Budapest University of Technology and Economics
Faculty of Electrical Engineering and Informatics
Department of Networked Systems and Services

Performance Analysis of Vehicular Ad-hoc Networks

Ph.D. DISSERTATION OF

Dhari Ali Mahmood Ghrairi

Scientific Supervisor
Prof.Dr. Gábor Horváth

October 19, 2020

DECLARATION

I, undersigned *Dhari Ali Mahmood Ghrairi* hereby declare that this Ph.D. dissertation was made by myself, and I only used the sources given at the end. Every part that was quoted word-for-word, or was taken over with the same content, I noted explicitly by giving the reference to the source.

Budapest, Hungary, October 19, 2020

Dhari Ali Mahmood Ghrairi

ABSTRACT

The dissertation presents different stochastic models in two main fields of self organized networks: in the field of packet schedulers of devices in telecommunication systems and in the field of vehicular ad-hoc networks. The proposed stochastic models focus on the performance evaluation; the accurate analysis models lead to improved efficiency of the system and reduced cost.

Several telecommunication devices such as routers and switches use Weighted Fair Queueing (WFQ) scheduler algorithm to differentiate between different traffic streams. The proposed solution provides a simple explicit approximation for mean response time in WFQ systems based on an algebraic approach.

The efficiency of Intelligent Transportation Systems (ITS) are affected by several factors (i.e. driving safety, collision avoidance, and etc.). ITS based on vehicular ad-hoc networks can be improved by proposing a mathematical model that describes the message propagation process along the road. We have introduced new contributions regarding the message propagation in VANET systems, the proposed model characterizes the stationary and the transient solutions of the message propagation distance. The traffic jam caused by an accident also affects the distance of the message propagation, so we also take the effect of the queueing system also into consideration in the proposed solution. We validated our analytical results with simulation (Veins and SUMO within OMNET++).

To cover another aspect of ITS, we also consider the roadside units (RSUs) to improve the message propagation efficiency, we present a model to derive the speed of the message propagation in the presence of RSUs. Regarding realistic data and our proposed solution, we show that the proper placement of RSUs can improve the message propagation speed drastically. The results can be used for network planning in VANETs with disconnected RSUs.

Furthermore, for the arrivals of vehicles the Poisson process is sufficiently accurate in many cases, but not always, so we proposed new contribution based on another arrival process called Markvian Arrival Process (MAP) to describe the inter-arrival times of the vehicles. We have derived accurate mathematical models to evaluate several properties of the message propagation distance, such as moments, complementary cumulative distribution functions (ccdf) of the stationary cluster length, stationary and transient distribution of the information distance. The analysis based on the robust mathematical model can improve the performance of telecommunication devices and VANET technology as well.

ACKNOWLEDGMENTS

First and foremost I would like to express my gratitude to my supervisor professor Gábor Horváth for his guidance, encouragement, and patience throughout my research. I would like to thank my colleagues in Department of Networked Systems and Services for their help and support during my research. I also want to thank the Stipendium Hungarian for funding my research. Last but not least, I would like to thank my parents, my family, my wife Nawres and my beautiful daughters (Maryam, Sarah and Sally) for their unconditional love, support, understanding, and encouragement.

CONTENTS

ACRONYMS

QOS	Quality Of Service
MANET	Mobile Ad-hoc Network
VANET	Vehicular Ad-hoc Network
V2V	Vehicle-to-Vehicle
V2I	Vehicle-to-Infrastructure
V2X	Vehicle-to-anythings
RSU	Road Side Unit
ITS	Intelligent Transportation System
MAP	Markovian Arrival Process
WFQ	Weight Fair Queueing
SCV	Squared Coefficient of Variation
SUMO	Simulation of Urban Mobility
COS	Class Of Service
DSCP	Diffserv Code Point
GPS	Generalized Processor Sharing
QBD	Quasi Birth-Death
FCFS	First-Come-First-Served
MAC	Media Access Control
pdf	probability density function
cdf	cumulative distribution function
DDE	Delayed Differential Equation
ccdf	complementary cumulative distribution function
PASTA	Poisson Arrival See Averages
CTMC	Continuous Time Markov Chain
LST	Laplace-Stieltjes Transform

NOTATIONS

μ	Service rate
λ_1, λ_2	Arrival rate of class 1 and class 2 customers
w_1, w_2	Weight of class 1 and class 2 customers
ρ_1, ρ_2	Utilization of class 1 and class 2 customers
ρ	The total utilization of the system
r	The asymmetry of the utilization
$E(T_1), E(T_2)$	Mean response time (waiting+service time) of class 1 and class 2
$E(T_{FCFS})$	Mean response time (waiting+service time) of two-class FCFS
$E(T_{Prio})$	Mean response time (waiting+service time),WFQ as priority queue
w^*	The weight of inflection point between class 1 and class 2
c	Parameter of control the shape in the given function
$\hat{E}(T_1)$	Scale of region $E(T_1)$ takes values between $E(T_{Prio})$ and $E(T_{FCFS})$
R	The radius of the radio transmission coverage
L	Length of a vehicle
v	The (constant) speed of the vehicles
λ	Vehicle arrival rate, in vehicles/second
ϑ	Exponentially distributed vehicle density(vehicle/meter)
$\mathcal{D}(t)$	The stochastic process of the information distance
$\underline{F}(t,x)$	Phase dependent transient ccdf
$\mathcal{G}$	Cluster length
$G(x)$	Phase independent ccdf of the cluster length
$E(\mathcal{G})$	The mean cluster length $\mathcal{G}$
$\mathcal{D}, E(\mathcal{D}), Var(\mathcal{D})$	The stationary information distance, its mean value and variance
$\mathcal{X}$	Length of M/D/1 queue
π	The (row) vector of stationary probabilities
$\mathcal{B}(t)$	The message propagation distance based on M/D/1 queue
v_{slow}	Slow vehicle speed at accident
D	Distance between any two adjacent RSUs
$\hat{R}$	Radio coverage of the RSUs
C	Asymptotic message propagation speed
$\mathcal{L}(t)$	Index of the farthest RSU informed at time t
$\mathcal{H}$	Duration of intervals when no inform vehicles
$\mathcal{Q}_k$	Time points where the farthest informed RSU
U_k	Positions of the RSUs, k=1,...
$\mathcal{G}'$	Cluster length when vehicle enters the coverage of an RSU
$\mathcal{S}_k$	Renewal points, k=1,...
$\mathcal{T}_k$	Time between renewal instants, k=1,...
$E(\mathcal{S})$	Mean distance of information jump from inform point
$E(\mathcal{T})$	The mean time between two renewal instants
c_S, c_T	Distance and the time taken since the last renewal point
$L_i^*(s)$	Laplace transform of the transient distribution

$\delta_{i,j}$	Kronecker delta
$L^*(s,z)$	Double tranform of the transient distribution
D'	Distance to the first RSU
$\boldsymbol{D_0}, \boldsymbol{D_1}$	The matrices characterizing the MAP of the vehicles arrivals process
$\boldsymbol{D}$	Generator matrix of the background process, $\boldsymbol{D} = \boldsymbol{D_0} + \boldsymbol{D_1}$
$\mathbb{1}$	Column vector of ones of appropriate size
$\underline{0}$	Column vector of zeros of appropriate size
I	Identity matrix of appropriate size
$\mathcal{J}(t)$	The phase process of the MAP at time t
$\boldsymbol{P}$	Phase transition probabilities for consecutive vehicle inter-arrivals
$\underline{\pi}$	The stationary phase distribution vector right after vehicle arrivals
$\underline{\alpha}$	The stationary probability vector of the MAP
$\boldsymbol{E}(\mathcal{G})$	Phase dependent of the mean value of cluster length
$\mathcal{T}_k, h(t), H(t)$	Random variable of kth vehicle inter arrival time, its pdf and its cdf
$\widehat{\mathcal{J}}(t)$	Phase of the MAP when cluster at time t will leave the accident
$\boldsymbol{G}(x)$	The phase dependent ccdf of the cluster length
$\boldsymbol{g}(x)$	The phase dependent pdf of the cluster length
$\boldsymbol{Z}$	Transition probabilities of MAP in beginning and end of a cluster
$\underline{\gamma}$	Stationary phase distribution of MAP at beginning of a cluster
$\mathcal{D}, E(\mathcal{D})$	The stationary information distance and its mean value
$\underline{F}(x)$	The phase-dependent ccdf of the stationary information distance
$F(x)$	The phase-independent ccdf of the stationary information distance
$\mathcal{H}, E(\mathcal{H})$	The random variable representing the distance between clusters
$\underline{\beta}(t)$	Transient phase dependent probabilities, no informed vehicles
$\underline{\beta}$	Stationary phase dependent probabilities, no informed vehicles
$\mathcal{C}$	The cycle time, the inter arrival time between two cluster heads
$\hat{\rho}$	The lag-1 correlation of vehicles inter-arrival times

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

1 INTRODUCTION

1.1 OVERVIEW

The user demand in computer networks has been increasing exponentially in the last decades. New technologies have been added to the network devices (i.e., routers, switches) of computer networks to manage the increasing demand in proper way [3]. The complexity in management, routing, use of network resources (i.e., bandwidth), and so on, are important challenges, they are the starting points of many research work. In computer and telecommunication networks the overall traffic is a mixture of packet flows having different quality demands. Some packets are urgent, while some others can tolerate delay better. Most modern communication protocols have a field in the packet header indicating to which class the packet is belonging to (like the class of service (CoS) field in the Ethernet frame header and the DiffServ code point (DSCP) in the IP header). Packet schedulers in the network devices (switches, routers) need to take this information into account to provide the necessary quality of service. A popular multi-class scheduling discipline for this purpose is the weighted fair queueing (WFQ) service. In such systems the packets belonging to different traffic classes are stored in separate queues before they get transmitted. The total service capacity is shared among the classes according to the weights associated with the queues: The higher the weight of a traffic class is, the higher service rate it gets. The WFQ schedulers are work conserving, thus the total service capacity is always distributed among the classes that are currently active. The weights provide a flexible way to express the importance of the traffic classes. However, the analysis of the WFQ schedulers is challenging.

