MINISTRY OF EDUCATION
TECHNICAL UNIVERSITY
OF CLUJ-NAPOCA, ROMANIA

Computer Science and Information Technology

PhD THESIS
– ABSTRACT –

Vehicular Traffic Prediction and Congestion Avoidance based on Range Query Data Structures

PhD Student:
Eng. Ioan Stan

PhD Supervisor:
Prof. Eng. Rodica POTOLEA, PhD

Examination committee:
Chair: Prof. Eng. **Vasile Teodor Dădârlat**, PhD – Technical University of Cluj-Napoca;
PhD Supervisor: Prof. Eng. **Rodica Potolea**, PhD – Technical University of Cluj-Napoca;
Members:
-Prof. Eng. **Vasile Manta**, PhD – Gheorghe Asachi Technical University of Iași;
-Prof. Mat. **Bazil Pârv**, PhD – Babeș-Bolyai University from Cluj-Napoca;
-Prof. Eng. **Mihaela Dînșoreanu**, PhD – Technical University of Cluj-Napoca.

– Cluj-Napoca –
2023

Introduction

This thesis presents a novel vehicular traffic prediction and congestion avoidance methodology. The main innovation of the methodology is represented by the vehicular traffic representation in time and space through on range query date structures. The range query data structures are queried by the route computation algorithms to provide in real time the current vehicular traffic state and potential traffic that can be generated during current route computation. The vehicular traffic modeled by range query data structures is employed on road network topologies representing urban areas from real world.

Context and Motivation

Vehicular traffic in urban areas represent one of the main bottlenecks of our daily lives. During the day, the urban area road network become congested and passengers spend extra time in traffic. From an overall perspective, billions of extra hours are spent in congestion, which results in hundreds of billions of dollars cost [1, 3]. For example, in the major United States urban areas 32% of the daily travel time occurred under congested traffic [9]. The number of vehicles that can run on the world road networks grows every year. This grow is happening faster than the road network infrastructure evolution and therefore, the vehicular urban traffic congestion is increasing.

Considering the above context, one way to improve vehicular traffic is to find approaches that reduce vehicular congestion using vehicular traffic data in an intelligent way. Such approaches were researched by many, from various perspectives:

- One approach for vehicular traffic congestion is to use the road network infrastructure devices, when available: cameras, speed cameras, traffic sensors at intersections, traffic lights [15, 23, 25]. Such devices are considered to be part of the Intelligent Transportation System infrastructure. This kind of solutions require access to the infrastructure's devices and are limited to specific urban areas.

- Traffic prediction based on vehicular data coming from traffic data providers or directly from vehicles on the road is another way that can be employed to reduce traffic congestion [4, 9, 39]. Such vehicular data may be from past [4, 39] or real time data [4, 9]. For this kind of approaches it is required to have access to the traffic data services or to be connected with real vehicles and road network infrastructure. This implies financial costs, are time consuming and may reduce the traffic prediction approach to a specific area, specific set of vehicles or specific type of data received from traffic data services. Systems like Waze [4] are exceptions from these limitations because they are location based services that utilize mobile phones platform as data providers. The effectiveness of Waze like systems depends on the drivers approval to share their location, the vehicular traffic representation and the route computation algorithms. On the other hand, such approaches tend to be employed in a generic way and do not consider the particularities of specific urban road networks.

- A specific type of systems for vehicular traffic prediction and congestion avoidance are the ones that are focused on a limited (local) road network (e.g., a road network intersection or a set of road network intersections) [16, 24, 26]. Such methods cannot be easily extended to a full urban road network due to the complexity of such scenarios in comparison with the limited road network utilized for evaluation and validation of the systems.
- Cost effective and research oriented traffic prediction methods are based on simulation. There are various tools that may be employed to simulate vehicular traffic [22, 37]. These methods may employ real map data or synthetic generated maps. To model solutions from real world related scenarios, the main requirement of such systems is to mimic the reality as much as possible. This requirement implies first to utilize real map data. Second, it is important to employ route computation algorithms that are effective to be applied in real settings. Third, the vehicular traffic model should be scalable enough to support large scale traffic simulation. The road network on which the vehicular traffic is simulated must cover a comprehensive set of road network topologies representing real world urban areas.

Objectives

Considering the above, we found necessary to develop a novel approach for vehicular traffic prediction and congestion avoidance that is effective, scalable, comprehensive and ready to be applied in real world scenarios, especially in urban areas. To avoid cost risks like in [2], the cost of our development should be minimal. For the achievement of such goal we defined the following objectives.

O1: Systematic analysis and classification of the vehicular traffic prediction and congestion avoidance systems in order to find vehicular domains that we should exploit. This analysis should consider large scale vehicular traffic models from urban areas that give the main criteria for vehicular traffic scenarios definition and evaluation.

O2: Definition of the main vehicular traffic concepts: vehicular density, road network topologies and infrastructure settings.

O3: Research and identification of a novel vehicular traffic prediction and congestion avoidance methodology by:

- O3.1: Asking the relevant questions to query and update vehicular traffic state in a collaborative manner.
- O3.2: Designing the **3 pillars** architecture of a consistent vehicular traffic congestion avoidance methodology:
    1. Road Network Topology
    2. Route Computation Algorithms
    3. Range Query Data Structures for vehicular traffic modelling
- O3.3: Designing the flow of the methodology.
- O3.4: Defining the main parameters of the vehicular traffic model employed by range query data structures and route computation algorithms.
- O3.5: Developing a vehicular traffic simulation mechanism that can be employed to calibrate vehicular traffic simulation in the context of an urban congestion avoidance system.

This methodology should be able to predict and avoid urban vehicular traffic congestion in an effective and scalable manner by employing the 3 pillars: the vehicular traffic should be modelled on any road network topology defined by using range query data structures. Range query data structures should support vehicular traffic modeling in two dimensions: **time** and **space**. Vehicular traffic state requests should be answered in **real time** on any road segments during any time interval/at any specific moment in time. The route computation algorithms should be collaborative through usage of the state of current vehicular traffic and possible vehicular traffic generated by current computing route. Route computation algorithms should use range query data structures to query vehicular traffic state.

O4: Proposal of collaborative route computation algorithms to be employed by the vehicular traffic congestion avoidance methodology. The route computation algorithms should consider the current state of vehicular traffic and the potential traffic that can be generated by the current route, to avoid traffic congestion.

