time. Further, they propose an online fleet management system to optimize traffic flow in the network.

Topology control in a VANET is challenging, as the network is temporal and transient. Koti and Kakkasageri [56] propose a dynamic algorithm based on stationary topological algorithms to address this concern. The approach enhances the practicality of packet transmission with guaranteed network connectivity.

Optimization of resources is critical to vehicular networks, as they are very limited. Throughput in the network needs to be maximized for optimal functionality. Li *et al.* [64] attempt to address this issue by creating an algorithm based on greedy algorithms for channel allocation.

F. PATTERN MINING

Security in the IoV is fundamental and a new form of malicious activity are likely to develop. Gündüz and Acarman [79] develop an investigation framework for IoV forensics based on the interaction provenance model. They store and identify evidence (e.g., communications) and verify its integrity. After experimentation, they ensured their system's applicability to the IoV and defense in adversarial scenarios.

Identification and interaction of objects and vehicles in the environment is another important aspect to the IoV. Hossain *et al.* [78] develop a multiple-object vehicle tracking system based on an affinity network of directed graphs. The accuracy and optimization of this system outperforms various clustering-based methods attempting to solve the same issue.

IV. NETWORK ANALYSIS OF THE LITERATURE

The network created in Figure 5 was created from the keywords listed in each author's paper. The nodes are created if a keyword is not present in the network yet. Once the nodes have been added to the network, edges are then created. Edges are added to link keywords used with in the same paper and if a keyword can be linked from a keyword from another paper.

After the network was created several network measures were applied. The first being node degree. The five nodes with the highest degree were: VANETs with a degree of 152, QoS with a degree of 28, graph theory with a degree of 27, optimization with a degree of 27, and delays with a degree of 23. The average node degree for the network was 10.91. From analyzing node degree, VANETs is by far the highest degree making it the most important node in the network.

The next network measure applied was, The average shortest path between keywords with in the network. The average shortest path was 2.22, meaning between each paper can be linked to another by an average of 2.22 keyword hops. The diameter of the network was 4, meaning the most separation between papers was just 4 keyword hops.

The next network measure applied was betweenness centrality, which measures the amount of times a node lies on a shortest path. The five nodes with the highest betweenness centralities were: VANETs at 0.89, heuristic algorithms at 0.086, IoT at 0.069, graph theory at 0.063, and ITS at 0.061. These nodes are common terms used when researching papers with in the IoV, so they do appear on more shortest paths between keywords.

The next network measure applied is Katz centrality, which measures the influence a node has with in a network. The top five nodes were: VANETs at 0.45, ITS at 0.37, IoT at 0.38, CRVANETs at 0.35, heuristic algorithms at 0.31. These nodes exhibit more influence throughout their neighbors and thus the network itself. These nodes are also common terminology when researching with in the IoV.

A. COMMUNITY DETECTION: LOUVIAN

Figure 6 was created by applying the Louvian community detection algorithm to the previously created keyword network in Figure 5. Communities are placed around a circle with the first community starting at zero degrees. The nodes represent keywords belonging to each community and are placed in a circular pattern around the community center. The nodes size depends on the degree of the node. The edge width varies depending on how many times a pair of keywords are used with in the same paper. The Louvian algorithm resulted

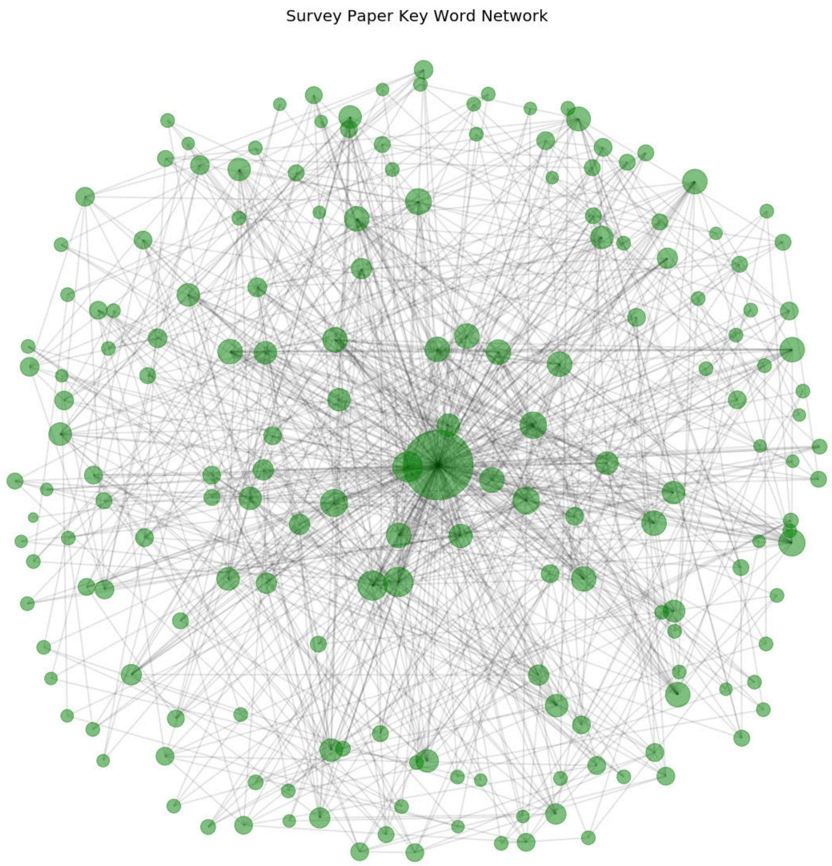


FIGURE 5. Keyword network constructed from surveyed papers. Node sizes vary based on the degree of the node. Edge width vary based on the frequency of the keyword pairs.

in eight communities. Keywords from each community can be seen in Table 4.

Figure 7 was created by counting how many times a keyword belonging to a specific category belonged to a community within Figure 6. Figure 7 shows the break category breakdown of each of the eight communities from Figure 6. Community names on the x-axis describe the major categories within the community. The V2V/Congestion community is comprised of mostly V2V papers and it contains the most congestion Papers of any other community. Offloading and optimization have a category of their own, meaning those communities only have keywords that belong to optimization and offloading papers. A common theme in the optimization category was, optimizing the routes taken. This can be seen in the community labeled routing/optimization.

B. COMMUNITY DETECTION: CLAUSET-NEWMAN-MOORE

Figure 8 was created by applying the Clauset-Newman-Moore algorithm to the previously created keyword graph network in Figure 5. Communities are distributed around a circle with the first community placed at the zero-degree mark. The nodes represent keywords belonging to each community and are placed in a circular pattern around the community's center. Node sizes vary based on the degree of the node; a higher degree makes the node bigger. Edge widths depend on the number of times a keyword pair is used with in the same paper, wider widths mean the keyword pairs appears more frequently with in a paper. The Cluaset-Newman-Moore algorithm resulted in six communities. Keywords from each community can be seen in Table 5.

Figure 9 was created by counting how many times a keyword belonging to a specific category belonged to a community with in Figure 8. Figure 9 shows the break category breakdown of each of the eight communities from Figure 6. Community names on the x-axis describe the major categories with in the community. The community label V2V is comprised of only keywords belonging to V2V papers.

C. COMMUNITY DETECTION RESULTS

When comparing Louvian and Clauset-Newman-Moore algorithms, Louvian provides more communities that reveal overlap among the categories of papers. For example, routing and optimization are common overlaps in categories, as finding the optimal path could have implications in most categories. Louvian also detected communities that are comprised

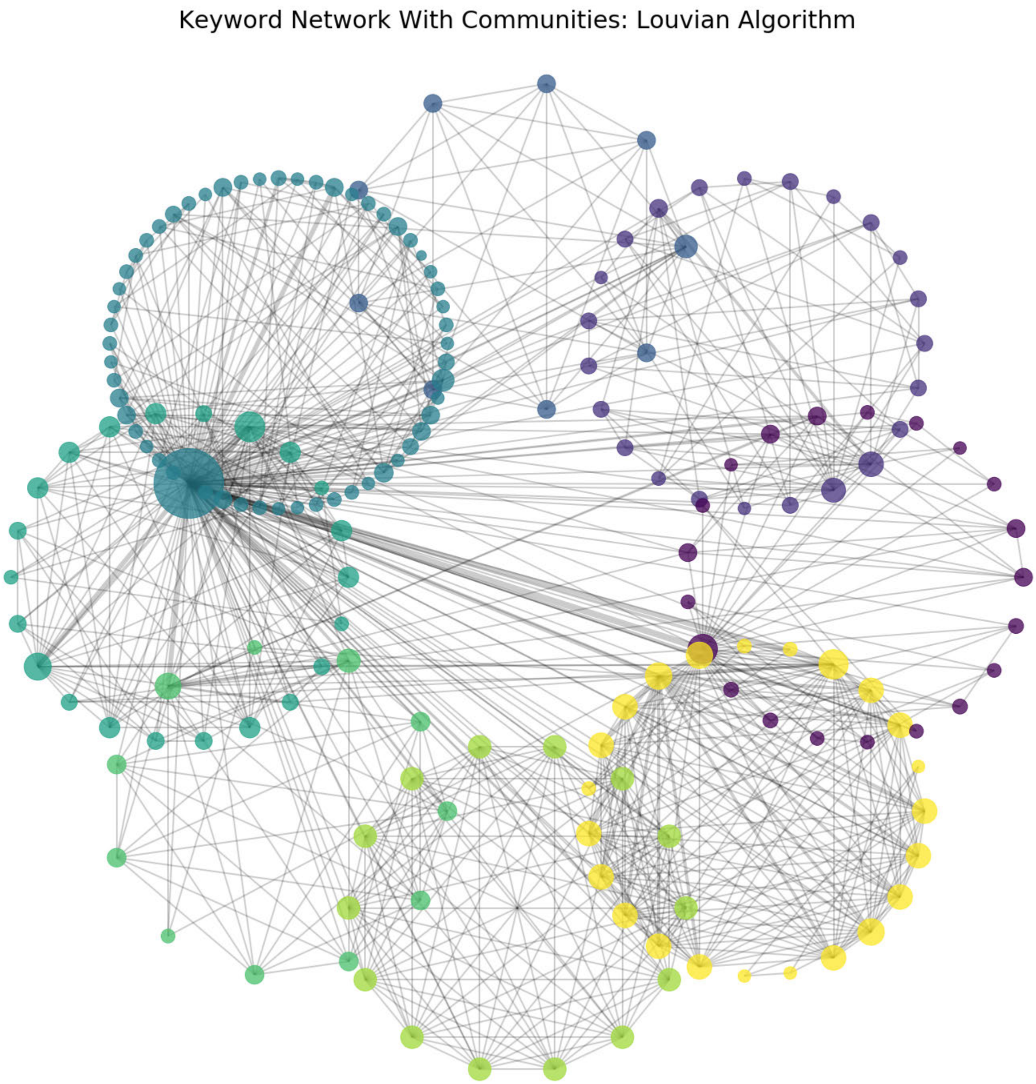


FIGURE 6. A network of communities created by applying the Louvian community detection algorithm to the network of keywords. Communities are placed around a circle with community one starting at the zero degree mark. Nodes are drawn in a circular pattern around the center of the community. Node sizes vary on the degree of the nodes. Edge width vary based on the frequency of keyword pairs.

of primarily one type of paper, offloading, V2V, and Optimization.