Moreover, the modern communication systems have appeared in many fields of our life. These technologies are present anywhere and anytime. One approach of commutation systems is called Mobile Ad-hoc Network (MANET) [13], which is a wireless system without central control, which consists of freely moving nodes. MANET is a self-configuring wireless network. A MANET comprises of mobile nodes, a router with multiple hosts and wireless communication devices. The telecommunication devices are based on wireless transmitters, receivers, and smart antennas within radio coverage. A new technology based on MANET is called vehicular ad-hoc network (VANET). VANET make the communication between the vehicles and exchange of information possible. Still, these communication services face many challenges due to different parameters (i.e., delay in received information, interference, etc.) lead to reduce the efficiency of communication systems [15]. In VANET, the vehicle-to-vehicle communication is often abbreviated as V2V, while the vehicle-to-infrastructure communication is usually referred to as V2I communication. As there are increasing number of vehicles on the roads, such a communication platform has many benefits, including

- decreasing the number of accidents caused by driver inattention,

- increasing the efficiency of transportation by platooning,
- propagating information about events on the highways,
- etc.,

thus the major objective is to increase both the efficiency of transportation and the road safety. The first step in realizing this objective is to equip vehicles with communication capabilities, thus, in the first period, only vehicles will be able to communicate, we can only talk about V2V communication. In this scenario, vehicles close enough to each other form so-called *clusters*, and only vehicles belonging to the same cluster can exchange messages with each other. The most efficient solution to improve the connectivity is to introduce infrastructure nodes, so-called *road-side units (RSUs)*, which can also participate in the message forwarding process (both V2V and V2I communication, [58]). Unfortunately, due to the high cost, complexity and lack of cooperation between government and private sectors the deployment of RSUs is slow [60]. There are two kinds of RSUs: in the *unconnected* case the RSUs are not able to communicate with each other, while in the *connected* case there is a direct communication channel (apart from the radio) between them. Due to the aforementioned factors, RSUs will probably be unconnected in the first deployed VANET systems.

1.2 PROBLEM STATEMENTS

Telecommunication devices in general and in VANETs are facing many *challenges that are caused by constraints* such as limited buffer capacity, narrow radio bandwidth, short radio coverage, cost of established communication Infrastructure. These limitations affect the performance, such as resource consumption, delay of data delivery, speed of message propagation, and network reliability. To overcome these limitations these systems should be optimized based on economical and user satisfaction. The mathematical analysis can play an important role in the optimization of these systems based on stochastic models, it can help in understanding the behavior of these systems better.

Our research work has been organized along the following four problem statements.

1.2.1 *Queueing system associated with network devices*

The weighted fair queueing (WFQ) service discipline provides a flexible way to share bandwidth among two or more traffic classes. Some variants of the basic WFQ principle are used in the practice in computer networks in routers, switches, etc. Unfortunately, the analytical modeling of the related queues turned out to be notoriously difficult.

The first problem that is associated with telecommunication devices (i.e., routers, switches, etc.) have many challenges and limitation as follows:

- There is no accurate formula for the mean response times that could be used to optimize the usage of network resources.
- Such a formula could enable the dynamic assignment of weights associated with the classes depending on the current network conditions.

1.2.2 *Alert message propagation in VANETs without RSUs*

To provide sufficient quality of service, it is important to develop analytical models to calculate various performance measures related to the message propagation. For instance, by computing how far the message propagates, and how long it takes to deliver a message to the vehicles, proposed model can help the drivers to take appropriate decisions on time. Having the message received, the driver can leave the road before reaching the accident and the corresponding traffic jam, or vehicles can maintain a suitable distance between each other to avoid accidents.

VANETs can have the following issues:

- The delivery delay of (alert) messages exchanged between the vehicles can be too high.
- There are no stochastic models available for the stationary and transient solutions of the alert message propagation distance.

1.2.3 *Alert message propagation in VANETs with disconnected RSUs*

The appropriate position of the road-side units (RSUs) can play an important role in the alert message propagation. Appropriate positioning of RSUs can increase the speed of message propagation and decrease the number of accidents, and that will lead to increased safety of the transportation system.

Problems to solve in VANETs with RSUs are:

- There is no formula providing the asymptotic message propagation speed as the function of the RSU distance. Such a formula could be useful to determine the optimal distance between RSUs and to calculate the effect of speed restrictions on message propagation.
- The transient distribution of the distance where the message is available is unknown.

1.2.4 *Failure of the Poisson process for modeling vehicle inter-arrival times*

The Poisson process has been used in the literature for a long time to model the inter-arrival times of the vehicles. The reason for this modeling choice is not its validated correctness, but its analytical simplicity; many useful performance measures can be expressed analytically when Poisson traffic is assumed. The Poisson process is, however, not a good model for the vehicular traffic in many cases. Due to the failure of the Poisson process in the modeling of vehicular traffic, different stochastic models, such Markovian Arrival Processes (MAPs) should be investigated.

Challenges in this field:

- MAPs are not commonly known by the VANET community.
- No results are available for the message propagation in case of MAP vehicle arrival process.

1.3 RESEARCH OBJECTIVE

As mentioned in Section 1.2, many challenges in the telecommunication system can be solved by introducing accurate analysis based on the mathematical model, that we used in our research work.

Telecommunication system devices have a stream of packets as input, like many other systems such as queueing systems, electrical systems, and many other examples. The accurate description of the input stream of these systems is important to obtain accurate performance measures, however, in many cases the Poisson arrival process is used due to its simplicity. We also use Poisson arrival process in Chapters 2, 3 and 4. Furthermore, when the Poisson model is not accurate enough, we can describe the input stream with the extension of the Poisson process, called the Markovian Arrival Process (MAP), as we proposed in Chapter 5.

The main objective of our research is to provide new and efficient performance analysis methods for self-organized networks. The proposed models make it possible to identify what kind of key performance indicators are available to make re-configuration decisions leading to minimal cost (i.e., channel bandwidth), improved quality of service, development of robust analytical models for the intelligent transportation system. The proposed solution describes the aforementioned problems in Section 1.2 as follows. *In the telecommunication systems, the packet scheduling algorithms* (i.e., weighted fair queueing) become an important part of buffer management in these devices to improve the performance.

In the network devices of self-organized networks the parameters of the WFQ packet scheduler can be adjusted on-the-fly to ensure that the quality requirements are met in changing traffic characteristics. To support this functionality it is important to have a fast algorithm for the analytical calculation of the main performance measures. Such an algorithm can also be used to check the sensitivity of the WFQ scheduler on the traffic characteristics. This objective is elaborated in Chapter 2.

Intelligent transportation systems (ITS) are becoming more and more important nowadays, motivating our research work to develop analytical models to investigate realistic communication systems based-on vehicular ad-hoc networks (VANETs). Providing assistance to drivers on the road based on alert messages received from other vehicles in an appropriate time can help in accident avoidance, as proposed in Chapter 3.

Furthermore, *road-side unit placed in the appropriate position on the highway* increase the efficiency of the transportation systems and road safety by contributing to the transferring of messages. High enough alert message propagation speed can decrease the number of accidents on the road. The solution proposed in Chapter 4 can be used to characterize the dynamics of communication and support the planning process of VANETs based-on disconnected RSUs.

Another objective of our work is *to introduce MAPs to model vehicle inter-arrival times*, when the Poisson arrival model turns out to be inaccurate. In most of the literature regarding VANET technology, the Poisson arrival model is considered as the vehicle's arrival process, but this model does not hold in several cases. Multiple statistical investigation (i.e., probability density function, SCV, and correlation) applied on the realistic data has shown that the Poisson model does not fit for VANET technology in many cases. The properties of the message propagation in VANET, including the

distance how far the messages can get, are affected by many parameters, such as the radio channel, the communication protocols but also the traffic model of the vehicles. This fact motivated our research work to use MAPs for vehicle arrivals as proposed in Chapter 5.

To give an understanding of the main parts of the dissertation and their relation with each other, the schematic diagram in Figure 1.1 shows the contributions of the dissertation.