O5: Design and development of various range query data structures that can be applied to model vehicular traffic in an efficient and scalable manner. These data structures represent the main innovation of this study: vehicular traffic modeling through range query data structures in order to support **vehicular traffic state query and update in real time**. The range query data structures should:

- Model vehicular traffic on any road segment (*space*).
- Model vehicular traffic in *time*. Can answer vehicular traffic state during **time intervals** or at a **specific moment in time**.
- Store the entire state of the vehicular traffic on a road network topology.
- Query the state of possible *future* vehicular traffic on a road segment during route computation.
- Update the state of vehicular traffic based on computed route.

O6: Adaption of a vehicular traffic simulation environment (simulator) to integrate the proposed methodology by implementing the proposed route computation algorithms and range query data structures. The simulation environment should be anchored in reality through real world parameters configuration and should be widely used in the literature to simulate vehicular traffic.

O7: Definition of a comprehensive set of vehicular traffic scenarios for methodology evaluation. The scenarios definition should be based on the following dimensions:

- Vehicular Density (from *free flow* to *high congestion* in urban areas)
- Road Network Topology
- Infrastructure Configuration (e.g., average traffic light delay)
- Route Computation Algorithm
- Vehicular Traffic Simulation Environment

O8: Simulation of the proposed vehicular traffic scenarios in urban areas using the adapted vehicular traffic simulator. The vehicular traffic simulation should use real world maps that correspond to proposed road network topologies.

O9: Evaluation and validation of the proposed vehicular traffic congestion avoidance methodology to show its usability, effectiveness and scalability. The evaluation measurements should be done on the results coming from simulation and by using metrics like vehicular speed, time spent in traffic, estimated time of travel, fuel consumption and CO_2 emissions.

Thesis Overview

The thesis is structured in 7 sections similar with the structure of this summary. Section 1 introduces the context and problematic of the thesis and establishes the fundamentals. The objectives of the research are defined and discussed to emphasize the reasoning of the study.

Section 2 presents the context of vehicular traffic and its main aspects. This section discusses existing approaches in the literature that target vehicular traffic prediction and congestion avoidance, especially in the context of large scale simulated traffic. Here are presented the main aspects of vehicular traffic and, based on them, defined the core vehicular traffic tuple of this study.

Section 3 presents the proposed vehicular traffic congestion avoidance methodology that is based on three main pillars: road network topology, route computation algorithms and range query data structures. The main novelty of the study is represented by the fact that range query data structures are modelling vehicular traffic in time and space so that the state of vehicular traffic can be queried at any moment for any road segment. The presented methodology has a flow that integrates the three main pillars and a vehicular traffic simulation mechanism that is presented in the same section.

The fourth section defines and discuss various route computation algorithms employed by the methodology to compute vehicular routes in a collaborative manner. The collaborative route computation approach supports vehicular traffic prediction and congestion avoidance. The collaborative route computation strategy employs the range query data structures to query and update the vehicular traffic state.

The main innovation of the study is presented in Section 5. In this section are described and evaluated the range query data structures employed by the proposed methodology to model vehicular traffic. The range query data structures are used to store, query and update vehicular traffic on road segments in time. Except the segment tree data structure, all the presented data structures in this section are novel and focused on usability, effectiveness and scalability of vehicular traffic modelling. The evaluation of the proposed range query data structures are done in this section by measuring their performance in terms of CPU, memory footprint and scalability.

Section 6 evaluates and validates the usability, effectiveness and scalability of the proposed methodology. The vehicular traffic scenarios employed by the evaluation are defined using the above mentioned 5 dimensions (vehicular density, road network topology, infrastructure configuration, route computation algorithm and vehicular traffic simulation environment) and are simulated on real world maps that corresponds to the proposed road network topologies. During the first stage of the research we used OsmAnd to generate and simulate routes and due to scalability limitations, for the second stage of the research we adapted SUMO [22] to generate routes and simulate specific vehicular traffic scenarios in urban areas. SUMO is a simulation environment anchored in reality and is widely used in the literature to simulate vehicular traffic. It was even used to forecast vehicular traffic in the city of Cologne. The results obtained by SUMO scenarios shows the usability, effectiveness and scalability of the vehicular traffic congestion avoidance methodology.

Last section of the study concludes the research by discussing the overall contributions, usability, effectiveness and scalability of the proposed methodology. Limitations and further research directions are also discussed in this section.

Vehicular Traffic

Vehicular traffic prediction and congestion avoidance have been discussed, analyzed, and evaluated in various studies in the literature. This gave a good starting point for our research. In this section we identified 5 vehicular communicating models that can be used by an Intelligent Transportation System:

- **M1**: Vehicle to Vehicle (V2V) Model
- **M2**: Vehicle to Cloud (V2C) Model
- **M3**: Vehicle to Infrastructure (V2I) Model
- **M4**: Vehicle to Everything (V2X) Model
- **M5**: Vehicle to Vehicle to Infrastructure (V2V2I) Model

Existing vehicular traffic prediction and congestion avoidance solutions for urban areas encounter some challenges,from which we present the ones relevant for this study: limited access to intelligent transportation systems and their devices, limited access to vehicular traffic data from traffic service providers, lack of connection to vehicles, locality of the solution (only some road segments or intersections), implementation costs, simulation based solutions that doesn't consider real settings.

A particular and special aspect of our research regarding vehicular traffic prediction and congestion avoidance is represented by the large scale urban vehicular traffic congestion simulation. Table 1 summarizes the existing solutions that simulate large scale vehicular traffic congestion scenarios in various cities. It can be observed that all the solutions from literature simulates moderate vehicular traffic congestion on one or at most two road network topologies from urban areas (for simplicity and space reasons K symbol is used as a suffix

Table 1 Large Scale Urban Traffic Congestion Simulation Solutions.

Solution	**Simulator**	**Road Network**	**Area Size**	**Number of Simulated Vehicles**	**Maximum of Concurrent Vehicles** *
VNS [14]	DIVERT	Porto City	N/A	130K	15K
Elbery [10, 11]	Integration	Los Angeles	133 km^2	563K+	30K
Farag [13]	Integration	Los Angeles	133 km^2	145K	30K
FOXS [6]	SUMO	Ottawa, Cologne	8 km^2 and 4 km^2	2.2K and 14K	N/A
SOPHIA [5]	SUMO	Cologne	approx. 25 km^2	46K	N/A
Re-RouTE [18]	SUMO	Los Angeles, Paris	25 km^2	1.25K to 6.25K	N/A
ATOM [20]	SUMO	Baltimore	approx. 1 km^2	7.2K	N/A
Stan [32, 33]	OsmAnd	New York, Cluj-Napoca	N/A	10K and 20K	N/A
CAVTTM [27]	SUMO	Barcelona, Bucharest, New York, Tokyo	65 km^2	50K+ and 100K+	49K

* Peak number of concurrent vehicles running on the roads simultaneously.

to denote thousands of vehicles). Our solutions (last 2 rows from table) simulate free flow vehicular traffic and moderate and high vehicular traffic congestion on a comprehensive set of road network topologies representing various urban areas from real world. Also, our solution is based on SUMO simulation and reaches up to 49,000 concurrent vehicles that run on the road network.