The Clauset-Newman-Moore algorithm provides more condensed communities, resulting in six communities, as opposed to eight communities detected by Louvian. Clauset-Newman-Moore provides the same communities comprised of one type of paper, V2V, optimization and offloading. The community labeled routing in Figure 9 is where the majority of routing papers belong. While the other communities are fairly balanced between five different types of papers.

V. A COMPARATIVE ANALYSIS OF SIMULATION TOOLS

Many IoV papers use simulators to validate their proposed method on the given application. There are numerous simulators and tools that can be used for the IoV, and most researchers use a combination. The most common simulation tools are Open Street Maps, Objective Modular Network Testbed in C++ (OMNeT++), Simulation of Urban Mobility (SUMO), Vehicles in Network Simulation (VEINS), Network Simulator (NS), and MATLAB. The different simulators and tools used for papers by year from 2015 to 2019 are shown in Figure 10.

A. OPEN STREET MAPS

Open Street Maps (OSM) is a free open source software created in 2004 by Steve Coast. OSM is a collaborative project where users can edit the map data. The geo-data created by the users is considered to be the primary achievement of the project. Since 2004, the number of OSM users has increased

TABLE 4. Louvian community keywords.

Community	Keywords
V2V/Congestion	SDVANETs, interference, capacity scaling law, graph theory, independent set, object detection, intelligent vehicles, radiofrequency interference, social-proximitybased mobility, interference relations, euclidean planar graph, LSGR, AODV, MANETs, NS235, synchronisation, temporal structure, time-ordered path, reachability, approximation algorithm, RSU deployment
V2V/Data mining	ITS, IoT, IoV Security, digital forensics, threat model, adversarial scenario, data dissemination, protocols, relays, complex networks, delay, social network of vehicles, cyber-physical systems, IoV, IoT architecture reference model, SAE J2735, traffic information collection, information diffusion, complex systems science, functional complexity, SOTDMA, probability of collision, probability of packet detection
Offloading	mobile edge computing, edge service, mobile communication, cloud computing, maintenance engineering, edge computing, heuristic algorithms, algorithm design and analysis service maintenance, skyline graph
V2V/Prediction	transmission delay, stochastic network calculus, routing algorithm, Dijkstra algorithm, VANETs, particle swarm optimization, multipath routing, learning automata, leapfrog, RSU placement, delay-bounded, space-time graph, bus-based routing, trajectory mining, time series analysis, effective throughput, network calculus, relay selection, Kuhn-Munkres algorithm, information propagation, broadcast efficiency, obstructing effect, effective coverage, street-centric routing, bus trajectory based graph, ant colony optimization, graphs, connectivity, analysis, complex networks measures, links, IVC, V2V, network topology, anomaly detection, network traffic, cnn, steady state distribution, SCF, rehealing delay, EMC, topology, signal processing algorithms, algorithm design and analysis, batteries, data plane security, trust management, trust model, 5G, temporal networks, centrality, node importance, simulation, invulnerability, temporal reachability graph, store-and-forward
V2V/Pre/Opt	ACO, MCQ, reliable routing, secure routing, QoS, delays, Evolving graph, routing reliability, vehicular networks, streaming media, standards, encoding, video coding, streaming media, standards, delays, encoding, video coding, roads, mathematical model, genetic algorithms, complexity theory
Routing/Optimization	evolving graph, routing protocols, VoEG model, Optimal number of clusters, cluster head, link reliability, routing, vehicle dynamics, road transportation, inference mechanisms, network reliability
Optimization	carpooling, recommender system, social VANETs, predictive models, Markov processes, wireless communication, conferences, mobile computing, social network services, geosocial, spatiotemporal, similarity detection, mobility modeling, Markov chain
V2V	clustering, stability, discrimination, merge, cognitive radio, signal detection, cooperative communication, correlation theory, diversity reception, optimization, probability, radio spectrum management, hybrid CSS scheme, double-threshold optimization problem, user selection problem, likelihood ratio testing, CRVANETs, hybrid cooperative spectrum sensing scheme, spatial-temporal correlation, cooperative spectrum sensing, soft combining, spatial temporal correlation, spatial-temporal diversity

to over two million. Each registered user has the ability to edit the OSM data through means of aerial photography and manual surveys, among other free sources. The OSM data is available under the open database license [85].

The data obtained from OSM can be used in such ways as producing paper maps, geo-coding of address and place names, electronic maps, and route planning. Some users of OSM include Facebook, Foursquare, and MapQuest [85].

While OSM is not a simulator itself, it is a tool that supplements simulations. By accessing the geo-data from the open database license, users can import that data into simulators to use real-world maps. Using real maps for simulations saves time on manually creating a map and allows for more realistic simulations.

B. OBJECTIVE MODULAR NETWORK TESTBED IN C++

The Objective Modular Network Testbed in C++ (OMNeT++) is primarily a network simulator based on C++. OMNeT++ uses C++ programming language to program the network simulator to determine what the nodes in the simulation do and then an omnetpp.ini file to specify the simulation parameters. OMNeT++ is not a network simulator by itself, but can be used to create network simulators for different types of networks such as ad-hoc networks, sensor networks, wired and wireless communication networks, etc. Data can then be easily exported into a CSV, JSON, OMNeT++ Scalar or Vector, or SQLite Scalar or Vector file to analyze the results using R, MATLAB, Python, or any other visualization programs. OMNeT++ has many simulation models and tools on their website and has easy to find tutorials for working with the program and has a user friendly interface. OMNeT++ works for customizing network simulators of various types, but does not offer vehicular movement features of other simulators, for example to simulate VANETs.

C. SIMULATION OF URBAN MOBILITY

The Simulation of Urban Mobility (SUMO) simulator is open source and designed specifically to simulate large road networks. The simulator includes structures for different types of roads (i.e. parking, surface roads, highways) and traffic management structures such as traffic lights, speed signs,

TABLE 5. Clauset-Newman-Moore community keywords.

Community	Keywords
Prediction/V2V	bus-based routing, Dijkstra algorithm, invulnerability, store-and-forward, trust management, 5G, stochastic network calculus, IVC, broadcast efficiency, simulation, SCF, centrality, trajectory mining, temporal networks, space-time graph, obstructing effect, complex networks measures, anomaly detection, learning automata, algorithm design and analysis, multipath routing, steady state distribution, particle swarm optimization, cnn, time series analysis, leapfrog, effective throughput, bus trajectory based graph, EMC, V2V, links, Kuhn-Munkres algorithm, street-centric routing, information propagation, relay selection, VANETs, effective coverage, connectivity, rehealing delay, node importance, data plane security, heuristic algorithms, delay-bounded, transmission delay, network calculus, routing algorithm, network traffic, topology, batteries, ant colony optimization, graphs, temporal reachability graph, trust model, RSU placement, analysis, network topology, signal processing algorithms
V2V/Data mining	reachability, functional complexity, intelligent vehicles, time-ordered path, complex networks, digital forensics, delay, optimization, IoV security, complex systems science, adversarial scenario, probability of packet detection, IoV, SOTDMA, temporal structure, relays, information diffusion, data dissemination, object detection, RSU deployment, cyber-physical systems, social network of vehicles, interference, graph theory, threat Model, radio frequency interference, IoT architecture reference model, interference relations, traffic informationc collection, IoT, independent set, ITS, SDVANETs, Approximation algorithm, euclidean planar graph, SAE J2735, protocols, capacity scaling law, probability of collision, social-proximitybased mobility
Routing	AODV, optimal number of clusters, delays, NS235, video coding, evolving graph, link reliability, video coding, cluster head, QoS, VoEG model, mathematical model, synchronisation, LSGR, encoding, streaming media, streaming media, inference mechanisms, standards, routing protocols, vehicle dynamics, MANETs, probability, reliable routing, vehicular networks, network reliability, secure routing, delays, road transportation, genetic algorithms, routing, encoding, MCQ, evolving graph, ACO, standards, roads, complexity theory, routing reliability
V2V	double-threshold optimization problem, cooperative spectrum sensing, radio spectrum management, diversity reception, merge, clustering, stability, cognitive radio, likelihood ratio testing, discrimination, signal detection, hybrid cooperative spectrum sensing scheme, correlation theory, spatial-temporal correlation, user selection problem, soft combining, spatial temporal correlation, cooperative communication, CRVANETs, hybrid CSS scheme, spatial-temporal diversity
Optimization	predictive models, mobile computing, geosocial, carpooling, Markov processes, social VANETs, spatiotemporal, wireless communication, recommender system, conferences, similarity detection, Markov chain, mobility modeling, social network services
Offloading	mobile communication, maintenance engineering, edge computing, edge service, skyline graph, mobile edge computing, cloud computing, algorithm design and analysis service maintenance

and sidewalks. The main simulator supports pedestrian, bicycle, railway, waterway, and standard vehicle simulation. Using a few files, road networks can be created along with other structures with each road having different speed limits, and another file can program routes for different vehicles in the simulation. SUMO includes additional features such as emergency vehicles, electric and hybrid vehicles, emissions, etc. The software has multiple extensions that can be used to enhance the simulator, and works well with OpenStreetMap to import maps around the world [86]. SUMO is easy to install and works well in simulating road networks but does not simulate network communication as required in VANETs.