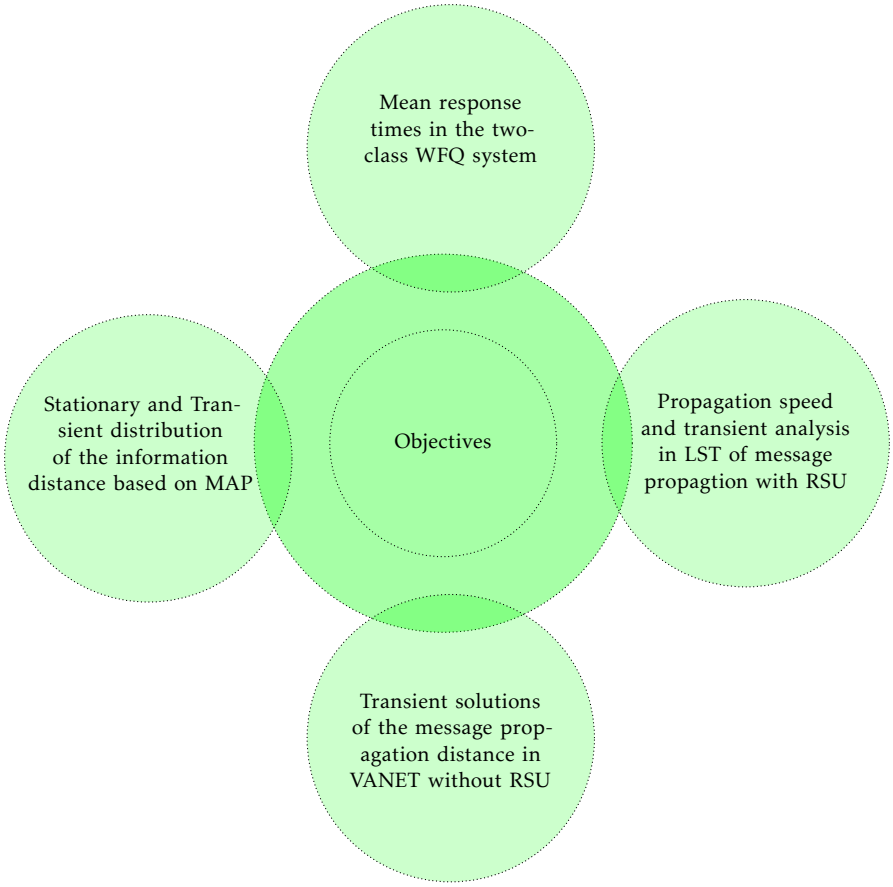


Figure 1.1: Contributions of the dissertation

CHAPTER TWO

2

A SIMPLE APPROXIMATION FOR THE RESPONSE TIMES IN THE TWO-CLASS WEIGHTED FAIR QUEUEING SYSTEM

In the recent decades, queueing systems become very important part in many fields [3]. In these systems customers (customers of a bank or packets in a communication system) arrive according to a kind of arrival process and are served. In some queueing systems these customers are classified according to their types. In the field of telecommunication, the packets arriving to a router or switch need to be put into different queues (depending on their class) before forwarding them to the appropriate destination. The weight associated with a class determines the portion of bandwidth it receives from the packet scheduler.

2.1 RELATED WORK

The fluid-based version of the scheduler, where the customers are infinitesimally small (considered as fluid drops), is often called generalized processor sharing (GPS), while the variant with discrete customers, also studied in this chapter, is called weighted fair queueing. According to the ideal weighted fair queueing (also referred to as the coupled processor model in [17], [66] and [50]) the multiple traffic classes can be served simultaneously, at the reduced service rate associated to them. This ideal WFQ is, however, impossible to implement in a real situation. Several packet-based approximations of the ideal WFQ appeared in the practice, including the Virtual Clock [73], Self-Clocked Fair Queueing [20], Deficit Round Robin [56], etc.

Although these WFQ-like schedulers are very popular, there are very few analytical results available in the literature. In the simplest scenario with Poisson arrival process and exponentially distributed service times, the Markov chain representing the number of customers in the system (class-wise) has a simple, regular structure, still, its stationary solution turned out to be a notoriously difficult problem. The only exact result we are aware of is [22], where the generating function was derived (the mathematical apparatus used in that paper demonstrates how difficult the problem is). Several approximations appeared as well to provide simpler, more tractable solution of the WFQ system. The result in [24] is based on the decoupling of the queues, while the QBD structure of the Markov chain is exploited in [2]. The idea in [55] is to transform the WFQ system to a priority queue.

In this chapter our aim is to provide a very simple, explicit approximation for the two-class ideal WFQ system, based on simulation results and curve fitting. Similar approach has been followed many times in the past: the KLB formula [33] for the approximation of the waiting time in G/G/1 queues and the formulas in [68] to approximate various properties of the departure process of G/G/1 queues were both successful and widely used results. Proposed approach is somewhat similar to [55], but that paper considers a slightly different system where the service of packets can not be preempted, and a step of the procedure needs the numerical solution of an equation. My formulas are

explicit, contain only basic operations, and can be easily implemented in a network device, making it possible to re-calculate the weights of the classes if the traffic situation changes.

2.2 CONCEPT OF WEIGHTED FAIR QUEUEING

The weighted fair queueing service discipline provides a flexible way to share bandwidth among two or more traffic classes. Some variants of the basic WFQ principle are used in the practice in computer networks in routers , switches, etc., as the schematic diagram shows in Figure 2.1. Unfortunately, the analytical modeling of the related queues turned out to be notoriously difficult. This chapter presents approximation expressions for the mean response times in a two-class (ideal) WFQ system with Poisson arrival process and exponentially distributed service times. The approximation is based on simulation. The results are very simple, explicit, yet reasonably accurate, ideal to use in self organizing networks where the weights associated with the different traffic classes need to be recalculated to adapt to the changing network conditions.

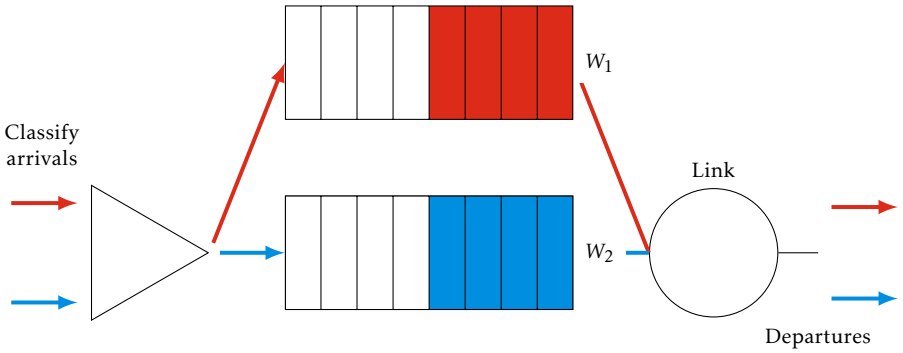


Figure 2.1: Schematic diagram for Weighted fair queueing diagram

2.3 MODEL DEFINITION

In this chapter we consider the two-class weighted fair queueing system. The customers are arriving according to a Poisson process with parameters λ_1 and λ_2, and are directed to two separate queues according to their class. The service times are exponentially distributed with (class independent) parameter μ. The server is shared among the two customer classes, controlled by weights w_1 and w_2. According to the ideal weighted fair queueing policy considered in the chapter, both the class 1 and class 2 queues are served in parallel, if both kinds of customers are present in the system: class 1 is served with rate $\mu \cdot w_1/(w_1+w_2)$, while class 2 is served with rate $\mu \cdot w_2/(w_1+w_2)$. If one of the queues is idle then the total service capacity is given to the other class.

The amount of work brought by class k customers to the system is $\rho_k = \lambda_k/\mu$. In this chapter we assume that the system is stable, hence for the total utilization $\rho = \rho_1 + \rho_2$ we have that $\rho < 1$. In the fully utilized case, the FCFS is a good approximation, the WFQ as we know shares the bandwidth between the classes, so if utilization is high, all classes will get the reduced service in the system.

The asymmetry of the utilization of the two customer classes can be characterized many ways. we found that the measure

$$r = \frac{\rho_1 - \rho_2}{\rho}, \tag{2.1}$$

$r \in (-1, 1)$, turned out to be a good choice, making the forthcoming expressions simpler.

2.3.1 *Analytical results used in the chapter*

To approximate the mean response time in the WFQ system, we are going to utilize the results of two closely related two-class queueing systems, that have exact mean response time results available.

One of these systems is the two-class FCFS (First Come First Serve) queue. In this system there is no capacity sharing and all demands are served according to the global arrival order independent on the class. The mean response time is given by [10]

$$E(T_{FCFS}) = \frac{1}{\mu - \lambda_1 - \lambda_2}. \tag{2.2}$$

The second queueing system necessary to our approximation is the two-class preemptive priority queue. Observe that the ideal WFQ server investigated in this chapter exhibits a kind of preemptive behavior: when a customer arrives to an idle queue, the service rate of the other class gets reduced immediately. When one of the weights, w_1 or w_2 is zero, then the WFQ behaves like a preemptive (resume) priority queue. If $w_1 = 0$, class 1 plays the role of the low priority class with mean response time given by

$$E(T_{Prio}) = E(T_{FCFS})\frac{1}{1 - \rho_2}. \tag{2.3}$$

Finally, the conservation law [10]

$$\rho_1 E(T_1) + \rho_2 E(T_2) = \rho E(T_{FCFS}) \tag{2.4}$$

allows us to focus on one of the classes only, the mean response time for the opposing class can be calculated from (2.4).

2.3.2 *Simulation of the system*

Due to the naming confusion and the existence of many variants of the WFQ system, and since we rely on simulation results heavily in the chapter, we briefly discuss the simulation of the queue studied in this chapter. Algorithm 1 provides the simplified simulation algorithm in a discrete event simulation system[1]

1 Our implementation is based on OmNet++ [64]

Algorithm 1 Discrete event simulation of the WFQ system

Event $end\_of\_service_i$:
    collect (current time - arrival time) to response time statistics
    remove customer from queue_i
    if queue_i is empty **then**
        call RescheduleServiceTimes
    else
        select next customer in queue_i
        service time $\leftarrow \text{Exp}(\mu \cdot share_i)$
        schedule $end\_of\_service_i$ **to** current time + service time
    end if
End
Event $arrival\_of\_class_i$:
    if queue_i is empty **then**
        add new customer to queue_i
        call RescheduleServiceTimes
    else
        add new customer to queue_i
    end if
End
procedure RescheduleServiceTimes
    for every non-empty queue_i **do**
        $share_i \leftarrow w_i / \sum_{\substack{\forall j:\text{queue}_j \\ \text{not empty}}} w_j$
        cancel event $end\_of\_service_i$
        service time $\leftarrow \text{Exp}(\mu \cdot share_i)$
        schedule $end\_of\_service_i$ **to** current time + service time
    end for
end procedure

In this algorithm there are two kinds of events to handle: arrival and service events. There are as many service events scheduled at the same time as many busy queues there are in the system. These events need to be re-scheduled when the busy state of the queues changes, that can occur in two situations: when a customer arrives into an empty queue and when a customer leaves its queue empty. In case of exponentially distributed service times re-scheduling the service events is simple, the memory-less property can be exploited.