To simulate large scale vehicular traffic scenarios in urban areas were defined vehicular traffic core elements that are presented below.

Vehicular Density on the road network can be classified as follows:

- **Free Flow** vehicular traffic corresponding to average vehicular speed above 25 km/h.

- **Moderate** vehicular traffic congestion corresponding to average vehicular speed between 10 km/h and 25 km/h.
- **High** vehicular traffic congestion corresponding to average vehicular speed below 10 km/h.

These vehicular density levels are **correlated with the number of simulated vehicles** and were modeled by simulating a large scale vehicular traffic representing between 10,000 and 100,0000 vehicles.

Another important characteristic regarding the vehicular traffic space is the *road network topology* of the urban area. In our study were defined four types of road network topology based on our work in [27, 31]:

- **T1**: *Grid Topology* representing almost ideal urban areas with dense road topology in the form of a grid, surrounded by highways and a proportional surface
- **T2**: *Unbalanced Grid Topology* representing urban areas with long roads, highways, grid topology, disproportional surface and pedestrian subareas
- **T3**: *Historical/Irregular Topology* representing urban areas with a small density of road infrastructure following random geometry based on historical roads and proportional surface
- **T4**: *Hybrid Topology* representing urban areas with rich road infrastructure that has many highways, large intersections and a proportional surface

For *infrastructure settings* we used the average traffic light delay to configure the vehicular scenarios.

These three vehicular traffic elements defines the tuple used in the study

$$<\textit{vehicular density, road network topology, infrastructure settings}> .$$

Any vehicle from traffic context follows the computed **route** between start and destination points of the vehicle. A route is composed by consecutive **road segments**. Any segment from the road network (map) is determined by 2 GPS points that represents road network intersections.

The findings in this section have been acknowledged by the scientific community through publication of 4 papers: 1 research paper at an ISI Q1 journal [32], 1 research paper at an ISI Q2 journal [27] and 2 papers at international conferences [29, 31].

Vehicular Traffic Congestion Avoidance Methodology

Most of the existing vehicular traffic prediction and congestion avoidance solutions encounter challenges that can't be controlled through their technical implementation. However, a particular aspect of these solutions is represented by the route computation strategy. This strategy is most of the time **individualistic** by trying to provide the optimal route for a specific vehicle. Considering this aspect, in this section was proposed a novel query model for vehicular traffic state. This model represent the main support for a **collaborative** route computation strategy. By this strategy it is intended to generate optimal routes for the global state of vehicular traffic that supports congestion avoidance. For this purpose were defined 3 types of vehicular traffic queries:

- **Q1**: queries the total number of vehicles that are on a road segment at any specific time interval
- **Q2**: queries the maximum number of vehicles on a road segment during any specific time interval
- **Q3**: queries the number of vehicles on a road segment at any specific moment

Querying vehicular traffic through the above queries is the base of the proposed methodology which is designed on 3 main pillars: *road network topology, route computation*

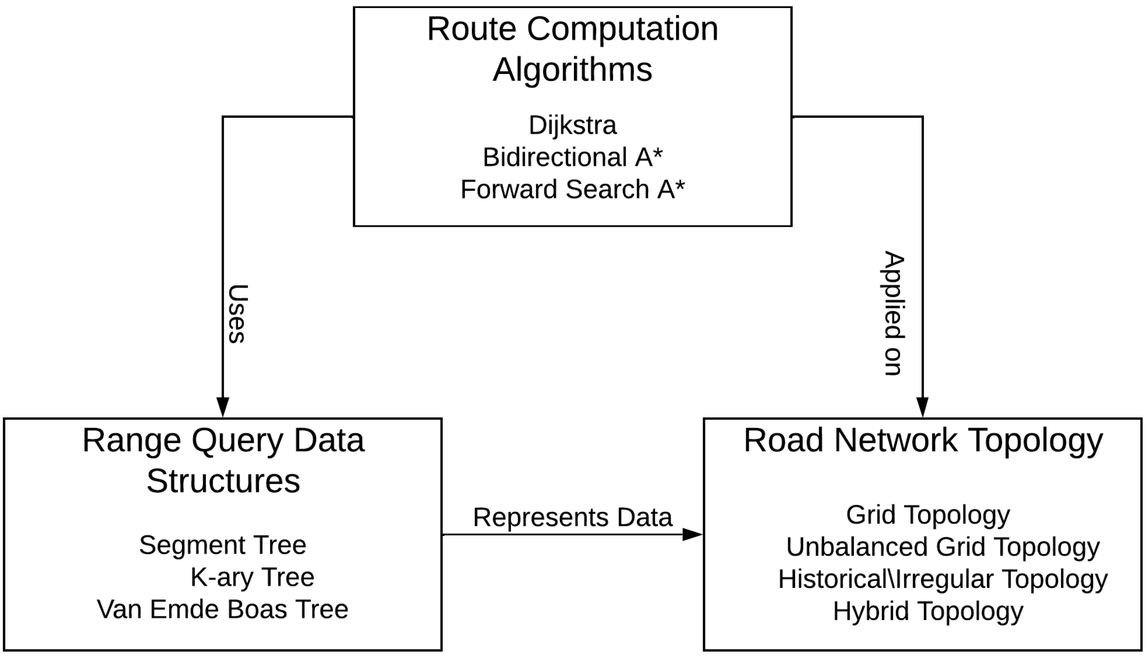

Fig. 1 Vehicular Traffic Congestion Avoidance Pillars

Table 2 Query - Range Query Data Structure Correlation

Query	Range Query Data Structure
Q1	*Segment Tree [33]*
Q2	*K-ary Interval Tree [32]*
Q3	*K-ary Entry Point Tree [32], Van Emde Boas Tree [34]*

algorithms and range query data structures. These pillars and their relation are shown in Figure 1: on a road network topology (from the defined ones) are generated the vehicular routes using collaborative route computation algorithms. The route computation algorithms are querying vehicular traffic state by considering the current vehicular traffic and the potential traffic that can be generated by the current computed route. In this way, the route computation algorithms can suggest alternative road segments in case of potential congestion on a road segment. Querying vehicular traffic state can be done based on the above defined queries and is based on vehicular traffic modeling through range query data structures. These data structures are employed to store, query and update vehicular traffic state on any **road segment (space) in time (time intervals or specific moments in time)**. Table 2 shows the correlation between above defined queries and proposed range query data structures.