D. VEHICLES IN NETWORK SIMULATION

Since OMNeT++ is a network simulator but lacks vehicular movement simulation, and SUMO is a road network simulator but lacks vehicular communications, Vehicles in Network Simulation (VEINS) is a software that combines the two programs to work in unison to simulate both components of a VANET. VEINS includes instructions about how to link the two simulators as well as links to the instructions for downloading and installing each of the programs. VEINS relies on models based on IEEE 802.11p and IEEE 1609.4 DSRC/WAVE network layers. VEINS allows for the addition of other parameters in the omnetpp file such as RSU placement and vehicles parameters that are not typically included in an OMNeT++ simulation. While VEINS itself does not simulate the network, if working with SUMO and OMNeT++ for simulating VANET, it connects the two programs to run simultaneously and collects the data on both the network and traffic components of the VANET [87].

E. NETWORK SIMULATOR

Network Simulator (NS) is a discrete-event computer network simulator. Its main purpose is for networking research, and supports simulation of routing, multi-cast protocols, and IP protocols over both wired and wireless networks. It also supports routing and queuing algorithms. In relation to the IoV, it allows for the implementation of IEEE 802.11p. The simulator is applicable for congestion control, as it allows researchers to understand mechanisms on a deeper level with the implementation of protocols, along with the documentation and logs of the simulations. NS requires multiple packages (i.e., Tcl/Tk with its header files, NAM, a C++ compiler) to be running in connection with itself. NS on its own requires 320 MB of disk space. Although the setup for

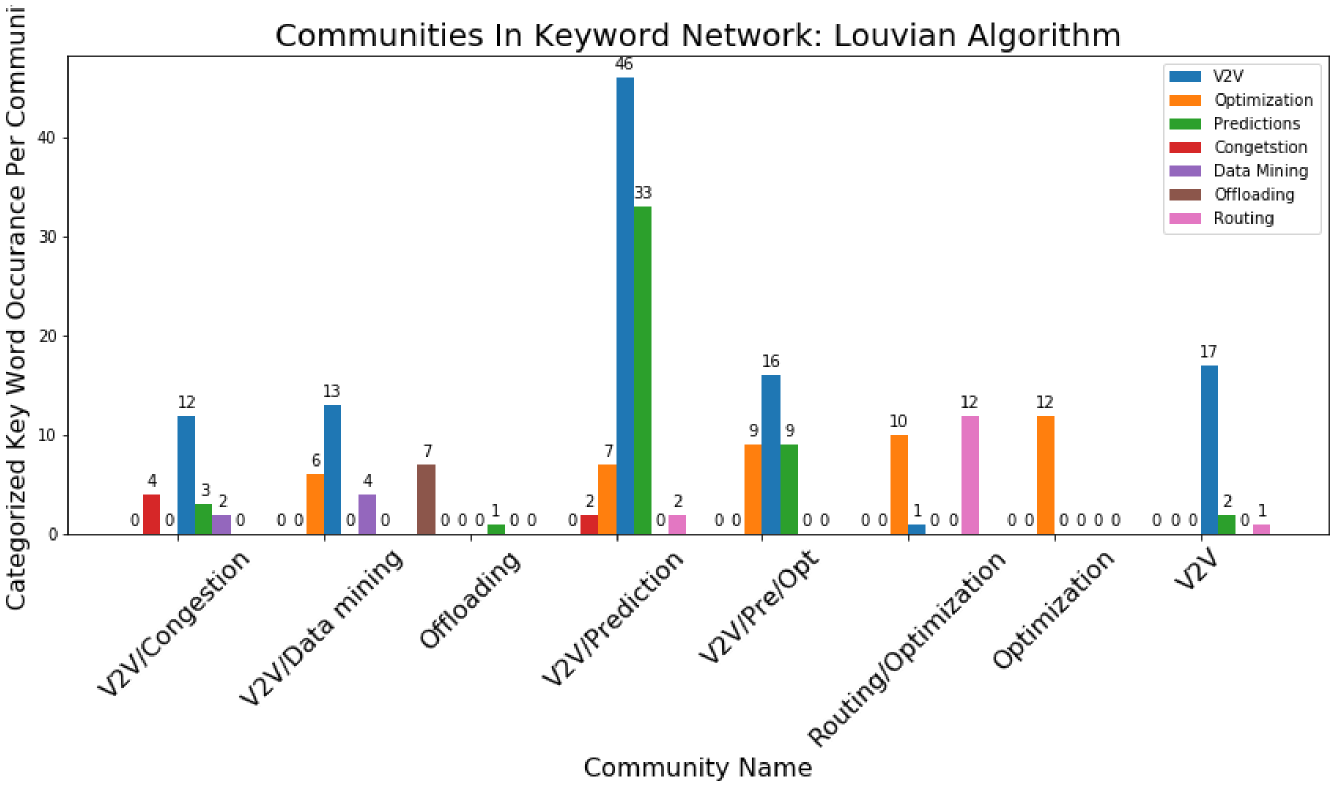


FIGURE 7. A bar chart used to describe the make up of each community in the Louvian community network, in terms of our categories. Labels for each community are based on the make up of the community.

the simulation is more complicated, it is developed for simulating networks, which is essential to conducting experiments for the IoV.

F. MATLAB

MATLAB is a matrix-based tool that is commonly used to visualize and analyze data. Visualization is accessible and convenient when using MATLAB's built-in graphics capabilities. Various data distributions, 2-D and 3-D plots, and animations are common methods used for analysis. MATLAB also has an extension, Simulink. Simulink takes input datasets from MATLAB and simulates it in real time. The results will then be output to MATLAB for visualization. MATLAB is organizational, where large-scale projects can be created and multiple people may work on it simultaneously. Adaptability is also available in MATLAB, as libraries in programming languages besides R (e.g., Python, Java, C++) can be imported, therefore allowing those languages to be used instead. Although the software is not directly developed as a simulator, it shows compatibility with its ability to conduct visualizations and its compatibility with ThingSpeak, an IoT analytics platform. The software runs on Windows, Linux, and Macintosh, and requires approximately 11 GB of memory.

G. SIMULATOR COMPARISON

NS is similar to OMNeT++, as they are both network simulators that can build simulations. Both NS and VEINS require other components to be downloaded to create a more realistic simulation. In addition, both support IEEE 802.11p, which is key to the Internet of Vehicles. SUMO does not compare to the other aforementioned test beds analyzed, however, it is a tool that can be used in unison with OMNeT++ and VEINS to simulate traffic. MATLAB and OSM are both programming tools, rather than simulators. MATLAB is primarily used for visualization of IoV datasets and shares insight with data analytics. OSM allows for both real-world data extraction as well as mapping.

Other less commonly used simulators in IoV papers include: Opportunistic Network Environment (ONE) simulator, SQLite, App Tune-up Kit application, VISSIM, MOVE, Abstract Notation Syntax One, MIXIM, VanetMobiSim, ndnSIM, and custom simulators using various tools developed by the researchers.

H. DATA SETS

Table 6 summarizes the different data sets in the papers analyzed in this survey. A third of the papers in the following table used OSM for information about road networks around the world [30], [52], [55], [68], [69], [82]. The majority of these papers used the OSM data with generated data to create synthetic data sets to test their proposed method, these are summarized in more detail in Section VI. However, some of these papers combined the OSM data with other collected data to use as a real data set [52], [69]. Other data sets focused on real-world data collected by various sources on

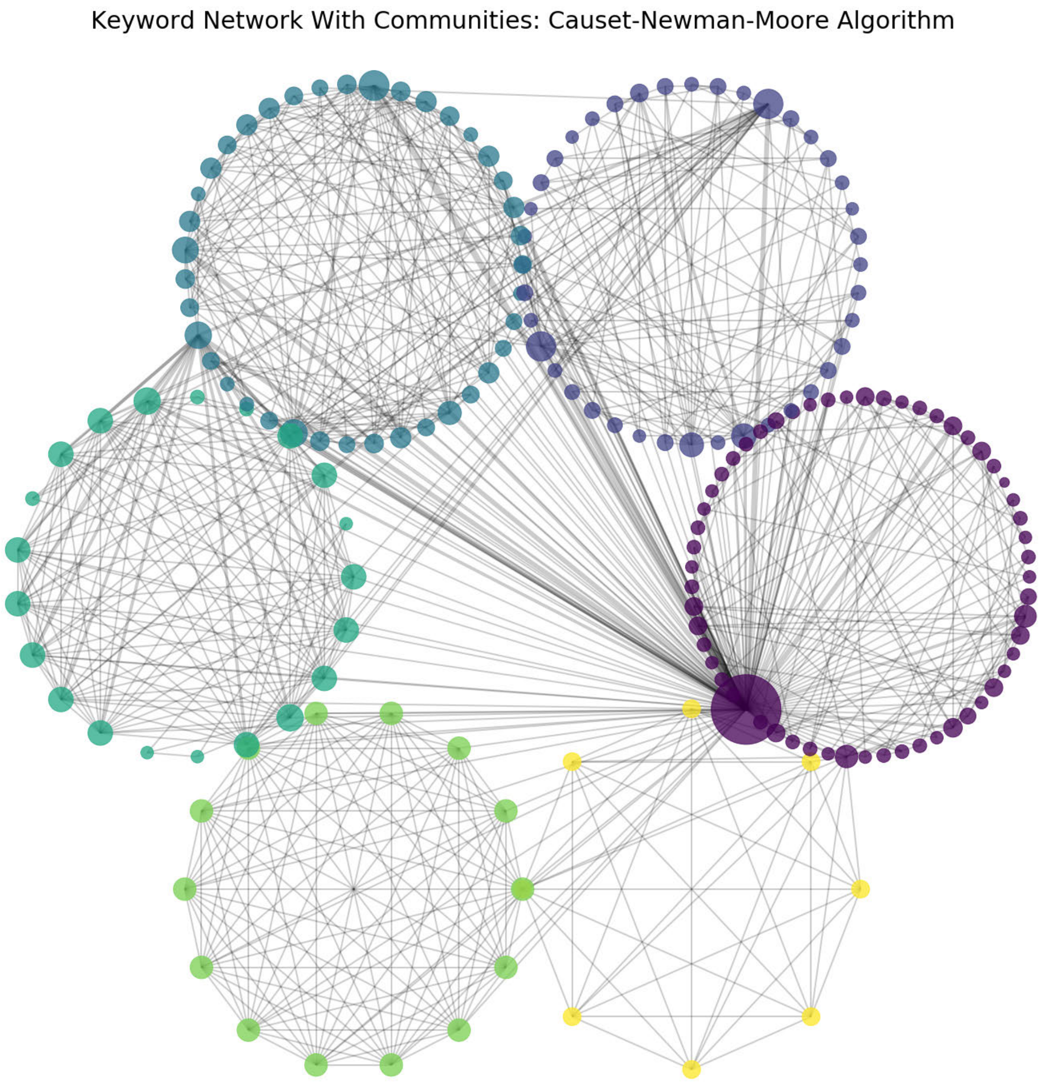


FIGURE 8. A network of communities created by applying the Clauset-Newman-Moore community detection algorithm to the network of keywords.Communities are placed around a circle with community one starting at the zero degree mark. Nodes are drawn in a circular pattern around the center of the community. Node sizes vary on the degree of the nodes. Edge width vary on the number of occurances the number of times a keyword pair is used.

cars, buses, and taxis and these are also classified as real data sets [26], [28], [36], [46], [62], [69], [70], [72], [73]. Two of the locations were used multiple times, but with different data sets, Beijing, China [36], [52], [62] and Cologne, Germany [28], [69]. These real data sets used in the papers are summarized in more detail. Data sets used that were not on road networks were also used as real data sets to test proposed methods and are also summarized below.