2.4 THE ANALYSIS OF THE RESPONSE TIMES

2.4.1 *The concept of the approximation*

Due to the conservation law (2.4) it is enough to focus on a single customer class, class 1, the mean response time for the other traffic class can be expressed from (2.4). An other feature of the system that we are going to exploit is that the two weight parameters

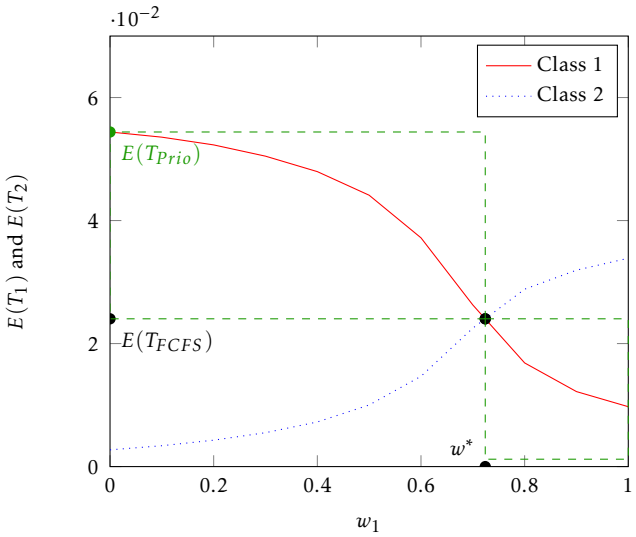


Figure 2.2: The mean response times as the function of w_1 ($\mu = 0.0012$)

w_1, w_2 defining the system are redundant. In the sequel, we are going to set $w_2 = 1$ and investigate the behavior of the system as the function of w_1.

Figure 2.2 depicts the mean response times as the function of w_1 in a particular example ($\rho_1 = 0.39, \rho_2 = 0.55$). Observe that if $w_1 = 0$ then the system behaves like a preemptive priority queue with class 1 being the low priority class, hence $E(T_1) = E(T_{Prio})$. At the other hand, when $w_1 \to \infty$, class 1 has exclusive access to the service capacity. The point where the curves of class 1 and class 2 meet plays an important role in our approximation. In this point $E(T_1) = E(T_2)$ holds, more precisely, (2.4) implies that $E(T_1) = E(T_2) = E(T_{FCFS})$. The weight belonging to this point is denoted by w^* in the sequel. Based on this point the plot of class 1 on the figure can be divided to two rectangular regions (denoted by dashed lines). Due to the symmetry of the system, we assume that $w_1 \leq w^*$ holds (the role of the two classes can be swapped in the opposite case), we are going to study only this case in the rest of the chapter, hence our aim is to approximate the behavior in the rectangle on the left.

Our approximation for the response times consists of two components:

- *The approximation of w^*.* This is the only unknown parameter to fully characterize the region marked by dashed lines in Figure 2.2. The top left point is given by $w_1 = 0, E(T_1) = E(T_{prio})$, and the bottom right point is located at $w_1 = w^*, E(T_1) = E(T_{FCFS})$.

- *The approximation of the shape of the response time curve.* Based on many simulation experiments we found that w^* is very close to the inflection point in most of the cases (except if the utilization is extremely low). Hence, $E(T_1)$ inside the dashed region is typically monotonous. The bend of the curve (referred to as the "shape parameter" in the sequel) depends on ρ_1 and ρ_2, and it is also subject to approximation. We found that with some parameters w^* is not in the inflexion point (although close to it); hence this assumption does not always hold, but it is a good approximation.

The next two subsections present the approximation of these two parameters.

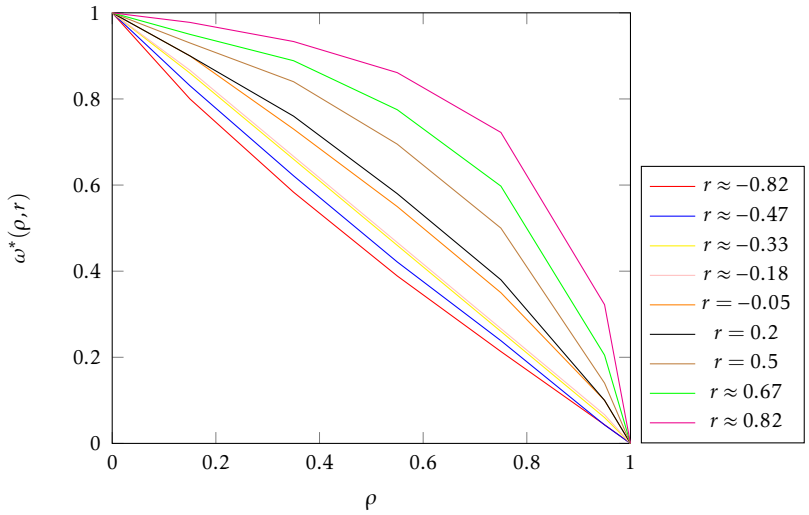


Figure 2.3: The parameter $\omega^*(\rho,r)$ at various ρ and r values

2.4.2 *Approximating the weight w^**

we have studied the behavior of w^* as the function of $\rho = \rho_1 + \rho_2$ with different r parameters (r characterizes the asymmetry, see (2.1)). we found that at the two extreme values of ρ the $w^*(\rho,r)$ tends to specific values:

- At $\rho \to 0$ w^* tends to 1,
- at $\rho \to 1$ w^* tends to $\rho_1/\rho_2 = \frac{1+r}{1-r}$.

The latter relation can be intuitively justified as follows: When ρ is almost one, the system is continuously busy so the class-1 queue is like an $M/M/1$ queue with arrival rate λ_1 and service rate $\frac{\mu w_1}{w_1+1}$. So, the mean response time of class-1 customers is $E(T_1(w_1)) = \frac{w_1+1}{w_1(\mu-\lambda_1)-\lambda_1}$. To obtain w^* one has to solve $E(T_1(w_1)) = \frac{1}{\mu-\lambda_1-\lambda_2}$, which, after a few calculations leads to $w_1 = \frac{1-\rho_2}{\rho_2} = \frac{\rho_1}{\rho_2}$.

To make the visual comparison easier, we scale w^* to the $[0,1]$ domain by introducing $\omega^*(\rho,r)$ as

$$\omega^*(\rho,r) = \frac{w^*(\rho,r) - \frac{\rho_1}{\rho_2}}{1 - \frac{\rho_1}{\rho_2}} = \frac{w^*(\rho,r) - \frac{1+r}{1-r}}{1 - \frac{1+r}{1-r}}. \tag{2.5}$$

Figure 2.3 depicts the shape of $\omega^*(\rho,r)$ as the function of ρ at various settings of r. According to our simulation experiments, these curves are (almost) symmetric to the $f(x) = x$ line. we found a family of functions, having the same symmetry, suitable for the approximation: the $f(x) = \frac{x-1}{cx-1}$ function. The c parameter of this function controls the shape: how much the curve bends towards the upper right corner of the rectangle defined by corners $(0,0)$ and $(1,1)$. Shape parameter c can be negative as well, in this case the curve bends towards the lower left corner instead.

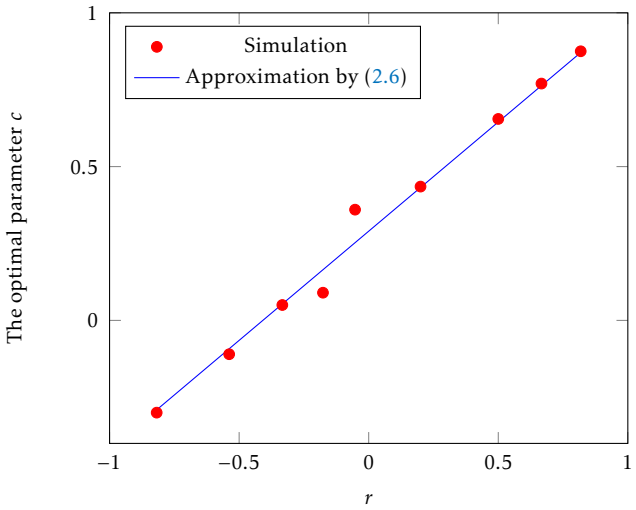


Figure 2.4: The optimal shape parameter c for the approximation of $\omega^*(\rho,r)$

For each r setting we determined the optimal shape c value leading to the most accurate approximation. Plotting these parameters as the function of r leads to a nearly linear function (Figure 2.4)

$$c = 0.71r + 0.29. \tag{2.6}$$

Putting together the pieces, the approximation for $\omega^*(\rho,r)$ is

$$\omega^*(\rho,r) = \frac{\rho-1}{c\rho-1} = \frac{\rho-1}{(0.71r+0.29)\rho-1} = \frac{\rho-1}{0.71(\rho_1-\rho_2)+0.29\rho-1}, \tag{2.7}$$

finally, the approximation for w^* is

$$w^* = \omega^*(\rho,r)(1-\frac{\rho_1}{\rho_2})+\frac{\rho_1}{\rho_2} = \left(1-\frac{\rho_1}{\rho_2}\right)\frac{\rho_1+\rho_2-1}{\rho_1-0.42\rho_2-1}+\frac{\rho_1}{\rho_2}. \tag{2.8}$$

Note that $\omega^*(\rho,r)$ (hence w^*) is always non-negative in the stability region since the minimum value is given at $\rho = 1$.