The proposed vehicular traffic prediction and congestion avoidance methodology integrates the 3 main pillars in a 4 steps flow as shown in Figure 2.

S1—Random generation of the start and destination points for each vehicular route in the context of a selected road network topology from a comprehensive set of road network topologies.

S2—The output of S1 is consumed by the Route Computation Algorithm (RCA) to obtain routes that reduce or avoid vehicular traffic congestion. During route computation, various thresholding mechanisms can predict and avoid vehicular traffic congestion.

S3—The computed routes in S2 are employed to run vehicles on the selected road network topology in a vehicular traffic simulation environment. The vehicular traffic simulation environment can be adapted to log vehicular traffic data that is parsed in the next step.

S4—The generated vehicular traffic logs generated in S3 are parsed to provide data for evaluating the congestion reduction and avoidance.

Implementation of the route computation algorithms and range query data structures was done based on a mathematical model defined in this section. Also, for vehicular traffic simulator calibration was used a simulation mechanism implemented and validated through VISSIM [36].

The findings in this section have been acknowledged by the scientific community through publication of 4 research papers, presented at an ISI Q1 journal [32], an ISI Q2 journal [27] and international conferences [28, 31].

Vehicular Route Computation Algorithms

Vehicular route computation was analyzed and discussed in many studies [8, 17, 19, 21, 30, 35, 40], reaching a high level of maturity. However, most of the existing route computation algorithms are following an individualistic strategy during route computation. In this section, besides implementation of the **Basic Route Computation Algorithm - BRCA** that follows an individualistic strategy, were designed novel route computation algorithms that are following a collaborative strategy:

- **Collaborative Route Computation Algorithm - CRCA**: in case the vehicular density on a road segment is reaching a threshold value of congestion risk, this algorithm suggests alternative road segments for the route so that the congestion may be avoided on the road segment with such risk.

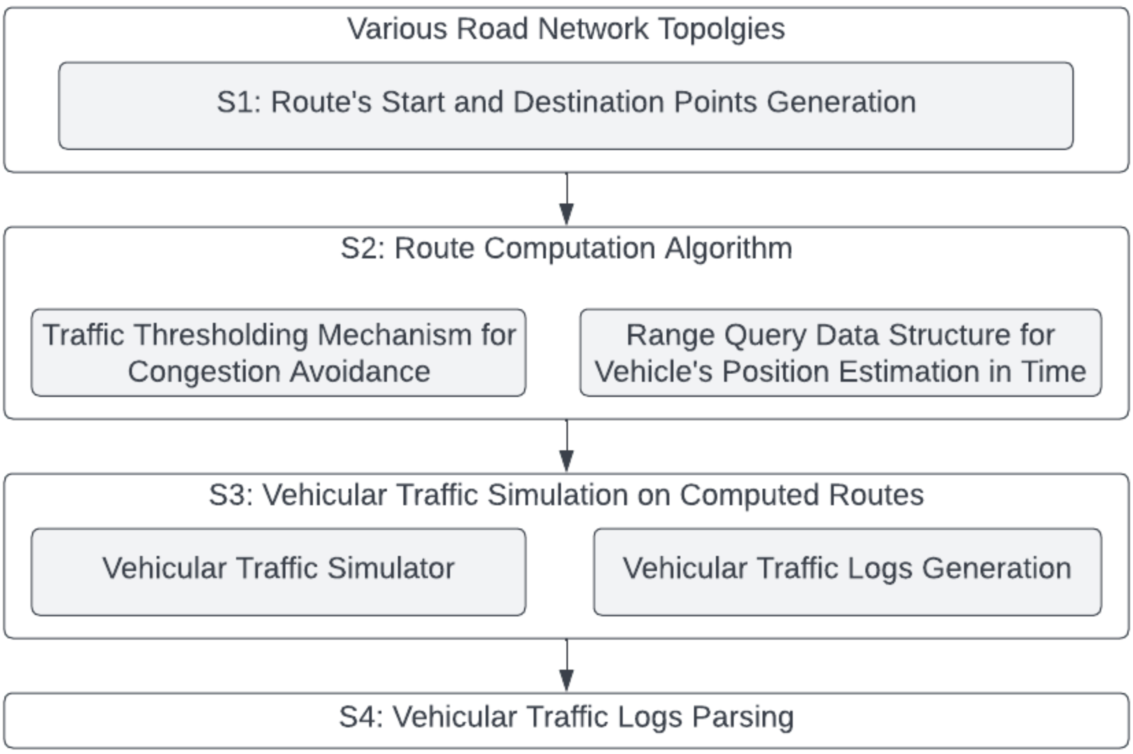

Fig. 2 Flow Diagram of Vehicular Traffic Congestion Avoidance System

Algorithm 1 Thresholding based CRCA

```
procedure TCRCA(p_s, p_d())
    init(queue)
    segment_s ← getSegment(p_s)
    segment_d ← getSegment(p_d)
    node ← segment_s
    queue.push(node)
    t_s ← t_0
    ETT ← t_0
    while queue is not empty do
        node ← queue.pop()
        if node == segment_d then
            for each segment in route(p_s, p_d) do
                speed ← ComputeSpeed(t_s, segment)
                d ← ComputeDelay(t_s, δ, segment)
                t_e ← t_s + L_segment / speed + d
                segment.Update(t_s, t_e)
                t_s ← t_e
            end for
        end if
        ETT ← getTravelTime(node)
        for each segment in N(node) do
            effort ← ComputeEffort(ETT, segment)
            if effort < E_segment then
                E_segment ← effort
            end if
        end for
    end while
end procedure
```

- **Full Bidirectional CRCA - FBCRCA**: this algorithm uses the same strategy as CRCA but considers all the alternatives during route computation (no heuristics).

- **Threshold based Collaborative Route Computation Algorithm - TCRCA**: the threshold introduced by CRCA was not sufficient to improve the vehicular traffic and avoid congestion in real world scenarios. Therefore, we found necessary to design a novel model based on more vehicular traffic density thresholds to be used during route computation algorithms. These thresholds are employed to compute **segment effort (cost), vehicular speed and delays introduced by vehicular traffic.** Vehicular traffic density has value in the range of [0..1]. Based on testing, this interval was split in many parts corresponding to vehicular traffic density thresholds.

TCRCA (presented in Algorithm 1) is efficient to reduce the time spent in traffic and to avoid congestion.