In Qiao *et al.*, two data sets of Beijing GPS taxi traces, one from 12 am to 1 am and the other from 7:30 am to 8:30 am, to analyze the temporal structure of VANETs [26]. The original data set included 2927 taxis with data, such as coordinates, collected every minute, all day on Jan. 5, 2009 [88]. Other data collected is not specified, but from the coordinates, speed and other variables can be calculated and used for testing IoV applications. Beijing is a good area for collecting real data due to the number of taxis used, however usage of data from only one day does not give multiple samples to testing variations in different parameters.

In Wang *et al.*, they use the Microsoft Research Asia T-Drive taxi data from the Beijing area to analyze traffic information collection and diffusion in IoVs [36]. The data set includes 10,357 taxis which creates about 15 million points and ranges a total trajectory distance of 9 million kilometers [89]. The data includes the ID, date, time, longitude, and latitude of each point and separates data into files that each contain the data from one taxi. The data set provides good real-world data on vehicle movement, especially in a more congested city such as Beijing, but does not provide any information such as vehicle speed, however, it could be calculated from the times and locations. The data set also covers an entire week which provides more variation and

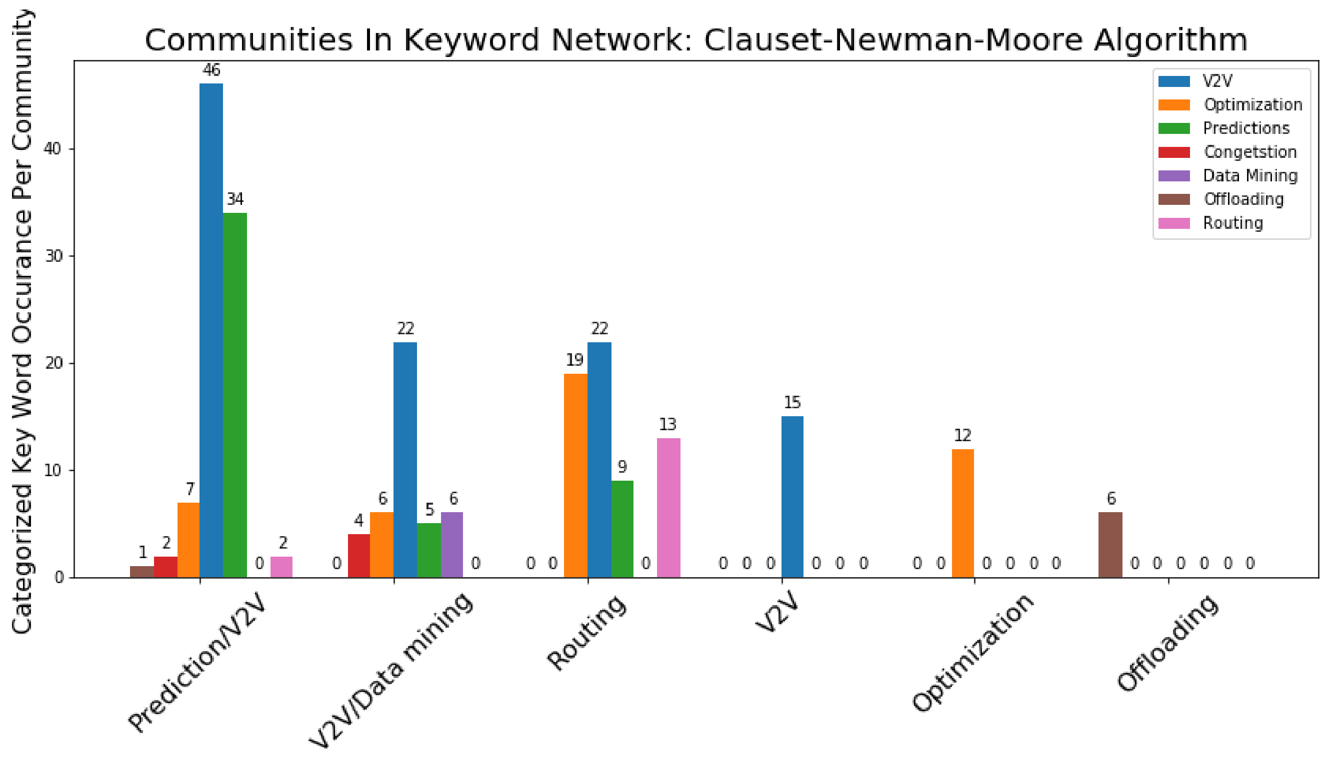


FIGURE 9. A bar chart used to describe the make up of each community in the Clauset-Newman-Moore community network, in terms of our categories. Labels for each community are based on the make up of the community.

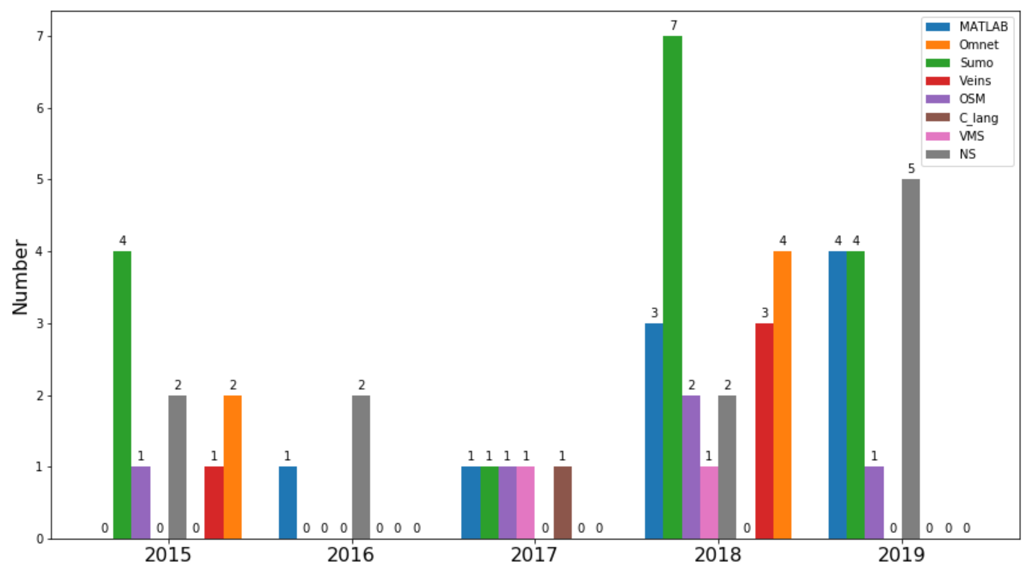


FIGURE 10. Frequency of the simulation tool 2015-2019.

multiple samples, but does not discuss how long the data was collected for during each day.

Zhang *et al.* use a quality of service data set to assess the quality of experience in mobile edge computing [80]. The paper uses the second version of two of the Quality of Web Service (QWS) data set which includes 2507 web services and their measurements [90]. The data set includes nine measurements of multiple benchmarks averaged over a six day period, as well as the service name and WDSL address. The benchmarks used are response time, availability, throughput,

TABLE 6. Summary of data sets.

Data Set Type	Data Set Used	Nodes	Year	Paper	Remarks
Synthetic	OSM [85]	NA	NA	[82]	
	OSM [85] & OBD2 DTC Sensory Data	NA	NA	[30]	
	OSM [85]	NA	NA	[55]	
	OSM [85]	NA	NA	[68]	
	Madrid Data Set	NA	NA	[46]	
Real	Beijing Taxi GPS Trace [88]	2927	2009	[26]	24 hour duration over one day
	Microsoft Research Asia T-Drive [89]	10357	2008	[36]	data collected over a week on a sample average interval of 177 seconds
	Quality of Web Service (QWS) [90]	2507	2007	[80]	
	Beijing Bus Routes		2013	[52]	every day of March 2013 from 6 AM to 10 PM in 20 second intervals
	California Transportation Agencies Performance Measurement System (CalTrans PeMS) [91]		2017/ 2018	[70]	real time data collected January to March 2017 and 2018
	Stanford Network Analysis Project (SNAP) Gowalla [92]	196591	2009-2010	[66]	user data collected from February 2009 to October 2010
	Beijing Taxi Trajectory Trace	12509	2012	[62]	data collected over 27 days from November 1 to 27, 2012
	KITTI Data Set		2012	[79]	2012
	Cologne Data Set	16658		[69]	23 hours
	TAPAS-Cologne Data Set	75600		[28]	6:00 AM to 8:00 AM
	England Data Set		2015	[72]	every Wednesday in May 2015
	Santander Data Set		2016	[73]	January to December 2016

successability, reliability, compliance, best practices, latency, and documentation. The data collected is good for analyzing communication and quality of service, but does not include mobility for the IoV, which is an essential component and affects the quality of IoV services.

OpenStreetMap road network topology and GPS data from three bus routes in Beijing China are used to evaluate Zhang *et al.*'s proposed method as well as a synthetic data set [52]. The GPS data is gathered from the three bus routes: No. 939, No. 944, and No. 983; collected between 6 AM and 10 PM on March 13 with reports every 20 seconds [52]. Each report contains a timestamp, ID, line number, current location, speed, and other information [52]. The data set used seems to have enough information to provide a thorough simulation to test an IoV application and includes important parameters that most other data sets include.