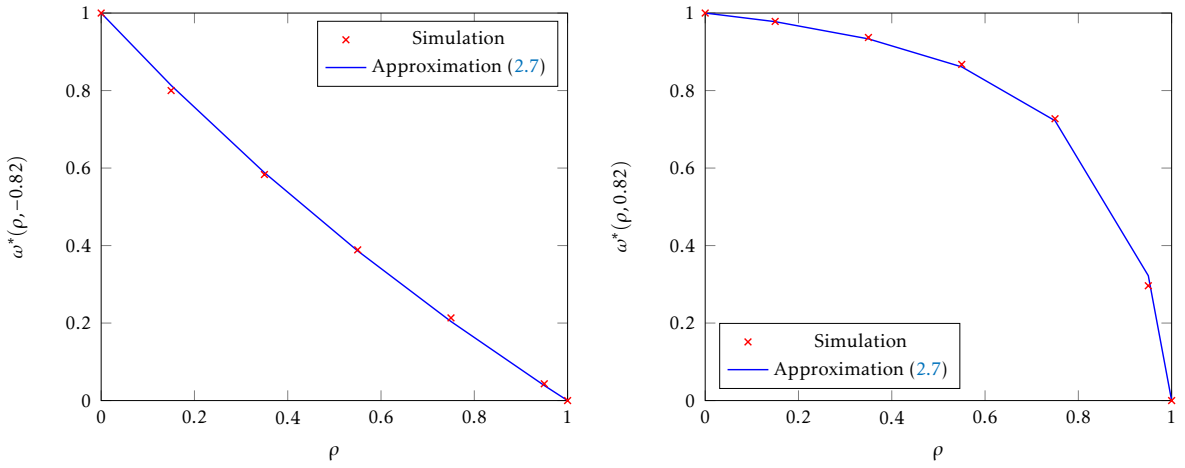
Figure 2.5 demonstrates how accurate this approximation is in two cases, for $r = -0.82$ and for $r = +0.82$, corresponding to $\rho_1/\rho_2 = 0.1$ and 10, respectively (the dashed line is the approximation, the marks indicate the simulation results).

2.4.3 *Approximating the shape*

Having w^* approximated, the next element of the solution is to approximate $E(T_1)$ when $w_1 \in [0,w^*]$ (the curve in the upper left dashed rectangle in Figure 2.2). In this region $E(T_1)$ takes values between $E(T_{Prio})$ and $E(T_{FCFS})$. Let us scale this region to the $[0,1]$ domain by defining $\hat{E}(T_1)$ as

$$\hat{E}(T_1) = \frac{E(T_1)-E(T_{FCFS})}{E(T_{Prio})-E(T_{FCFS})} \tag{2.9}$$

and investigate its dependence on the parameters of the system. Depending on the class 1 and class 2 load the curve representing $\hat{E}(T_1)$ as the function of w_1/w^* bends



a) $r = -0.82$ b) $r = 0.82$

Figure 2.5: The accuracy of the approximation of parameter ω^*

towards the lower left or towards the upper right corner of the unit rectangle. The function $f(x) = \frac{x-1}{cx-1}$, introduced and used in the previous section, turned out to be suitable to approximate this curve as well. As before, the question is how to set the shape parameter c to make the approximation accurate.

First we studied the symmetric case with $\rho_1 = \rho_2$ ($r = 0$). Investigating the plot depicting the optimal shape parameter as the function of the total load $\rho = \rho_1 + \rho_2$ we found that it changes between -1 and 1 and that it can be approximated by $c = 2\rho^{2/3} - 1$ very accurately (see Figure 2.6).

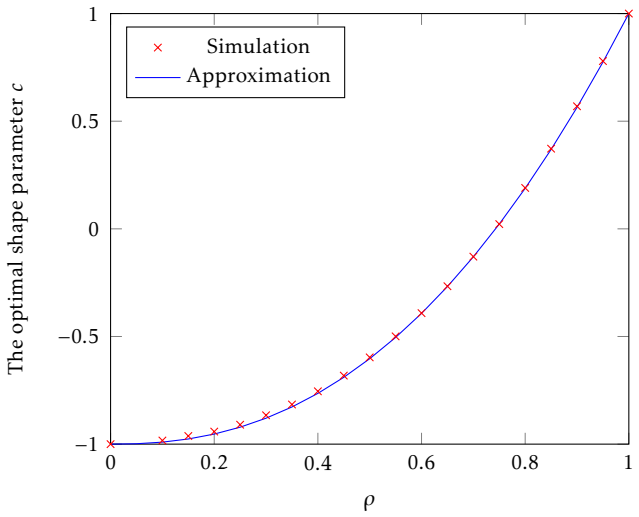


Figure 2.6: The optimal shape parameter as the function of the total load

Hence, we was looking for the approximation in the non-symmetric ($\rho_1 \neq \rho_2$) case in the form of $c = 2\rho^{g(r)} - 1$ as well. The empirical analysis of the exponent revealed that

$$g(r) = 6 \cdot |r - 0.25|^{3.2} + 2 \tag{2.10}$$

is a relatively accurate approximation of the simulation results with less than 5% error (see Figure 2.7), although it gives $g(0) = 2.071$ instead of $2/3$ for $r = 0$.

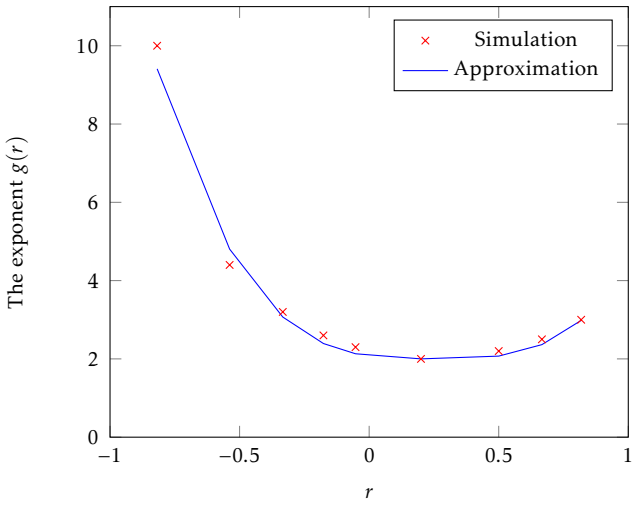


Figure 2.7: The approximation of the exponent $g(r)$

Altogether, the mean scaled response time $\hat{E}(T_1)$ as the function of the scaled weight w_1/w^* is approximated by

$$\hat{E}(T_1) = \frac{w_1/w^* - 1}{(2\rho^{6\cdot|r-0.25|^{3.2}+2} - 1)w_1/w^* - 1}. \tag{2.11}$$

The complete algorithm including the selection of the role of class 1 and the approximation of both mean response times is presented in Algorithm 2. While the resulting formula is explicit, the algorithm brakes down the solution to multiple steps for simplicity.

There is possibility to use another distribution (other than exponential) in our simulation, but the analytical model would be difficult to extend to general inter-arrival time and service time distributions.

2.5 NUMERICAL RESULT

In this section we demonstrate the behavior of our approximation method and compare it with the procedure published in [24]. This comparison is not completely fair, though, since [24] considers a more general system where the inter-arrival and service times can be non-exponential as well.

In general, the proposed approximation managed to achieve very accurate results. In the extreme cases, when the utilization is high and the load is very asymmetric, the accuracy is worse, while in the more "balanced" cases the accuracy is better. Among the scenarios we investigated, the results were the worst with parameters $\rho = 0.95, r = -0.82$. The mean response times as the function of w_1 are depicted in Figure 2.8. The reason of the sub-optimal performance is that under such a high load the inflection point of the curve does not coincide with w^*. However, the results are still much better than the ones obtained by [24].

Figures 2.9 and 2.10 present the typical accuracy of the proposed method. The weights w^*, where the curves cross each other, are captured almost exactly. The approximation of the bend of the curve has some error, but it is much more accurate than the error of [24].