The novelty and findings in this section have been acknowledged by the scientific community through publication of 4 research papers, presented at an ISI Q1 journal [32], an ISI Q2 journal [27] and international conferences [30, 33].

Data Structures for Vehicular Traffic Modelling

To store, query and update vehicular traffic state on road segments in *real time* were designed and developed 5 novel range query data structures. Considering that any vehicle that follows a route runs on a specific road segment for a time interval, the vehicular traffic state can be modeled in time (time interval or specific moment in time) by considering the number of vehicles on any road segment. Figure 3 shows a simplified route that was initially determined by the start and destination points and is composed by road segments. A vehicle is running on a specific road segment during a time interval. The number of vehicles that are on a specific road segment (from the road network) during a time interval/specific moment in time, represents the vehicular density on that segment. The vehicular density can be stored, queried and actualized in real time by using range query data structures. This represents the vehicular traffic modeling on road network topologies.

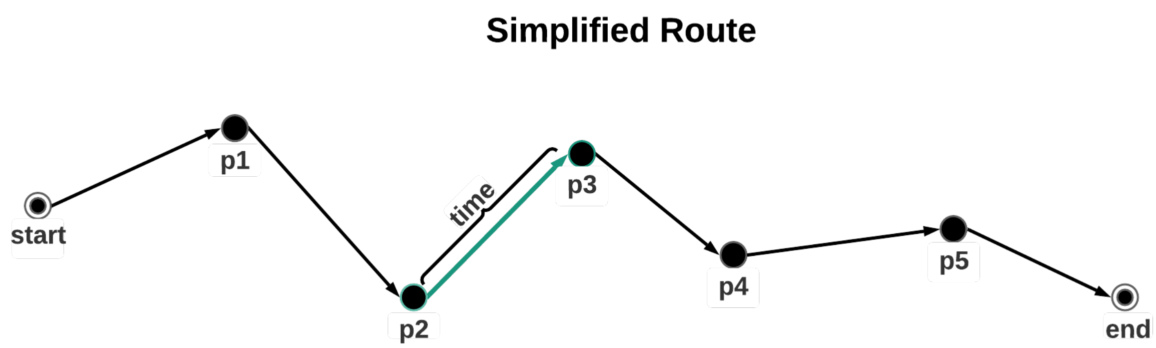

Fig. 3 Representation of a Simplified Route in M2 Context

In this study were employed 3 types of range query data structures:

- **Segment Tree** is a height-balanced binary tree that can model vehicular traffic on road segments during **a time interval**
- **K-ary Tree based** is a rooted tree data structure where each node has no more than k indexed children. These data structures were designed and developed to model

Table 3 Complexities of Range Query Data Structures

Operation	**Segment Tree**	**KI Tree**	**KEP Tree**	**VEB Tree**
Construction	O(n)	O(n)	O(n)	O(n)
Search	O(log n)	$O(k \cdot \log_k n)$	$O(k \cdot \log_k n)$	O(log log n)
Insert	O(log n)	$O(k \cdot \log_k n + \log_k^2 n)$	$O(\log_k n)$	O(log log n)
Delete	O(log n)	$O(k \cdot \log_k n + \log_k^2 n)$	$O(\log_k n)$	O(log log n)
Space	O(n)	O(n)	O(n)	O(n)

vehicular traffic on road segments during **a time interval** or at a **specific moment in time**.

- **Van Emde Boas Tree (VEB Tree) based** - is a rooted tree data structure based on associative arrays with m-bit integer keys. These data structures were designed and developed to model vehicular traffic on road segments at a **specific moment in time**. They have the best complexity on basic operations.

Table 3 shows the basic operations complexity of the range query data structures designed, developed and implemented in this study.

To query the vehicular traffic state is employed the **query** operation that corresponds to the basic operation of *search*. For vehicular traffic state **update** operation were combined two basic operations: *insert* and *delete*. To efficiently store the data represented by range query data structure were used indexed arrays.

The evaluation of the proposed data structures was done based on CPU performance, memory footprint and scalability. The results shown that, except segment tree, all the proposed range query data structures perform well and are scalable to model vehicular traffic in urban areas, even in case of large scale vehicular traffic. While KEP tree data structure perform very well on both query and update operations, the Partial Sums VEB Tree has the best practical performance for query operation.

The innovation and findings in this section have been acknowledged by the scientific community through publication of 3 research papers, presented at an ISI Q1 journal [32] and international conferences [33, 34].

Methodology Evaluation and Validation

To evaluate the proposed methodology were defined various vehicular traffic scenarios in urban areas (especially vehicular traffic congestion). These scenarios were defined based on 5 dimensions:

- Vehicular Density (from *free flow* to *high congestion* in urban areas)
- Road Network Topology
- Infrastructure Configuration (e.g., average traffic light delay)
- Route Computation Algorithm
- Vehicular Traffic Simulation Environment

During first stage of the research were simulated 5 urban traffic scenarios based on OsmAnd. In this scenarios were simulated between **10,000 and 20,000** vehicles. Depending on the time interval when the vehicles start to run on the roads, generating free flow or congestion. These scenarios were used to evaluate and validate the performance and scalability of the proposed range query data structures. On the other hand, the improvement of the time spent in vehicular traffic and on the congestion avoidance were minimal (below 3%).

Due to low performance and scalability limitations of OsmAnd, in the second stage of the research was employed another vehicular traffic simulator that is well anchored in reality, used widely in the literature and scalable: SUMO [22]. Moreover, for route computation was used a novel and effective route computation algorithm that reduced the time spent in vehicular traffic and reduced congestion: TCRCA.

Table 4 Comparison of Vehicular Traffic Congestion Avoidance Solutions

Solution	$\mathbf{t_{total}}\%$	$\mathbf{fuel_{total}}\%$	$\mathbf{CO_2}\%$
CAVTTM [27]	**69**	**61**	**61**
FOXS [6]	2.6	N/A	N/A
SOPHIA [5]	15	N/A	25.95
Re-RouTE [18]	65	N/A	N/A
Elouni [12]	41	20	N/A
Chunjiang [38]	31	N/A	N/A
CACC [7]	48.3	N/A	N/A

With SUMO were generated **26 scenarios** of urban vehicular traffic that mostly represents congestion (moderate and high) and a free flow scenario. Were simulated between **10,000 and 100,000** vehicles to model large scale vehicular traffic. These vehicles started to run during a given time interval. In this way we simulated a high number of urban traffic scenarios that run on a comprehensive set of road network topologies. Were reached up to **49,000** concurrent vehicles on a road network. Another noteworthy result is the fact that for congestion scenarios the time spent in traffic was reduced by up to **69%** and fuel consumption by up to **61%**. Considering the results shown in Table 4 we can state that, in comparison with existing solutions in the literature, the proposed methodology provides the best improvement of time spent in traffic, fuel consumption and CO_2 emissions reduction.