Ge *et al.* use California Transportation Agencies Performance Measurement System (CalTrans PeMS) data sets PEMSD7, for district 7, and PEMSD4, for district 4, of Los Angeles [70]. They select 204 during Jan. to Mar. 2018 and 325 sensors during Jan. to Mar. 2017 from PEMSD7 and PEMSD4, respectively, and extract speed information to test their traffic speed prediction [91]. In both data sets, the periods of Jan. 1 to Mar. 14 are used for the training sets and Mar. 15 to Mar. 31 are used for the testing set [70]. The CalTrans PeMS data is collected from real-time sensors located around the freeway network in California and can be accessed through CalTrans. This data set includes extensive real data that can be used for testing many applications in the IoV due to the number of sensors and extensiveness of the collection area. However, the data collected is not specified which is important for the use of the data set.

The Stanford Network Analysis Project (SNAP) Gowalla data set is used in Elbery *et al.* to train and test their proposed carpool prediction method [66]. The training data is from Feb. 2009 to Aug. 31, 2010 and future data to test the method is from Aug 31, 2010 to Oct. 2010. SNAP is a collection of large network data sets including social, autonomous systems, temporal, and various other networks. Gowalla is a location-based social networking website, and the data set consists of 196591 nodes and 950327 edges and includes the user ID, check-in time, latitude, longitude, and location ID [92]. This data set provides data that works with analyzing connectivity between users which works for carpool prediction, but does not include any vehicular data for testing or simulating an IoV and also does not specify the data collected from users. Using the majority of the first part of a data set for training and less of the second half the data set for testing predictions works well for testing a prediction application.

In Cui *et al.*, another Beijing taxi data set is used, due to their similarity to autonomous vehicles, to analyze their proposed methods of using network calculus and a Kuhn-Munkres algorithm [62]. The data set contains trajectories of 12509 taxis in Beijing from Nov. 1 to 27, 2012 which resulted in 785.4 million entries each containing the time, location, ID, and taxi state [62]. This data set takes into account the taxi state (i.e. vacant, loaded, parking, not in service, or other) which allows for altering the data set to fit the needs of the application. However, it does not discuss how many vehicles had data collected on, although it can be assumed that since it is Beijing taxi data, there is an extensive number of vehicles the data was collected on.

In Chen *et al.*, a synthetic data set based in Madrid was used go approach maximizing message dissemination [46]. The data set is a generated realistic mobility trace in Madrid; it has a traffic flow rate of 3,600 vehicles per hour with a speed limit of 120 kilometers per hour. Three lanes were used in the simulation, which allows for an evaluation of how the scheme performs in wide, more populated areas. By having a large area to test their algorithm in, the authors are able to analyze how popular broadcasting are received among the vehicles. A limitation of this paper is that more rural and sparsely populated areas are not included in the evaluation, so the algorithm analysis is not comprehensive.

Gündüz and Acarman [79] use the KITTI data set to approach improving real-time object detection by vehicles. The KITTI data set is composed of 7,481 training images and 7,518 test images with 80,256 labeled objects. The data is collected from Karlsruhe, a mid-sized city; both rural areas and highways are located in the city, which provides for an adaptive use of the data set. The authors select images that contain a maximum of 15 cars and 30 pedestrians per image so they may evaluate how vehicles using their algorithm will respond to different densities of objects in their paths.

In Yu *et al.*, [69] traffic data collected in Cologne is used to develop and test an algorithm to estimate the speed of other vehicles. The data set used in the work implement 16,658 sparsely connected nodes as well as 37,034 edges. It stems from a larger set of information containing 3.54 billion GPS and speed records of over 700,000 individual trips taken over 23 hours. Speeds are randomly selected from the speed records for the vehicles in their simulation; this enables the authors to conduct experiments that rigorously tests the ability of their algorithm to estimate traffic speed.

In Tao *et al.* [72], a real data set from England was used to analyze how an algorithm centered around predicting short-term traffic flow would respond in a vehicular network. The data set contains traffic flow data with intervals of 15 minutes; the times the data was collected were all Wednesdays in May of 2015. Seven different roads were selected to extract the data from. This enables the authors to test their scheme in places with varying amounts of traffic flow to obtain a more comprehensive analysis of their scheme. A limitation of this approach is the times data was collected. Specifically choosing Wednesdays in May with unstated time slots does not provide an accurate representation of how traffic will always flow. More defined and consistent data collection times need to be established on different days to gain a more inclusive and comprehensive data set to experiment on.

Loulloudes *et al.* [28] use the TAPAS-Cologne dataset to describe patterns and characteristics of how vehicles behave in the network. The portion of the dataset used is during a traffic rush period from 6:00 AM to 8:00 AM; the area in question is 33 × 35 kilometers with 134,645 roads and 42,148 intersections. 75,600 unique vehicles were used to analyze vehicular behavior; by using such a large amount of vehicles, a more accurate representation of the behavior in the network can be demonstrated. To properly evaluate all vehicular behavior in a network, more time slots with less traffic flow should also be analyzed to identify behavior of vehicles in a more sparse area.

Kim *et al.* [73] use a data set from Santander, Spain to support their claim that embedded topology improves the process of vehicles learning traffic features. The data set comes from case studies in the SETA EU project [93] and contains real traffic speed data. Measurements for the data set were taken every 15 minutes for unstated times throughout the year 2016. The first nine months of data was used as a training set and the following three months were used to evaluate their scheme. This great amount of data allows for a deeper and more realistic analysis of how vehicles will benefit from embedded topology. A limitation of this paper is that the time slots are unstated which does not provide an idea of what the sample of vehicular data is representing.

TABLE 7. Summary of experimentation methods.

Experimentation	Application	Paper
Vehicle Simulation	V2V Interaction	[29], [82], [39], [58], [25], [27], [41], [40], [35], [37], [42], [30], [57], [94], [47], [50], [43], [49], [45], [34], [31], [44], [48], [33]
	Optimization	[58], [25], [57], [60], [24], [51], [53], [56], [55], [54], [64]
	Predictions	[67], [71], [65], [25]
	Routing	[61], [76], [63], [71]
	Congestion Control	[75], [76], [77]
Dataset	V2V Interaction	[36], [28], [26], [46], [50]
	Predictions	[66], [69], [70], [72], [73]
	Offloading	[80], [62]
	Data Mining	[79]
Mapping	V2V Interaction	[32]
	Predictions	[68]
	Optimization	[68]
	Congestion Control	[74]
	Data Mining	[78]
Routing/Trajectory	Optimization	[59], [52]
Theoretical	V2V Interaction	[38], [46]

I. SIMULATOR-APPLICATION RELATIONSHIP

Table 7 refers to a table relating IoV-related applications discussed in the paper to simulators used by the authors. For V2V interactions, MATLAB, SUMO, and NS were the most popular simulation tools; they are also among the most commonly used simulators discussed in this paper. This can be attributed to the different methods of experimentation in the category of V2V interaction, which would call for different operation tools for different experimentation (e.g., dataset versus vehicular simulation). For optimization in the IoV, OMNeT++ was the most commonly used tool. Simulations centered around routing were performed in Network Simulator. Despite efforts to find relevant papers from 2016 that incorporated simulation tools, there was a noticed trend that

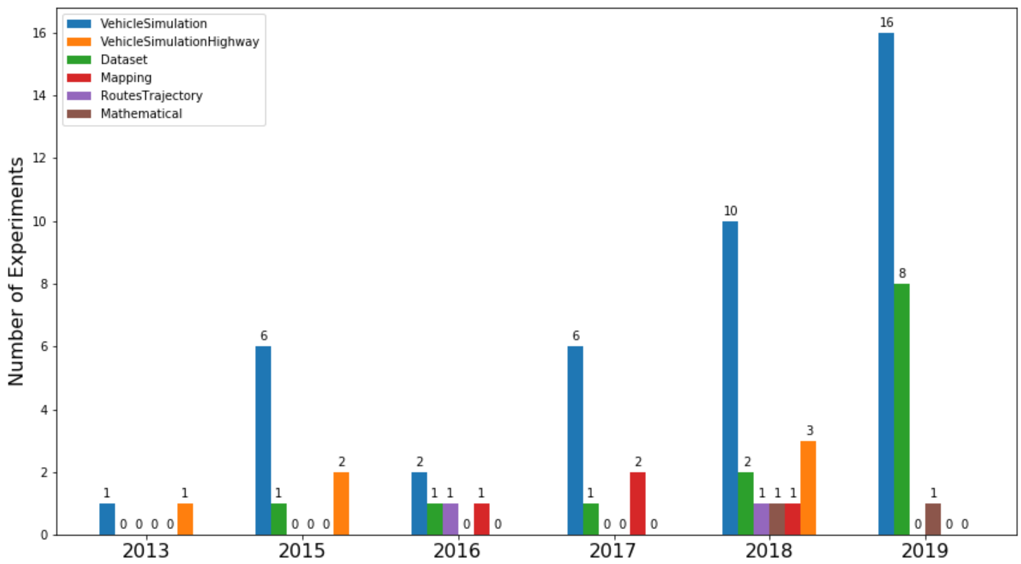


FIGURE 11. **Different IoV experimentation 2013-2019.**

focused on data-centric evaluations rather than conducting experiments using simulations.

J. EXPERIMENTATION METHODS

Most papers opt to perform vehicular simulations (54% of papers in this survey), with many authors opting for a multi-lane street or highway as the simulation's map layout. These simulations typically constituted of three different scenarios: sparse traffic, medium-level traffic, and dense traffic. This way, the application of their algorithm or methodology can be evaluated at different levels of traffic density to provide a more well-rounded analysis of its performance.

Datasets extracted from transportation sources (e.g., CalTrans) are the next most-common form of experimentation. This can be attributed to the availability of large sets of data points. Following the applications in the proposed taxonomy, the information from the datasets (which may include: vehicle paths, speeds, number of vehicles, etc.) is visualized and analyzed.

Less commonly used experimentation methods are: using common city routes or maps, and synthetic datasets.

VI. DISCUSSION ON SIMULATION STRENGTHS AND LIMITATIONS

In the category of congestion control, none of the papers use datasets to test their proposed method. This is linked to the fact that the traffic and congestion have already happened in a given dataset and therefore, it would be inefficient to control the traffic. Therefore vehicle simulation is neccessary for these papers. All but one of the congestion control papers use NS2, and the one that doesn't uses VISSIM. Three papers that use NS2 also use SUMO to simulate the vehicles mobility part since NS2 simulates the network part. The last paper uses MOVE and Matlab as well as NS2 and SUMO. Some of the congestion control papers vary cars, speeds, and densities, but none of them include RSUs or vary the duration of the simulation which is important to thoroughly test the method. The congestion control papers are able to reduce network congestion, improve end-to-end delays and packet ratio, however they do not consider how their method works in a lower density network, or it does not work as effectively in a less dense network. It is relevant to find a method of congestion control that works in all densities of traffic.