Algorithm 2 The approximation of the mean response time

function $E(T_1), E(T_2) = \text{WFQRESPONSETIME}(\lambda_1, \lambda_2, \mu)$
  $\rho_1 \leftarrow \lambda_1/\mu,\ \rho_2 \leftarrow \lambda_2/\mu,\ \rho \leftarrow \rho_1 + \rho_2$
  $w^* \leftarrow \left(1 - \frac{\rho_1}{\rho_2}\right)\frac{\rho_1+\rho_2-1}{\rho_1-0.42\rho_2-1} + \frac{\rho_1}{\rho_2}$
  if $w_1/w_2 < w^*$ **then**
    $r \leftarrow (\rho_1 - \rho_2)/\rho$
    $c \leftarrow 2\rho^{6\cdot|r-0.25|^{3.2}+2} - 1$
    $w \leftarrow w_1/w_2$
    $E(T_{FCFS}) \leftarrow \frac{1}{\mu_1-\lambda_1-\lambda_2}$
    $E(T_{Prio}) \leftarrow E(T_{FCFS})/(1-\rho_2)$
    $\hat{E}(T_1) \leftarrow (w/w^* - 1)/(c \cdot w/w^* - 1)$
    $E(T_1) \leftarrow \hat{E}(T_1)E(T_{Prio}) + (1-\hat{E}(T_1))E(T_{FCFS})$
    $E(T_2) \leftarrow (\rho E(T_{FCFS}) - \rho_1 E(T_1))/\rho_2$
  else
    $r \leftarrow (\rho_2 - \rho_1)/\rho$
    $c \leftarrow 2\rho^{6\cdot|r-0.25|^{3.2}+2} - 1$
    $w^* \leftarrow 1/w^*$
    $w \leftarrow w_2/w_1$
    $E(T_{FCFS}) \leftarrow \frac{1}{\mu_1-\lambda_1-\lambda_2}$
    $E(T_{Prio}) \leftarrow E(T_{FCFS})/(1-\rho_1)$
    $\hat{E}(T_2) \leftarrow (w/w^* - 1)/(c \cdot w/w^* - 1)$
    $E(T_2) \leftarrow \hat{E}(T_2)E(T_{Prio}) + (1-\hat{E}(T_2))E(T_{FCFS})$
    $E(T_1) \leftarrow (\rho E(T_{FCFS}) - \rho_2 E(T_2))/\rho_1$
  end if
  return $E(T_1), E(T_2)$
end function

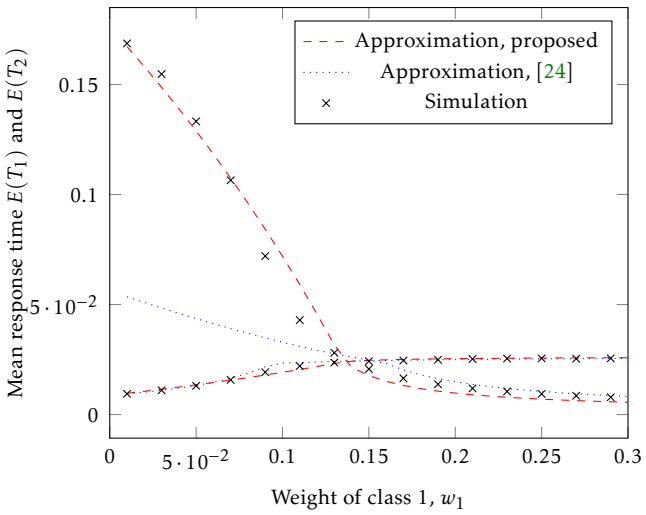


Figure 2.8: The worst results obtained by the approximation, $\rho = 0.95, r = -0.82$ ($\mu = 0.0012$)

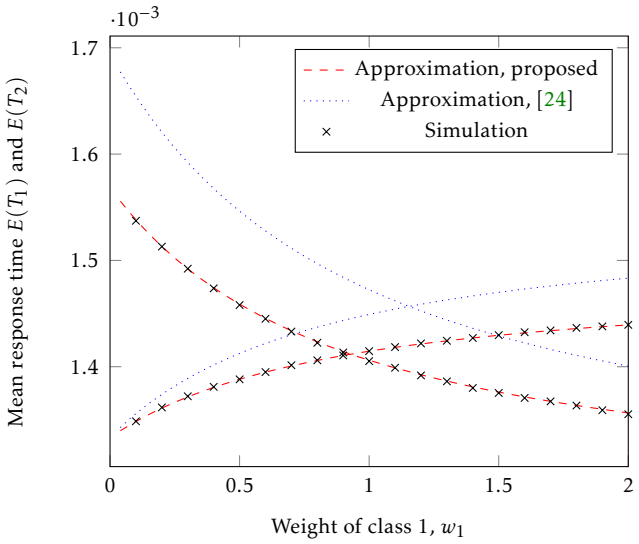


Figure 2.9: Comparison with parameters $\rho = 0.15$ and $r = -0.33$ ($\mu = 0.0012$)

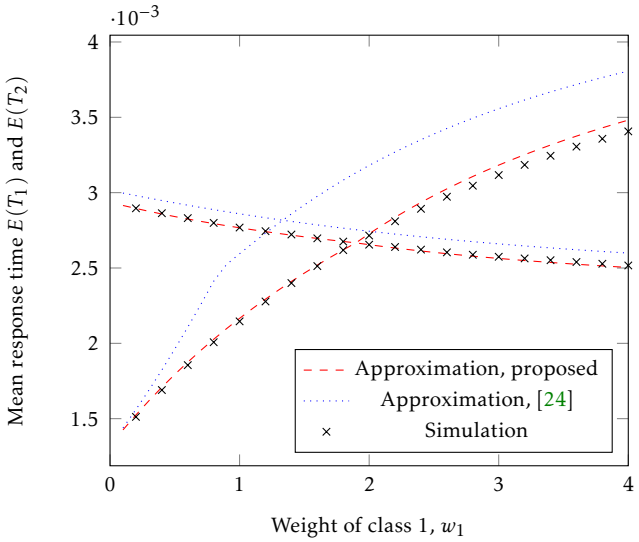


Figure 2.10: Comparison with parameters $\rho = 0.55$ and $r = 0.67$ ($\mu = 0.0012$)

2.6 CONCLUSION AND APPLICATION OF PROPOSED MODEL

2.6.1 *Conclusion*

In this chapter we have presented a simple explicit approximation formula for the mean response times in the two-class weighted fair queueing system. While there are some queueing considerations behind the results, the approximation is mostly based on an algebraic approach. Some decisions on how to approximate the behavior seem sometimes ad-hoc, however the accuracy of the approximation is reasonable, much better than the method found in the literature studying the same system.

2.6.2 *Application of proposed queueing model for WFQ*

The rapid change in telecommunication systems posed many challenges according to the users' requirements. Several decades ago the network applications were much simpler than the networks applications nowadays. These applications have shifted from sending files, e-mails and reading web pages into watching videos, listening audio, playing online games and more. The packet schedulers of the existing networks devices of the Internet such as routers, switches that are parts should take the different requirements of the aforementioned applications into account. Our solution presents a simple approximation formula to evaluate mean response time in the WFQ scheduler to set the weights of the scheduler or to implement adaptive behavior.

CHAPTER THREE

3

ALERT MESSAGE PROPAGATION ON THE HIGHWAY IN VANET WITHOUT RSUS

Intelligent transportation systems are becoming more and more important nowadays. In these systems, vehicles and possibly the infrastructure communicate with each other by vehicular ad-hoc networks. VANETs are being deployed and widely used in urban as well as in highway applications. Several standard use cases have been identified over the last decade (i.e., alert messages, car following support, data exchange between vehicles, etc.). In this chapter, we focus on the alert message propagation on the highway. We derive the stationary and the transient solution of the message propagation distance by constant vehicle speed. Since these messages frequently indicate an accident on the road leading to a traffic jam, we extend the model to take the queueing system due to the traffic jam also into consideration. Our analytical results are compared with SUMO-/Veins-based simulations.

3.1 BACKGROUND

There are two main techniques to obtain performance measures in a VANET systems. The typical approach is based on discrete-event simulation, with SUMO/Omnet++ [34, 62] being the most popular tool for this purpose. Analytical models have several advantages over simulations: they are typically faster to evaluate and provide a better, more direct insight into the behavior of the system; on the other hand, they are much more difficult to develop. In this chapter we study the message passing process, referred to as "propagating information about events" in the list above in Section 1.1. In our scenario an event occurs on the highway (e.g., an accident), due to which alert messages are emitted continuously. When a vehicle receives the message, it starts emitting it, too, and all vehicles in its radio transmission range will receive the information.

There are many factors affecting the efficiency of the message delivery. In this chapter, we assume that the communication is based on the IEEE 802.11p WAVE protocol. In this protocol, the information is carried by various channels defined for different purposes ([21]). In our work, we do not analyze the details of the radio transmission, but assume an idealistic behavior, namely that all messages are successfully transmitted immediately to all vehicles in the transmission range of the sender. We also assume that vehicles traveling in the opposite direction do not contribute to the message delivery ([39], [51]). The message transmission considered in this chapter is one-dimensional, which is a proper model for highways, but not for urban scenarios. While these assumptions seem restrictive, there are surprisingly few results available in the literature for even such a simple model. There are some papers on the stationary analysis of the message propagation distance ([77], [43]), however, according to our best knowledge, there are no results available on the analysis of the transient behavior.

Nevertheless, the analysis of the transient behavior is crucial in many safety scenarios. When the speed of a vehicle changes suddenly, or when an accident occurs, it is essential

to know how fast the related alert message reaches a given distance. Answering this question needs transient analysis, which is the main topic of this chapter.

3.2 RELATED WORK AND USE CASES IN VANET

In the recent decades, a large number of research work appeared related to VANETs, proposing new protocols, studying the properties of the radio channel, analyzing various performance measures, etc. In this section we summarize only those which are closely related to the topic of this chapter, that is the analysis of the message propagation.

The message propagation delay and the number of vehicles receiving the messages depend on many factors. When road-side units (RSUs) are present, the message delivery is more efficient, since RSUs can collaborate to increase the throughput of the network. A cooperative load balancing solution is proposed in [5] that prioritizes requests based on their urgency. However, in this chapter, we assume that no RSUs are present, hence the messages can be passed only between vehicles.