The novelty, and results from this section have been acknowledged by the scientific community through publication of 4 research papers, presented at an ISI Q2 journal [27] and international conferences [29, 30, 33].

Thesis Contributions

The main goal of this study was to design and develop a novel methodology for vehicular traffic prediction and congestion avoidance that should be effective, scalable, comprehensive and applicable on real urban traffic scenarios. This goal was reached by achieving the following contributions from thesis sections.

The main contributions in Section 2 are:

- **C2.1** Systematic analysis and classification of the vehicular traffic prediction and congestion avoidance systems with particular focus on large scale vehicular traffic scenarios. This analysis includes a critical and comparative analysis of existing vehicular traffic simulation tools and identification of a vehicular traffic simulatior that fulfills our objectives.

- **C2.2** Definition of the main vehicular traffic concepts modeled as a tuple
  <*vehicular density, road network topology, infrastructure settings*> .

The main contribution of Section 3 is **C3** - Design and development of a novel, effective and scalable vehicular traffic prediction and congestion avoidance methodology:

- **C3.1** Defined collaborative vehicular traffic.

- **C3.2** Identified 3 vehicular traffic queries that can be used to query and update vehicular traffic state on road segments in time.

- **C3.3** Designed the vehicular traffic prediction and congestion avoidance architecture based on 3 pillars *Road Network Topology, Route Computation Algorithms and Range Query Data Structures for vehicular traffic modelling*.
- **C3.4** Integrated the methodology's pillars into a flow that implements and evaluates the methodology.
- **C3.5** Defined the mathematical model employed by route computation algorithms and range query data structures.
- **C3.5** Developed a vehicular traffic calibration mechanism for simulation.

In Section 4 were designed and developed route computation algorithms that can be applied to predict vehicular traffic and avoid congestion (**C4**):

- **C4.1** Implemented Basic Route Computation Algorithm (BRCA) that uses individualistic route search strategy. The algorithm was integrated into OsmAnd.
- **C4.2** Designed and developed a novel Collaborative Route Computation Algorithm (CRCA) that considers vehicular traffic state during route computation. The algorithm was integrated into OsmAnd.
- **C4.3** Designed and developed a novel Full Bidirectional Collaborative Route Computation Algorithm (FBCRCA) that considers vehicular traffic state during route computation and all route alternatives. The algorithm was integrated into OsmAnd.
- **C4.4** Designed and developed a novel Threshold based Collaborative Route Computation Algorithm (TCRCA) that showed to be effective in reducing time spent in traffic and fuel consumption. The algorithm was integrated into SUMO.

The main innovation of this thesis is **C5** represented by the design and development of range query data structures to efficiently model vehicular traffic on road segments in time. The traffic model is used to store, query and update vehicular traffic state in *real time*. This was achieved through the below contributions:

- **C5.1** Implemented Segment tree to store, query and update vehicular traffic on road segments during time intervals. The Segment tree was integrated into OsmAnd.
- **C5.2** Designed and developed a novel and scalable K-ary Interval tree to store, query and update vehicular traffic on road segments during time intervals. The K-ary Interval tree was integrated into OsmAnd.

- **C5.3** Designed and developed a novel and scalable K-ary Entry Point tree to store, query and update vehicular traffic on road segments at a specific moment in time. The K-ary Entry Point tree was integrated into OsmAnd and SUMO.
- **C5.4** Designed and developed a novel and scalable Basic Augmented VEB tree to store, query and update vehicular traffic on road segments at a specific moment in time. The Basic Augmented VEB tree was integrated into OsmAnd.
- **C5.5** Designed and developed a novel and scalable Layered VEB tree to store, query and update vehicular traffic on road segments at a specific moment in time. The Layered VEB tree was integrated into OsmAnd.
- **C5.6** Designed and developed a novel and scalable Partial Sums VEB tree to store, query and update vehicular traffic on road segments at a specific moment in time. The Partial Sums VEB tree was integrated into OsmAnd.
- **C5.7** Evaluation and validation of **CPU performance, memory usage and scalability** of the proposed range query data structures.

The methodology evaluation in Section 6 shown the following contributions:

- **C6.1** Defined vehicular traffic scenarios that are applied on a comprehensive set of road network topologies representing urban areas from the world. The scenarios definition was based on 5 dimensions: vehicular density, road network topology, average traffic light delay, route computation algorithm and vehicular traffic simulation environment.
- **C6.2** Simulated 5 vehicular traffic scenarios in OsmAnd which is a real navigation solution and 26 large scale vehicular traffic scenarios in SUMO which is a scalable simulation environment anchored in reality, utilized in the literature as research tool and utilized in real world to forecast vehicular traffic in City of Cologne. The large scale vehicular traffic scenarios simulated between 10K and 100K+ vehicles (50K+ for moderate congestion and 100K+ for high congestion). This study covered the **most** comprehensive set of large scale vehicular traffic scenarios simulated in the literature.
- **C6.3** Reached **49K** concurrent vehicles at a specific moment on a road network topology (**most** number of concurrent vehicles known in the literature).
- **C6.4** Evaluated the results coming from simulation and found that the proposed vehicular traffic prediction and congestion avoidance methodology is **scalable** to model large scale traffic scenarios and **effective** to improve time spent in traffic, fuel consumption and CO_2 emissions:

- for the scenarios with 50K+ vehicles it reduced the time spent in traffic by more than **69%**, fuel consumption by up to **61%** and CO2 emissions with more than **61%** (the highest improvement know in the literature);
- for the scenarios with 100K+ vehicles it reduced the time spent in traffic by more than **32%**, fuel consumption by up to **31%** and CO2 emissions by more than **31%**.

- **C6.5** Balanced the vehicular traffic in a Grid road network topology and improved the number of simulated vehicles in free flow by **200%**.

In Table 5 are correlated the objectives stated in the first section with the contributions summarized in the current section.

Table 5 Objectives - Contributions Correlation

Objective	**Contribution(s)**	**Short Description**
O1	C2.1	Classification of vehicular traffic prediction and congestion avoidance solutions.
O2	C2.2	Main aspects of vehicular traffic.
O3 (O3.1 - O3.5)	C3 (C3.1 - C3.5)	Vehicular traffic prediction and congestion avoidance methodology.
O4	C4 (C4.1 - C4.4)	Route Computation Algorithms.
O5	C5 (C5.1 - C5.7)	Range Query Data Structures.
O6	C4; C5	Vehicular Simulation Environment Adaption.
O7	C6.1	Vehicular Traffic Scenarios.
O8	C6.2; C6.3	Simulation of Vehicular Traffic Scenarios.
O9	C6.3-C6.5	Methodology Evaluation: Usability, Effectiveness and Scalability.