For the offloading papers, two papers use datasets. One uses Matlab, and the other paper uses NS3 with Matlab. Only one of the papers takes into account varying the number of cars, but neither test multiple densities or durations. Since the amount of processing that takes place can vary based on the number of cars and the time, it is important for these experiments to include consider environments. It is also good to use datasets for testing how the method works in the real world, but variation in the datasets used is also important. One paper demonstrated an improved computation time, and the other reduces the travel time and energy consumption, and the last paper increases efficiency, but each one could be improved in efficiency, applied to different traffic densities, and work with a communication standard.

Of the predictions papers, five of them use datasets, three of them use OMNeT++ and two use SUMO, with one using the VEINS combination. Some of the papers also use OSM, PyTorch, NS3, and Sklearn. Two of the papers vary the number of cars, one paper varies the speeds, and one paper varies the density. Since predictions need to take into account various situations, the experiments need to include more variations in duration and scenarios, among others. The papers that use datasets work well since predictions should be tested in a real world environment, however they should also be evaluated for varying traffic conditions to make sure the

TABLE 8. Overview of simulation parameters for predictions.

Paper	Vehicles	RSU	Speeds	Density	Duration	Scenario	Lanes	Simulations
[65]	✓		✓	✓		Highway	3	✓
[68]	✓		✓	✓	4000 s.	City	2+	✓
[25]			✓			Highway	3	✓

TABLE 9. Overview of simulation parameters for V2V interactions.

Paper	Vehicles	RSU	Speeds	Density	Duration	Scenario	Lanes	Simulations
[27]	✓			✓	2000 s.	City	2	
[57]	✓		✓			Highway		
[58]	✓			✓		Highway	3	
[25]			✓			Highway	3	✓
[33]					50 s.	City		
[39]	✓			✓		Highway		✓
[40]			✓		500 s.	City	4	
[82]					33 s.	Highway		✓
[29]	✓		✓	✓	1000 s.	Highway	4	
[28]	✓		✓	✓	7200 s.	City/Highway	2+	✓
[37]			✓		500 s.			
[35]	✓		✓	✓	500 s.		4	
[48]		✓	✓		1800 s.			
[46]	✓	✓	✓	✓			4	
[31]					200 s.	Intersection	3	
[45]	✓	✓		✓		Intersection	3	
[43]	✓		✓	✓	500 s.	Highway	4+	✓
[47]			✓		180 s.	Highway	4	
[50]	✓		✓	✓	120 s.	Highway	12	✓

method works in all scenarios. The current prediction papers perform well in RMSE, transmission ratio, end-to-end delay, route lifetime, number of request ratios, producing a complete speed map faster, more accurate, and work better in higher density scenarios. However, papers can still be improved under overhead, computational complexity, and different road scenarios and structures. Table 8 summarizes the simulations of the different predictions papers.

Table 9 is a breakdown of each parameter for the simulation (number of vehicles, number of RSU, vehicle speeds, density of vehicles, duration of simulations, scenario, number of lanes, and number of simulations ran). Each column in Table 9 represents whether a parameter was varied during the simulation. For example, the first row has a check mark under the column referring to vehicles, which means that simulation used different number of vehicles for each simulation. The more thorough simulations have more check marks and other information available.

The three most thorough simulations were discussed in [47], [50], and [28]. All three of these simulations were ran for at least two minutes and thoroughly validated using multiple simulations under different circumstances. For example, [28] ran their simulation for 7200 seconds multiple times while varying the number of vehicles, vehicle speeds, density of vehicles, and used multiple lanes.

A common theme among the papers in Table 9 is the constant number of RSUs throughout simulations. Another common theme is short simulations. Short simulations may not yield enough data to thoroughly test their approaches. A final common theme throughout the simulations ran is not running multiple simulations for each test. Running multiple simulations gives a well-rounded set of scenarios to test and lead to better results.

For the optimization papers, many of the authors performed simulations with multiple factors taken into account. The simulation conducted by Wang *et al.* is thoroughly conducted. They ran 100 simulations with a duration of 168 hours. The speeds of the vehicles were varied from 20 km/hr to 60 km/hr, with a communication range of 40 km. They included four performance metrics: average delivery

TABLE 10. Overview of simulation parameters for resource optimization.

Paper	Vehicles	RSU	Speeds	Density	Duration	Scenario	Lanes	Simulations
[53]	✓		✓	✓		Highway		
[54]	✓			✓	80 s.			
[59]	✓	✓	✓	✓	7200 s.	City	2+	✓
[52]	✓			✓	5400 s.	city	2+	✓
[57]	✓		✓	✓				
[58]				✓		Highway	3	✓
[55]		✓			24 h.	City	2+	✓
[56]			✓			City	2+	
[51]	✓	✓	✓	✓				✓
[60]			✓		168 h.	City		✓

ratio, average delivery delay, average communication cost, and access ratio, which enabled a well-rounded evaluation. Table 10 summarizes these parameters for the resource optimization papers.

There are still gaps in the simulations. A commonality between simulations is that the number of RSUs are only varied in papers whose aim is to optimize RSU placement. Other papers not optimizing RSU placement should also include tests where there are different numbers of RSUs for varying numbers of vehicles. Furthermore, the duration of the simulations need to be extended to gather larger amounts of data to base their results off of [30].

VII. CONCLUSION

The concept of the Internet of Vehicles (IoV) is becoming an increasingly relevant topic as it allows for more efficient traffic networks and better transportation. Due to its popularity and the importance of many of its features, there is a broad library of research on the IoV. This paper summarizes the research projects that utilize network or graph theory, and presents its different applications to the IoV, as well as a discussion on the strengths and limitations and the simulations conducted. This paper paves the path for understanding the existing research on the IoV. This paper also summarizes the different simulation tools most relevant to IoV research and how they were used in the papers to provide an analysis on which simulators are most appropriate for various applications. Finally, the paper presents summary of findings and discusses their key contributions to the body of research. The shortcomings on the conducted simulations are also discussed, which acts as the baseline for future research to be performed on.

We identified the relationships among the keywords and the categories by applying two community detection methods, Louvian and Clauset-Newman-Moore. These findings visualize the trends in the literature and can guide researchers on the existing challenges. Such an approach to surveying the literature is a novel method, revealing many patterns that may be hidden to the naked eye.

REFERENCES

[1] S. Tayeb, M. Pirouz, and S. Latifi, "A Raspberry-Pi prototype of smart transportation," in *Proc. 25th Int. Conf. Syst. Eng. (ICSEng)*, Aug. 2017, pp. 176–182.

[2] S. Tayeb, M. Pirouz, G. Esguerra, K. Ghobadi, J. Huang, R. Hill, D. Lawson, S. Li, T. Zhan, J. Zhan, and S. Latifi, "Securing the positioning signals of autonomous vehicles," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2017, pp. 4522–4528.

[3] S. Tayeb, S. Latifi, and Y. Kim, "A survey on IoT communication and computation frameworks: An industrial perspective," in *Proc. IEEE 7th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2017, pp. 1–6.

[4] H. Wang, J. Hou, and N. Chen, "A survey of vehicle re-identification based on deep learning," *IEEE Access*, vol. 7, pp. 172443–172469, 2019.

[5] J. Zhu, H. Zeng, Y. Du, Z. Lei, L. Zheng, and C. Cai, "Joint feature and similarity deep learning for vehicle re-identification," *IEEE Access*, vol. 6, pp. 43724–43731, 2018.

[6] C. Wang, X. Gaimu, C. Li, H. Zou, and W. Wang, "Smart mobile crowdsensing with urban vehicles: A deep reinforcement learning perspective," *IEEE Access*, vol. 7, pp. 37334–37341, 2019.

[7] G. Loukas, T. Vuong, R. Heartfield, G. Sakellari, Y. Yoon, and D. Gan, "Cloud-based cyber-physical intrusion detection for vehicles using deep learning," *IEEE Access*, vol. 6, pp. 3491–3508, 2018.

[8] M. H. Basiri, M. Pirani, N. L. Azad, and S. Fischmeister, "Security of vehicle platooning: A game-theoretic approach," *IEEE Access*, vol. 7, pp. 185565–185579, 2019.

[9] O. Kaiwartya, A. H. Abdullah, Y. Cao, A. Altameem, M. Prasad, C.-T. Lin, and X. Liu, "Internet of vehicles: Motivation, layered architecture, network model, challenges, and future aspects," *IEEE Access*, vol. 4, pp. 5356–5373, 2016.

[10] M. E. J. Newman, *Networks*. London, U.K.: Oxford Univ. Press, 2018.

[11] X. Wang, Z. Ning, M. Zhou, X. Hu, L. Wang, Y. Zhang, F. R. Yu, and B. Hu, "Privacy-preserving content dissemination for vehicular social networks: Challenges and solutions," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1314–1345, 2nd Quart., 2019.

[12] Z. Ning, X. Hu, Z. Chen, M. Zhou, B. Hu, J. Cheng, and M. S. Obaidat, "A cooperative quality-aware service access system for social Internet of vehicles," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2506–2517, Aug. 2018.

[13] G. Xiong, F. Zhu, X. Liu, X. Dong, W. Huang, S. Chen, and K. Zhao, "Cyber-physical-social system in intelligent transportation," *IEEE/CAA J. Autom. Sinica*, vol. 2, no. 3, pp. 320–333, Jul. 2015.

[14] W. Xu, H. Zhou, N. Cheng, F. Lyu, W. Shi, J. Chen, and X. Shen, "Internet of vehicles in big data era," *IEEE/CAA J. Autom. Sinica*, vol. 5, no. 1, pp. 19–35, Jan. 2018.

[15] X. Wang, C. Wang, J. Zhang, M. Zhou, and C. Jiang, "Improved rule installation for real-time query service in software-defined Internet of vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 2, pp. 225–235, Feb. 2017.

[16] Y. Ma, Z. Wang, H. Yang, and L. Yang, "Artificial intelligence applications in the development of autonomous vehicles: A survey," *IEEE/CAA J. Autom. Sinica*, vol. 7, no. 2, pp. 315–329, Mar. 2020.