Several published VANET papers investigate the effect of the radio channel and the media access control (MAC) on the message transmission delay. The survey [61] lists the most widely used propagation models (including deterministic and probabilistic models) to analyze the probability of successful message delivery given the parameters of the channel. Like our proposed model, [53] also focuses on the message delivery delays, but it assumes a dense-enough (urban) network such that there is always at least one vehicle present in the transmission range. Our model does not have this restriction. Furthermore, [53] considers the properties of the physical layer and the MAC layer (fading, interference, packet loss, etc.) as the primary source of delays, while our solution builds on a higher level model with idealistic physical and MAC layers. The MAC layer is analyzed in [72], too, where the authors consider a Markov model for the contention window to obtain the packet reception rate taking the potential collisions due to concurrent transmissions and queue overflow into consideration.

Protocol layers above the MAC have been studied in the literature as well. A new broadcast protocol has been introduced in [9] with a new forwarding scheme, that is capable of achieving very high message propagation speed. In ad-hoc networks, the broadcast communication sometimes leads to a so-called storm problem. In paper [65] the authors proposed a new protocol, called DRIVE, based on local one-hop neighbor information that achieves short delays and low overhead.

In this chapter, we investigate the same performance measures as the aforementioned published papers, namely the message propagation distance and the message propagation speed (or the delay). However, we ignore all the effects of the underlying physical/MAC/network layers, which were analyzed in details in the papers listed above. We assume idealistic behavior. Thus, the messages inside a given range are always delivered successfully, and with zero transmission delay. Even with these restrictive assumptions, the exact mathematical analysis is challenging. The following couple of papers pose the same assumptions and analyze the dynamics of the message passing in VANETs with purely mathematical tools.

In [76] a slightly different problem is solved. That model aims to determine the optimal lifetime of the messages, to ensure that all the vehicles in a given zone receive the information. The authors assume a fixed transmission range, Poisson arrival process for the vehicles, and normal distributed speed, but the results are not explicit and are approximate when there are vehicles with different speeds. In my proposed model, the message lifetime is considered to be infinite. Hence, messages are never dropped even if they have not been forwarded for a long time.

The message propagation distance in one-dimension is studied in [43]. Making use of an interesting approach, based on the workload process of a G/D/∞ queue, the probability density function (pdf) of the message propagation distance is derived in Laplace transform domain. From the Laplace transform the mean value is determined, and the probability density function (pdf) in time domain is obtained explicitly by inverse transformation. We study the same problem in our model, by a different approach: we develop a differential equation for the pdf of the message propagation distance. With this approach, we were able to derive a differential equation for the transient behavior as well, which has never been published before.

In [77] the one-dimensional scenario of [43] has been extended to a two-dimensional lattice, for urban environment. This proposed model still considers the stationary behavior only and presents no results on the transient.

3.3 DESCRIPTION OF THE SYSTEM

In this chapter we assume that vehicles enter the highway according to a Poisson process with rate λ. While there are situations when this assumption does not hold, for low to moderate traffic intensity the Poisson process turned out to be a valid model of the traffic [52, 69]. In our model all cars have the same constant speed v. While, based on real traffic measurements [12], the speed of the vehicles is frequently modeled by a normal distribution [69], several research papers make use of the constant speed assumption to make the analysis tractable. This is our main motivation as well. Even with the constant speed assumption the derivations get fairly complicated, yet enable us to draw some interesting conclusions. Making the model more general by allowing different vehicle speed is among our future plans. Under these restrictions we have that the distance between two adjacent vehicles is exponentially distributed with parameter $\vartheta = \lambda / v$. The direction of the vehicles is constant, and it is one-directional as shown in Figure 3.1.

For the radio communication, we assume that all vehicles communicate according to the IEEE 802.11p standard. Several VANET's papers have been published in the past on the analysis of the message reception probabilities as the function of the inter-vehicle distance at radio level. The models developed can be classified into deterministic or probabilistic models [61]. We use the deterministic propagation model in this chapter (free space propagation model, also referred to as Friis model, [19]), that defines a coverage radius R, and assumes that all messages are successfully transmitted to all vehicles inside the coverage radius. This model is used in many VANET simulation frameworks, including the Simulation of Urban MObility (SUMO, [34]) and Vehicles in Network Simulation (Veins, [57]).

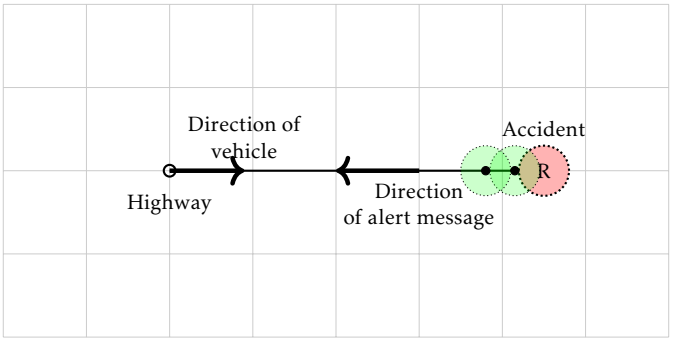


Figure 3.1: Propagation alert meaage in the Highway

For a similar problem, the stationary message propagation distance has been derived both in [77], [43]. We arrive at the same results in Sections 3.3.1 and 3.3.2. Our new contribution, the transient analysis is described in Section 3.3.3.

3.3.1 *Clusters of informed vehicles*

A set of vehicles, where the distance between subsequent vehicles in the set is less than R, is called a *cluster* in this chapter. When the head (the first vehicle) of the cluster receives a message, all the vehicles in the cluster will receive it eventually. In our model, we assume that all vehicles will be informed immediately, although in reality, the speed of the information propagation depends on the beacon periods, too. (Later, Section 3.3.4 studies the consequences of this assumption.)

Let us introduce random variable $\mathcal{G}$, referred to as the *cluster length*, that plays an important role in the analysis. $\mathcal{G}$ represents the distance between the position of the first vehicle in the cluster and the last position where the alert information is available, that is, the position of the last vehicle plus R (see Figure 3.2). In the rest of the section the complementary cumulative distribution function (ccdf) of $\mathcal{G}$, denoted by $G(x) = P(\mathcal{G} > x)$, is derived.

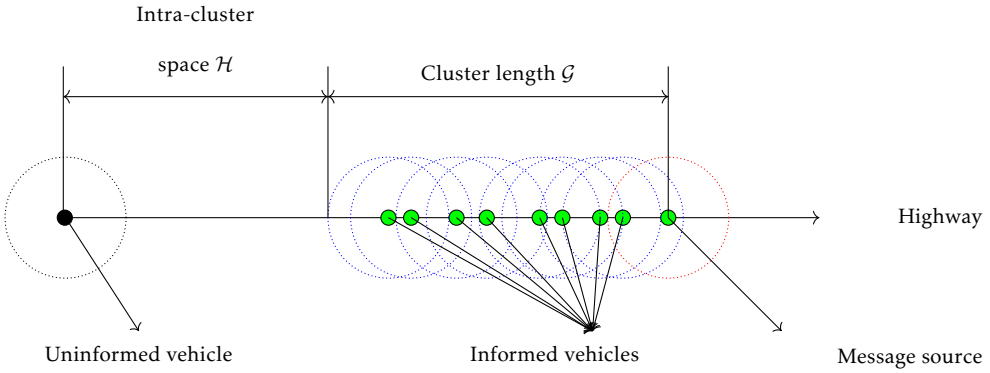


Figure 3.2: A cluster of informed vehicles

From the stochastic interpretation of the system, $G(x)$ satisfies the recursive expression

$$G(x) = \begin{cases} 1, & \text{if } x \le R, \\ \int_{y=0}^{R} \vartheta e^{-\vartheta y} G(x-y)\, dy, & \text{if } x > R. \end{cases} \tag{3.1}$$

The first term corresponds to the case when x is close enough to the message source to receive the message directly. In the second case the vehicle that is the closest to the message source falls into the coverage area (in distance y); it receives the message and starts broadcasting it, hence the message has to take only the remaining $x-y$ distance to reach the target.

The next theorem provides the mean of $\mathcal{G}$ based on this recursion.

Theorem 1. *The mean cluster length $E(\mathcal{G})$ can be expressed by*

$$E(\mathcal{G}) = \frac{1}{\vartheta}(e^{\vartheta R} - 1). \tag{3.2}$$

Proof. To obtain the mean value the integral of the ccdf is calculated:

$$\begin{aligned} E(\mathcal{G}) &= \int_{x=0}^{\infty} G(x)\, dx \\ &= \int_{0}^{R} 1\, dx + \int_{R}^{\infty} \int_{0}^{R} \vartheta e^{-\vartheta y} G(x-y)\, dy\, dx \\ &= R + \int_{0}^{R} \vartheta e^{-\vartheta y} \left[\int_{y}^{\infty} G(x-y)\, dx - \int_{y}^{R} G(x-y)\, dx \right] dy \\ &= R + (1 - e^{-\vartheta R}) E(\mathcal{G}) - R(1 - e^{-\vartheta R}) \\ &\quad + (1 - e^{-\vartheta R})/\vartheta - R e^{-\vartheta R}, \end{aligned}$$

from which (3.2) follows. □

The differential equation providing the ccdf $G(x)$ itself is given by the following theorem.

Theorem 2. *The ccdf of $\mathcal{G}$ is the solution of the delayed differential equation (DDE)*

$$\frac{d}{dx} G(x) = -\vartheta e^{-\vartheta R} G(x-R), \quad x > R, \tag{3.3}$$

with boundary condition $G(x) = 1, x \le R$.