Due to cost and efficiency reasons of the methodology development, during this study was used simulation. This doesn't limit the usability, but opens new directions of development of the methodology in real world (as are presented in the further work of the thesis).

Dissemination

The contributions of this thesis have been published in prestigious international journals and conferences in computer science and intelligent transportation systems domains. Below is presented a summary of our research activity and the list of publications:

ISI Journals	**2** (first author) 1 × Q1, IF 3.847 - Sensors 1 × Q2, IF 2.838 - Applied Sciences
International conferences	**6** (first author)
International conferences before PhD stage	**2** (second author)
Total Publications	**10**
Citations	**30**
H-index	**4**

Publications at ISI Journals

Stan, I.; Ghere, D.A.; Dan, P.I.; Potolea, R. Urban Congestion Avoidance Methodology Based on Vehicular Traffic Thresholding. Applied Sciences 2023, 13, 2143. https://doi.org/10.3390/app13042143; **Q2, IF 2.838**

Stan I, Suciu V, Potolea R. Scalable Data Model for Traffic Congestion Avoidance in a Vehicle to Cloud Infrastructure. Sensors. 2021; 21(15):5074. https://doi.org/10.3390/s21155074; **Q1, IF 3.847**

Publications at International Conferences

I. Stan, B. -P. Ungur and R. Potolea, "Augmented Van Emde Boas Tree for Connected Vehicles Traffic Modeling," 2021 IEEE 17th International Conference on Intelligent Computer Communication and Processing (ICCP), 2021, pp. 199-208, doi:10.1109/ICCP53602.2021.9733462.

Ioan Stan, Raul Ghisa, and Rodica Potolea. Urban Traffic Simulation Methodology for Connected Vehicles Congestion Avoidance. In Proceedings of the 22nd International Conference on Information Integration and Web-based Applications; Services (iiWAS 2020). Association for Computing Machinery, New York, NY, USA, 305–312. https://doi.org/10.1145/3428757.3429102

I. Stan, V. Suciu and R. Potolea, "Smart Driving Methodology for Connected Cars," 2019 23rd International Conference on System Theory, Control and Computing (ICSTCC), 2019, pp. 608-613, doi: 10.1109/ICSTCC.2019.8885450.

I. Stan, V. Suciu and R. Potolea, Routing Algorithms in Connected Cars Context, Collection of open conferences in research transport (2019). Vol. 2019, 354 URL https://www.scipedia.com/public/Stan_et_al_2019a

Stan, I., Potolea, R. (2018). Connected Cars Traffic Flow Balancing Based on Classification and Calibration. In: Groza, A., Prasath, R. Mining Intelligence and Knowledge Exploration. MIKE 2018. Lecture Notes in Computer Science, vol 11308. Springer, Cham. https://doi.org/10.1007/978-3-030-05918-7_16

I. Stan, D. Toderici and R. Potolea, "Segment Trees based Traffic Congestion Avoidance in Connected Cars Context," 2018 IEEE 14th International Conference on Intelligent Computer Communication and Processing (ICCP), 2018, pp. 137-143, doi: 10.1109/ICCP.2018.8516609.

Publications Before PhD Stage

A. Colesa and I. Stan, "Improving the Responsiveness of Replicated Virtualized Services in Case of Overloaded Replicas Connectivity," 2012 14th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, 2012, pp. 287-294, doi: 10.1109/SYNASC.2012.13.

A. Colesa, I. Stan and I. Ignat, "Transparent Fault-Tolerance Based on Asynchronous Virtual Machine Replication," 2010 12th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, 2010, pp. 442-448, doi: 10.1109/SYNASC.2010.58.

Selected References

[1] (2017). Statistics of vehicle in use between 2006 to 2015. Available online: https://www.statista.com/statistics/281134/number-of-vehicles-in-use-worldwide/ (last accessed online on 2020-07-07).

[2] (2019). Report: £300m traffic jam-busting scheme made some journeys longer. Available online: https://www.aol.co.uk/news/2019/04/07/a-300m-traffic-jam-busting-scheme-made-some-journeys-longer-a/ (last accessed on 2019-05-13).

[3] (2021). Urban mobility report. Available online: https://mobility.tamu.edu/umr/report/ (last accessed on 16 April 2021).

[4] (2023). Routing server - waze. Accesible online: https://wiki.waze.com/wiki/Routing{_}server (last accessed on 2023-02-12).

[5] Akabane, A. T., Immich, R., Pazzi, R. W., Madeira, E. R. M., and Villas, L. A. (2019). Exploiting vehicular social networks and dynamic clustering to enhance urban mobility management. *Sensors*, 19(16).

[6] Brennand, C. A. R. L., Filho, G. P. R., Maia, G., Cunha, F., Guidoni, D. L., and Villas, L. A. (2019). Towards a fog-enabled intelligent transportation system to reduce traffic jam. *Sensors*, 19(18).

[7] Cao, Z., Lu, L., Chen, C., and Chen, X. (2021). Modeling and simulating urban traffic flow mixed with regular and connected vehicles. *IEEE Access*, 9:10392–10399.

[8] Chen, C. L. P., Zhou, J., and Zhao, W. (2012). A real-time vehicle navigation algorithm in sensor network environments. *IEEE Transactions on Intelligent Transportation Systems*, 13(4):1657–1666.

[9] Chu, K., Saigal, R., and Saitou, K. (2019). Real-time traffic prediction and probing strategy for lagrangian traffic data. *IEEE Transactions on Intelligent Transportation Systems*, 20(2):497–506.

[10] Elbery, A. and Elnainay, M. (2019). Vehicular communication and mobility sustainability: the mutual impacts in large-scale smart cities.

[11] Elbery, A., Rakha, H. A., and ElNainay, M. (2019). Large-scale modeling of vanet and transportation systems. In Hamdar, S. H., editor, *Traffic and Granular Flow '17*, pages 517–526, Cham. Springer International Publishing.

[12] Elouni, M., Abdelghaffar, H. M., and Rakha, H. A. (2021). Adaptive traffic signal control: Game-theoretic decentralized vs. centralized perimeter control. *Sensors*, 21(1).