[17] P. M. Kebria, A. Khosravi, S. M. Salaken, and S. Nahavandi, "Deep imitation learning for autonomous vehicles based on convolutional neural networks," *IEEE/CAA J. Autom. Sinica*, vol. 7, no. 1, pp. 82–95, Jan. 2020.
[18] J. Cheng, J. Cheng, M. Zhou, F. Liu, S. Gao, and C. Liu, "Routing in Internet of vehicles: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 5, pp. 2339–2352, Oct. 2015.
[19] H. Guo, D. Cao, H. Chen, C. Lv, H. Wang, and S. Yang, "Vehicle dynamic state estimation: State of the art schemes and perspectives," *IEEE/CAA J. Autom. Sinica*, vol. 5, no. 2, pp. 418–431, Mar. 2018.
[20] J. Cheng, H. Mi, Z. Huang, S. Gao, D. Zang, and C. Liu, "Connectivity modeling and analysis for Internet of vehicles in urban road scene," *IEEE Access*, vol. 6, pp. 2692–2702, 2018.
[21] J. Cheng, H. Yan, A. Zhou, C. Liu, D. Cheng, S. Gao, D. Zang, and D. Cheng, "Location prediction model based on the Internet of vehicles for assistance to medical vehicles," *IEEE Access*, vol. 8, pp. 10754–10767, 2020.
[22] F. Trueblood, S. Gill, R. Wong, S. Tayeb, and M. Pirouz, "A data-centric approach to taming the message dissemination in the Internet of vehicles," in *Proc. 10th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2020, pp. 207–214.
[23] L. Qi, M. Zhou, and W. Luan, "A dynamic road incident information delivery strategy to reduce urban traffic congestion," *IEEE/CAA J. Autom. Sinica*, vol. 5, no. 5, pp. 934–945, Sep. 2018.
[24] R. Kochhar and H. Mandoria, "Performance study of VANET using ant based routing algorithms," in *Proc. 2nd Int. Conf. Comput. Sustain. Global Develop. (INDIACom)*, 2015, pp. 1803–1806.
[25] M. H. Eiza and Q. Ni, "An evolving graph-based reliable routing scheme for VANETs," *IEEE Trans. Veh. Technol.*, vol. 62, no. 4, pp. 1493–1504, May 2013.
[26] L. Qiao, Y. Shi, and S. Chen, "An empirical study on the temporal structural characteristics of VANETs on a taxi GPS dataset," *IEEE Access*, vol. 5, pp. 722–731, 2017.
[27] J. Costa, D. Rosario, A. M. de Souza, L. A. Villas, and E. Cerqueira, "Data dissemination based on complex networks' metrics for distributed traffic management systems," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Jun. 2018, pp. 01062–01067.
[28] N. Loulloudes, G. Pallis, and M. D. Dikaiakos, "The dynamics of vehicular networks in large-scale urban environments," in *Proc. IEEE Conf. Collaboration Internet Comput. (CIC)*, Oct. 2015, pp. 192–199.
[29] M. N. Avcil and M. Soyturk, "ReSCUE: Relatively stable clustering for unbiased environments in VANETs," in *Proc. Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Aug. 2015, pp. 1049–1055.
[30] K. M. Alam, M. Saini, and A. E. Saddik, "Toward social Internet of vehicles: Concept, architecture, and applications," *IEEE Access*, vol. 3, pp. 343–357, 2015.
[31] M. Renukadevi, C. Balasubramanian, and M. D. A. Nithya, "A reliable application level broadcasting protocol for VANET," in *Proc. Int. Conf. Soft-Comput. Netw. Secur. (ICSNS)*, Feb. 2015, pp. 1–7.
[32] Y. Huang, M. Chen, Z. Cai, X. Guan, T. Ohtsuki, and Y. Zhang, "Graph theory based capacity analysis for vehicular ad hoc networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–5.
[33] V. Kumar and A. S. Baghel, "Comparison between AODV protocol and LSGR protocol in VANET," in *Proc. 3rd Int. Conf. Comput. Sustain. Global Develop. (INDIACom)*, 2016, pp. 3026–3030.
[34] S. Tan, X. Li, and Q. Dong, "A trust management system for securing data plane of ad-hoc networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 9, pp. 7579–7592, Sep. 2016.
[35] H. Feng, C. Li, and Y. Xu, "Invulnerability analysis of vehicular ad hoc networks based on temporal networks," in *Proc. 2nd IEEE Int. Conf. Comput. Commun. (ICCC)*, Oct. 2016, pp. 2198–2202.
[36] J. Wang, C. Jiang, Z. Han, Y. Ren, and L. Hanzo, "Internet of vehicles: Sensing-aided transportation information collection and diffusion," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 3813–3825, May 2018.
[37] H. Feng, J. Zhang, J. Wang, and Y. Xu, "Time evolution of the importance of nodes in VANET based on temporal networks," in *Proc. 3rd IEEE Int. Conf. Comput. Commun. (ICCC)*, Dec. 2017, pp. 1210–1214.
[38] L. Nie, Y. Li, and X. Kong, "Spatio-temporal network traffic estimation and anomaly detection based on convolutional neural network in vehicular ad-hoc networks," *IEEE Access*, vol. 6, pp. 40168–40176, 2018.
[39] J.-J. Huang and Y.-T. Tseng, "The steady-state distribution of rehealing delay in an intermittently connected highway VANET," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 10010–10021, Oct. 2018.
[40] H. Feng, J. Zhang, J. Wang, and Y. Xu, "Dynamic analysis of VANET using temporal reachability graph," in *Proc. IEEE 17th Int. Conf. Commun. Technol. (ICCT)*, Oct. 2017, pp. 783–787.
[41] X. Li, T. Song, Y. Zhang, G. Chen, and J. Hu, "A hybrid cooperative spectrum sensing scheme based on spatial-temporal correlation for CR-VANET," in *Proc. IEEE 87th Veh. Technol. Conf. (VTC Spring)*, Jun. 2018, pp. 1–6.
[42] X. Li, T. Song, Y. Zhang, G. Chen, and J. Hu, "Hybrid cooperative spectrum sensing scheme based on spatial-temporal correlation in cognitive radio enabled VANET," *IET Commun.*, vol. 13, no. 1, pp. 36–44, Jan. 2019.
[43] F. Abbas, P. Fan, and Z. Khan, "A novel low-latency V2V resource allocation scheme based on cellular V2X communications," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 6, pp. 2185–2197, Jun. 2019.
[44] G. Luo, Q. Yuan, H. Zhou, N. Cheng, Z. Liu, F. Yang, and X. S. Shen, "Cooperative vehicular content distribution in edge computing assisted 5G-VANET," *China Commun.*, vol. 15, no. 7, pp. 1–17, Jul. 2018.
[45] G. Luo, H. Zhou, N. Cheng, Q. Yuan, J. Li, F. Yang, and X. S. Shen, "Software defined cooperative data sharing in edge computing assisted 5G-VANET," *IEEE Trans. Mobile Comput.*, early access, Nov. 12, 2019, doi: 10.1109/TMC.2019.2953163.
[46] C. Chen, J. Hu, T. Qiu, M. Atiquzzaman, and Z. Ren, "CVCG: Cooperative V2V-aided transmission scheme based on coalitional game for popular content distribution in vehicular ad-hoc networks," *IEEE Trans. Mobile Comput.*, vol. 18, no. 12, pp. 2811–2828, Dec. 2019.
[47] G. Manisha, G. S. R. E. Selvan, and M. P. Ramkumar, "Pending interest lifetime mechanism for vehicular named data networks," in *Proc. Int. Conf. Vis. Towards Emerg. Trends Commun. Netw. (ViTECoN)*, Mar. 2019, pp. 1–6.
[48] M. Wazid, P. Bagga, A. K. Das, S. Shetty, J. J. P. C. Rodrigues, and Y. Park, "AKM-IoV: Authenticated key management protocol in fog computing-based Internet of vehicles deployment," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8804–8817, Oct. 2019.
[49] A. Yang, J. Weng, N. Cheng, J. Ni, X. Lin, and X. Shen, "DeQoS attack: Degrading quality of service in VANETs and its mitigation," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4834–4845, May 2019.
[50] B. Toghi, M. Saifuddin, M. O. Mughal, and Y. P. Fallah, "Spatio-temporal dynamics of cellular V2X communication in dense vehicular networks," in *Proc. IEEE 2nd Connected Automat. Vehicles Symp. (CAVS)*, Sep. 2019, pp. 1–5.
[51] S. Mehar, S. M. Senouci, A. Kies, and M. M. Zoulikha, "An optimized roadside units (RSU) placement for delay-sensitive applications in vehicular networks," in *Proc. 12th Annu. IEEE Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2015, pp. 121–127.
[52] F. Zhang, B. Jin, Z. Wang, H. Liu, J. Hu, and L. Zhang, "On geocasting over urban bus-based networks by mining trajectories," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 6, pp. 1734–1747, Jun. 2016.
[53] S. Peng, R. Chai, Q. Chen, and Y. Qin, "Minimum end-to-end transmission delay based routing algorithm for VANETs," in *Proc. 9th Int. Conf. Adv. Infocomm Technol. (ICAIT)*, Nov. 2017, pp. 176–181.
[54] V. Saritha, P. V. Krishna, S. Misra, and M. S. Obaidat, "Learning automata based optimized multipath routingusing leapfrog algorithm for VANETs," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–5.
[55] D. Kim, Y. Velasco, W. Wang, R. Uma, R. Hussain, and S. Lee, "A new comprehensive RSU installation strategy for cost-efficient VANET deployment," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 4200–4211, May 2017.
[56] R. B. Koti and M. S. Kakkasageri, "Dynamic topology control in multiple clustered vehicular ad hoc networks," in *Proc. Int. Conf. Signal Process., Commun., Power Embedded Syst. (SCOPES)*, Oct. 2016, pp. 1371–1375.
[57] R. Chai, Y. Qin, S. Peng, and Q. Chen, "Transmission performance evaluation and optimal selection of relay vehicles in VANETs," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2017, pp. 1–6.
[58] R. Huang, J. Wu, C. Long, Y. Zhu, B. Li, and Y.-B. Lin, "Mitigate the obstructing effect of vehicles on the propagation of VANETs safety-related information," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 5558–5569, Jul. 2018.
[59] H. Bao, Q. Liu, C. Huang, and X. Jia, "Minimal road-side unit placement for delay-bounded applications in bus ad-hoc networks," in *Proc. IEEE 36th Int. Perform. Comput. Commun. Conf. (IPCCC)*, Dec. 2017, pp. 1–7.
[60] X. Wang, Z. Ning, X. Hu, L. Wang, B. Hu, J. Cheng, and V. C. M. Leung, "Optimizing content dissemination for real-time traffic management in large-scale Internet of vehicle systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1093–1105, Feb. 2019.