[13] Farag, M. M. G., Rakha, H. A., Mazied, E. A., and Rao, J. (2021). Integration large-scale modeling framework of direct cellular vehicle-to-all (c-v2x) applications. *Sensors*, 21(6).

[14] Fernandes, R., Vieira, F., and Ferreira, M. (2012). Vns: An integrated framework for vehicular networks simulation. In *2012 IEEE Vehicular Networking Conference (VNC)*, pages 195–202.

[15] Figueiredo, L., Jesus, I., Machado, J. A. T., Ferreira, J. R., and de Carvalho, J. L. M. (2001). Towards the development of intelligent transportation systems. In *ITSC 2001. 2001 IEEE Intelligent Transportation Systems. Proceedings (Cat. No.01TH8585)*, pages 1206–1211.

[16] Fu, X., Gao, H., Cai, H., Wang, Z., and Chen, W. (2021). How to improve urban intelligent traffic? a case study using traffic signal timing optimization model based on swarm intelligence algorithm. *Sensors*, 21(8).

[17] Gendreau, M., Potvin, J.-Y., Bräumlaysy, O., Hasle, G., and Løkketangen, A. (2008). *Metaheuristics for the Vehicle Routing Problem and Its Extensions: A Categorized Bibliography*, volume 43, pages 143–169. Springer US.

[18] Guidoni, D. L., Maia, G., Souza, F. S. H., Villas, L. A., and Loureiro, A. A. F. (2020). Vehicular traffic management based on traffic engineering for vehicular ad hoc networks. *IEEE Access*, 8:45167–45183.

[19] Kanoh, H. and Hara, K. (2008). Hybrid genetic algorithm for dynamic multi-objective route planning with predicted traffic in a real-world road network. pages 657–664.

[20] Lee, S., Younis, M., Murali, A., and Lee, M. (2019). Dynamic local vehicular flow optimization using real-time traffic conditions at multiple road intersections. *IEEE Access*, 7:28137–28157.

[21] Liao, T.-Y. and Hu, T.-Y. (2011). An object-oriented evaluation framework for dynamic vehicle routing problems under real-time information. *Expert Systems with Applications*, 38(10):12548 – 12558.

[22] Lopez, P. A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.-P., Hilbrich, R., Lücken, L., Rummel, J., Wagner, P., and Wießner, E. (2018). Microscopic traffic simulation using sumo. In *The 21st IEEE International Conference on Intelligent Transportation Systems*. IEEE.

[23] Martinez, F., Toh, C.-K., Cano, J.-C., Calafate, C., and Manzoni, P. (2010). Emergency services in future intelligent transportation systems based on vehicular communication networks. *IEEE Intell. Transport. Syst. Mag.*, 2:6–20.

[24] Olayode, I. O., Tartibu, L. K., and Okwu, M. O. (2021). Prediction and modeling of traffic flow of human-driven vehicles at a signalized road intersection using artificial neural network model: A south african road transportation system scenario. *Transportation Engineering*, 6:100095.

[25] Papadimitratos, P., Fortelle, A. D. L., Evenssen, K., Brignolo, R., and Cosenza, S. (2009). Vehicular communication systems: Enabling technologies, applications, and future outlook on intelligent transportation. *IEEE Communications Magazine*, 47(11):84–95.

[26] Qi, L., Zhou, M., and Luan, W. (2017). Impact of driving behavior on traffic delay at a congested signalized intersection. *IEEE Transactions on Intelligent Transportation Systems*, 18(7):1882–1893.

[27] Stan, I., Ghere, D. A., Dan, P. I., and Potolea, R. (2023). Urban congestion avoidance methodology based on vehicular traffic thresholding. *Applied Sciences*, 13(4).

[28] Stan, I., Ghisa, R., and Potolea, R. (2020). Urban traffic simulation methodology for connected vehicles congestion avoidance. In *iiWAS '20: The 22nd International Conference on Information Integration and Web-based Applications & Services, Virtual Event / Chiang Mai, Thailand, November 30 - December 2, 2020*, pages 305–312. ACM.

[29] Stan, I. and Potolea, R. (2018). Connected cars traffic flow balancing based on classification and calibration. In *Mining Intelligence and Knowledge Exploration - 6th International Conference, MIKE 2018, Cluj-Napoca, Romania, December 20-22, 2018, Proceedings*, pages 177–188.

[30] Stan., I., Suciu., V., and Potolea., R. (2019). Routing algorithms in connected cars context. In *Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management - Volume 2: KEOD,*, pages 290–297. INSTICC, SciTePress.

[31] Stan, I., Suciu, V., and Potolea, R. (2019). Smart driving methodology for connected cars. In *2019 23rd International Conference on System Theory, Control and Computing (ICSTCC)*, pages 608–613.

[32] Stan, I., Suciu, V., and Potolea, R. (2021a). Scalable data model for traffic congestion avoidance in a vehicle to cloud infrastructure. *Sensors*, 21(15).

[33] Stan, I., Toderici, D., and Potolea, R. (2018). Segment trees based traffic congestion avoidance in connected cars context. *2018 IEEE 14th International Conference on Intelligent Computer Communication and Processing (ICCP)*, pages 137–143.

[34] Stan, I., Ungur, B.-P., and Potolea, R. (2021b). Augmented van emde boas tree for connected vehicles traffic modeling. In *2021 IEEE 17th International Conference on Intelligent Computer Communication and Processing (ICCP)*, pages 199–208.

[35] Tatoimir, B., Rothkrantz, L., and Suson, A. (2010). Travel time prediction for dynamic routing using ant based control. pages 1069 – 1078.

[36] Vissim, P. (2020). Traffic simulation software | ptv vissim | ptv group. Accessed: 2020-07-08.

[37] Wegener, A., Piorkowski, M., Raya, M., Hellbrück, H., Fischer, S., and Hubaux, J.-P. (2008). Traci: An interface for coupling road traffic and network simulators. *Proceedings of the 11th Communications and Networking Simulation Symposium, CNS'08*.

[38] Wu, C., Zhou, S., and Xiao, L. (2020). Dynamic path planning based on improved ant colony algorithm in traffic congestion. *IEEE Access*, 8:180773–180783.

[39] Wu, Y., Tan, H., Qin, L., Ran, B., and Jiang, Z. (2018). A hybrid deep learning based traffic flow prediction method and its understanding. *Transportation Research Part C: Emerging Technologies*, 90:166–180.

[40] Zavin, A., Sharif, A., Ibnat, A., Abdullah, W. M., and Islam, M. N. (2017). Towards developing an intelligent system to suggest optimal path based on historic and real-time traffic data. In *2017 20th International Conference of Computer and Information Technology (ICCIT)*, pages 1–6.