[61] A. A. Kashani, M. Ghanbari, and A. M. Rahmani, "Improving the performance of opportunistic routing protocol using the evidence theory for VANETs in highways," *IET Commun.*, vol. 13, no. 20, pp. 3360–3368, Dec. 2019.

[62] Q. Cui, Y. Wang, K.-C. Chen, W. Ni, I.-C. Lin, X. Tao, and P. Zhang, "Big data analytics and network calculus enabling intelligent management of autonomous vehicles in a smart city," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2021–2034, Apr. 2019.

[63] L. Zhao, Z. Li, J. Li, A. Al-Dubai, G. Min, and A. Y. Zomaya, "A temporal-information-based adaptive routing algorithm for software defined vehicular networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[64] R. Li, P. Zhu, and L. Jin, "Channel allocation scheme based on greedy algorithm in cognitive vehicular networks," in *Proc. IEEE 3rd Inf. Technol., Netw., Electron. Automat. Control Conf. (ITNEC)*, Mar. 2019, pp. 803–807.

[65] M. H. Eiza, T. Owens, and Q. Ni, "Secure and robust multi-constrained QoS aware routing algorithm for VANETs," *IEEE Trans. Dependable Secure Comput.*, vol. 13, no. 1, pp. 32–45, Jan. 2016.

[66] A. Elbery, M. El Nainay, and H. Rakha, "Proactive and reactive carpooling recommendation system based on spatiotemporal and geosocial data," in *Proc. IEEE 12th Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, Oct. 2016, pp. 1–8.

[67] K. Pattanayak, A. Chatterjee, M. Dzaferagic, S. S. Das, and N. Marchetti, "A functional complexity framework for dynamic resource allocation in VANETs," in *Proc. 14th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Jun. 2018, pp. 458–463.

[68] G. Sun, Y. Zhang, D. Liao, H. Yu, X. Du, and M. Guizani, "Bus-trajectory-based street-centric routing for message delivery in urban vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7550–7563, Aug. 2018.

[69] J. J. Q. Yu and J. Gu, "Real-time traffic speed estimation with graph convolutional generative autoencoder," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 10, pp. 3940–3951, Oct. 2019.

[70] L. Ge, H. Li, J. Liu, and A. Zhou, "Temporal graph convolutional networks for traffic speed prediction considering external factors," in *Proc. 20th IEEE Int. Conf. Mobile Data Manage. (MDM)*, Jun. 2019, pp. 234–242.

[71] Z. Khan, P. Fan, S. Fang, and F. Abbas, "An unsupervised cluster-based VANET-oriented evolving graph (CVoEG) model and associated reliable routing scheme," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 10, pp. 3844–3859, Oct. 2019.

[72] Y. Tao, P. Sun, and A. Boukerche, "A novel travel-delay aware short-term vehicular traffic flow prediction scheme for VANET," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2019, pp. 1–6.

[73] Y. Kim, P. Wang, and L. Mihaylova, "Structural recurrent neural network for traffic speed prediction," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 5207–5211.

[74] X. Guan, Y. Huang, M. Chen, H. Wu, T. Ohtsuki, and Y. Zhang, "Exploiting interference for capacity improvement in software-defined vehicular networks," *IEEE Access*, vol. 5, pp. 10662–10673, 2017.

[75] Y. Xia, X. Qin, B. Liu, and P. Zhang, "A greedy traffic light and queue aware routing protocol for urban VANETs," *China Commun.*, vol. 15, no. 7, pp. 77–87, Jul. 2018.

[76] F. Perronnet, J. Buisson, A. Lombard, A. Abbas-Turki, M. Ahmane, and A. El Moudni, "Deadlock prevention of self-driving vehicles in a network of intersections," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 11, pp. 4219–4233, Nov. 2019.

[77] A. Di Maio, M. R. Palattella, and T. Engel, "Multi-flow congestion-aware routing in software-defined vehicular networks," in *Proc. IEEE 90th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2019, pp. 1–6.

[78] M. Hossain, R. Hasan, and S. Zawoad, "Trust-IoV: A trustworthy forensic investigation framework for the Internet of vehicles (IoV)," in *Proc. IEEE Int. Congr. Internet Things (ICIOT)*, Jun. 2017, pp. 25–32.

[79] G. Gunduz and T. Acarman, "Efficient multi-object tracking by strong associations on temporal window," *IEEE Trans. Intell. Vehicles*, vol. 4, no. 3, pp. 447–455, Sep. 2019.

[80] Y. Zhang, J. Li, Z. Zhou, and X. Liu, "Efficient dynamic service maintenance for edge services," *IEEE Access*, vol. 6, pp. 8829–8840, 2018.

[81] G. Li, Q. Sun, L. Boukhatem, J. Wu, and J. Yang, "Intelligent vehicle-to-vehicle charging navigation for mobile electric vehicles via VANET-based communication," *IEEE Access*, vol. 7, pp. 170888–170906, 2019.

[82] M. A. Labiod, M. Gharbi, F.-X. Coudoux, P. Corlay, and N. Doghmane, "Cross-layer approach dedicated to HEVC low delay temporal prediction structure streaming over VANETs," in *Proc. Int. Conf. Smart Commun. Netw. Technol. (SaCoNeT)*, Oct. 2018, pp. 120–125.

[83] M. Y. Arafat and S. Moh, "Routing protocols for unmanned aerial vehicle networks: A survey," *IEEE Access*, vol. 7, pp. 99694–99720, 2019.

[84] R. Geng, X. Wang, and J. Liu, "A software defined networking-oriented security scheme for vehicle networks," *IEEE Access*, vol. 6, pp. 58195–58203, 2018.

[85] (2017). *OpenStreetMap Contributors*. [Online]. Available: https://planet.osm.org and https://www.openstreetmap.org

[86] P. A. Lopez, E. Wiessner, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flotterod, R. Hilbrich, L. Lucken, J. Rummel, and P. Wagner, "Microscopic traffic simulation using SUMO," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 2575–2582. [Online]. Available: https://elib.dlr.de/124092/

[87] C. Sommer, R. German, and F. Dressler, "Bidirectionally coupled network and road traffic simulation for improved IVC analysis," *IEEE Trans. Mobile Comput.*, vol. 10, no. 1, pp. 3–15, Jan. 2011.

[88] M. Sathiamoorthy, A. G. Dimakis, B. Krishnamachari, and F. Bai, "Distributed storage codes reduce latency in vehicular networks," *IEEE Trans. Mobile Comput.*, vol. 13, no. 9, pp. 2016–2027, Sep. 2014.

[89] J. Yuan, Y. Zheng, C. Zhang, W. Xie, X. Xie, G. Sun, and Y. Huang, "T-drive: Driving directions based on taxi trajectories," in *Proc. 18th SIGSPATIAL Int. Conf. Adv. Geograph. Inf. Syst. (GIS)*, 2010, pp. 99–108.

[90] E. Al-Masri and Q. H. Mahmoud, "Investigating Web services on the world wide Web," in *Proc. 17th Int. Conf. World Wide Web*, 2008, pp. 795–804.

[91] *Caltrans Pems*. Accessed: Feb. 15, 2020. [Online]. Available: http://pems.dot.ca.gov/

[92] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: User movement in location-based social networks," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2011, pp. 1082–1090.

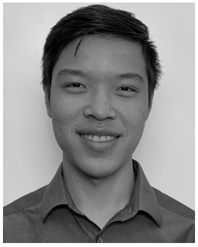
[93] E. Seta, "Project, a ubiquitous data and service ecosystem for better metropolitan mobility," SETA, Southampton, U.K., Tech. Rep., 2020. Accessed: Feb. 15, 2020. [Online]. Available: http://setamobility.weebly.com/case-studies.html

[94] M. Ranaweera, A. Seneviratne, D. Rey, M. Saberi, and V. V. Dixit, "Anomalous data detection in vehicular networks using traffic flow theory," in *Proc. IEEE 90th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2019, pp. 1–5.

SHAHAB TAYEB (Member, IEEE) received the Ph.D. degree. He is a Faculty Member of the Department of Electrical and Computer Engineering, Lyles College of Engineering, California State University, Fresno. Through funding from the Fresno State Transportation Institute, his research team has been working on the security of the network backbone for connected and autonomous vehicles over the past years. His research incorporates machine learning techniques and data analytic approaches to tackle the detection of zero-day attacks. His research expertise and interests include network security and privacy, particularly in the context of the Internet of Vehicles. He has been the recipient of several national awards, including the U.S. Congressional Commendation for STEM Mentorship.

SUMANJIT GILL is currently pursuing the degree in computer engineering with California State University, Fresno, through the Smittcamp Family Honors College Program. She has worked both as a Supplemental Instruction Leader and an Instructional Student Assistant of computer engineering at California State University. She is currently a Research Assistant with the Department of Electrical and Computer Engineering, California State University. Her current research interests include algorithmic and protocol analyses in the Internet of Vehicles. She is also a member of the Society of Women Engineers.

FLETCHER TRUEBLOOD received the bachelor's degree in computer science from California State University, Fresno, in 2019. He has been working as a Research Assistant with California State University. His recent work on message dissemination within the Internet of Vehicles was published at the proceedings of the IEEE CCWC. His current research interests are in the Internet of Vehicles and complex networks.



ROBERT WONG is currently pursuing the degree in computer engineering with California State University, Fresno, through the Smittcamp Family Honors College Program. He has been working as an Instructional Student Assistant and is currently an Undergraduate Research Assistant with the Department of Electrical and Computer Engineering, California State University. His research interests include the Internet of Vehicles and clean energy development. He is a member of the Eta Kappa Nu, Tau Beta Pi, and Phi Kappa Phi.

MATIN PIROUZ (Member, IEEE) received the Ph.D. degree. She is a Faculty Member of the Department of Computer Science and Engineering, California State University, Fresno. Her current project includes applying prescriptive and descriptive analyses. Her research has been funded by the National Science Foundation and other federal, state, and private funding agencies. Her current research interests include CS Education, big data analytics, social network analysis, and data mining.