Proof. For $G(x+R)$ the second case of (3.1) is valid (since $x > 0$), that, by swapping the variables in the integral, can be transformed to

$$\begin{aligned} G(x+R) &= \int_{y=0}^{R} \vartheta e^{-\vartheta y} G(x+R-y)\, dy \\ &= \int_{u=x}^{x+R} \vartheta e^{-\vartheta(x+R-u)} G(u)\, du, \end{aligned} \tag{3.4}$$

which, after taking the derivative and making use of the Leibniz integral rule, leads to

$$\frac{d}{dx}G(x+R) = \vartheta G(x+R) - \vartheta e^{-\vartheta R}G(x)$$
$$\underbrace{-\vartheta \int_{x}^{x+R} \vartheta e^{-\vartheta(x+R-u)}G(u)du}_{G(x+R)} = -\vartheta e^{-\vartheta R}G(x),$$

which establishes the theorem. □

3.3.2 *The stationary solution of $\mathcal{D}(t)$*

The stationary solution $\mathcal{D} = \lim_{t\to\infty}\mathcal{D}(t)$ is the distance between the event, A, and the position where the information is available, as seen by an external observer at a random point of time. With other words, when an external observer takes a snapshot of the system, $\mathcal{D}$ represents the distance the message can travel through a chain of vehicles being closer than R to each other, starting in A.

Observe that this quantity is almost the same as the cluster length $\mathcal{G}$ defined in Section 3.3.1. The only slight difference is that in this case, the head of the cluster is the source at A instead of a vehicle, which does not need different mathematical treatment. Hence we have that $\mathcal{D}$ is in fact equal to $\mathcal{G}$, thus $E(\mathcal{D}) = E(\mathcal{G})$ and $F(x) = P(\mathcal{D} > x) = G(x)$ hold.

3.3.3 *The transient analysis of $\mathcal{D}(t)$*

After the analysis of the stationary solution we also investigate the transient behavior of the process $\mathcal{D}(t)$, that is, we characterize the ccdf $F(t,x) = P(\mathcal{D}(t) > x)$ for $x \geq R$. Note that this distribution has probability mass at $x = R$, $\mathcal{D}(t) = R$ occurs when there are no vehicles on the highway having the message received.

Theorem 3. *The transient ccdf $F(t,x)$ satisfies the partial differential equation (PDE) for $x > R$*

$$\frac{\partial}{\partial t}F(t,x) - v\frac{\partial}{\partial x}F(t,x) = \lambda(1 - F(t,R_+))G(x-R), \tag{3.5}$$

where $F(t,R_+)$ denotes $\lim_{x\searrow R}F(t,x)$. For $x \leq R$ we have $F(t,x) = 1$.

Proof. To prove the theorem we describe the evolution of $\mathcal{D}(t)$ in the infinitesimally small time period $(t, t+\Delta)$. There are two possibilities leading to $\mathcal{D}(t+\Delta) > x, x > R$:

- either the position of the last informed vehicle was beyond $x + \Delta \cdot v$ at time t,
- or there was no informed vehicle present at time t (hence $\mathcal{D}(t) = R$, having probability $1 - F(t,R_+)$), a new vehicle entered the range R of the event (with probability $\lambda\Delta$), received the message, and through a chain of vehicles the information eventually reached beyond distance $x - R$ (the corresponding probability is $G(x-R)$).

All the remaining events have probability $o(\Delta)$, for which $\lim_{\Delta\to 0} o(\Delta)/\Delta = 0$ holds. More formally,

$$\begin{aligned} F(t+\Delta,x) = {} & F(t,x+\Delta\cdot v) \\ & + (1-F(t,R_+))\cdot\lambda\Delta\cdot G(x-R) \\ & + o(\Delta), \end{aligned} \tag{3.6}$$

which, after dividing by Δ, some transformation, and taking the limit $\Delta \to 0$ leads to (3.5).

□

To show that $G(x)$ (and Theorem 2) is the stationary solution of $F(t,x)$ indeed, we take the limit $t \to \infty$ and substitute $G(x) = \lim_{t\to\infty} F(t,x)$, leading to

$$\frac{d}{dx}G(x) = -\vartheta(1-G(R_+))G(x-R), \tag{3.7}$$

since $\lambda/v = \vartheta$. Taking the limit $x \to R$ in (3.3) and exploiting that $G(x-R)$ becomes 1 in this range yields $G(R_+) = 1-e^{-\vartheta R}$, that, together with (3.7), gives the same DDE as shown in Theorem 2.

To actually compute the transient probabilities the differential equation (3.5) has to be evaluated numerically. (3.6) provides one possibility for this purpose, with an appropriately small Δ it enables to obtain the probabilities progressively.

An other, particularly interesting quantity to study is the evolution of the mean value $E(\mathcal{D}(t))$ in time. The mean value is given by the integral of the ccdf, hence

$$E(\mathcal{D}(t)) = \int_0^\infty F(t,x)\,dx = R + \int_R^\infty F(t,x)\,dx.$$

Integrating both sides of (3.5) with regards to x from R to ∞ gives

$$\frac{d}{dt}(E(\mathcal{D}(t))-R) - v\Big[F(t,x)\Big]_R^\infty = \lambda(1-F(t,R_+))E(\mathcal{D}),$$

that, making use of (3.2), simplifies to

$$\frac{d}{dt}E(\mathcal{D}(t)) = v(e^{\vartheta R}(1-F(t,R_+))-1). \tag{3.8}$$

Similar to the mean value, the variance $Var(\mathcal{D}(t))$ can be calculated similarly.

3.3.4 *Numerical examples*

In the first example the mean message propagation distance $E(\mathcal{D})$ is depicted as the function of the vehicle arrival rate λ and radio transmission range R (see Figure 3.3). The speed of the vehicles is set to $v = 36$ m/s throughout this section. The results are as expected: the higher the vehicle density and the wider the transmission range is, the longer is the message propagation distance. Figure 3.3 confirms that the radio coverage plays a crucial role in the message propagation, increasing R to twice the value increases the message propagation distance by several orders of magnitudes.

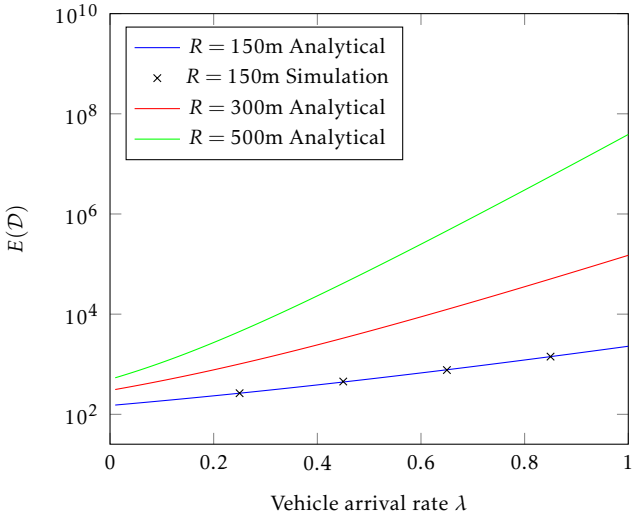


Figure 3.3: The mean message propagation distance

Next, we plot the ccdf $F(x)$ by $\lambda = 0.65$ and various R parameters in Figure 3.4.

To investigate the effect of the simplifying assumptions making the scenario idealistic (including the infinitely large acceleration and deceleration, infinitesimally small beacon period), we developed various simulation tools in the SUMO/Veins/OMNeT++ framework [34, 57, 62]). These simulation tools are

- An idealized simulation tool respecting the same assumptions as our analytical models. The only purpose of this simulator is to verify our analytical results. Since the performance measures obtained by this simulator matched our formulas up to a very small error, we do not include the corresponding results on our plots.
- The realistic simulation where the radio communications are realistic. The message is transferred by beacons, with the beacon period set to 50ms. The behavior of the vehicles (including the acceleration, deceleration, car-following model, etc.) is also realistic, for the parameters see Table 3.1.
- A realistic simulation where the vehicles have different speed, according to a normal distribution with variance $\sigma^2 = 8$.

In Figure 3.4 the results obtained by the detailed simulations and the analytical results are compared, by parameters $R = 150, \lambda = 0.65$. As reflected by the plots, assuming immediate message transmission between the vehicles and assuming constant vehicle speed have little impact on the shape of the ccdf.

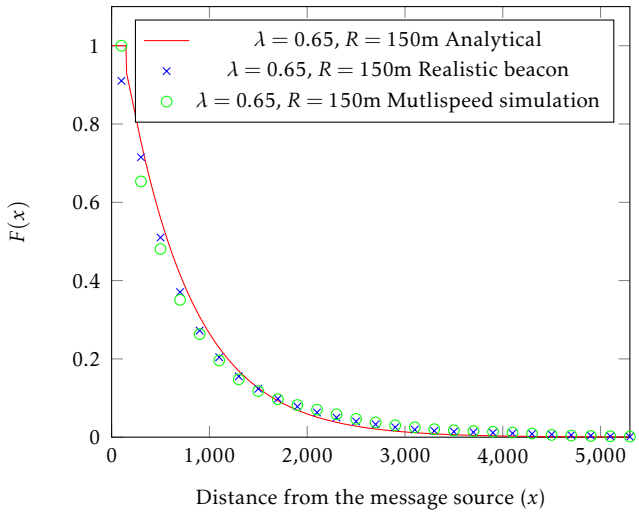


Figure 3.4: The comparison of the analytical and simulation-based results

Finally, we demonstrate some studies that can be carried out using the results for the transient behavior. we note that the transient behavior is especially difficult to investigate efficiently by simulation: to compute the message propagation probabilities at a certain time t a huge number of independent simulations must be performed up to time t with different random seeds. Our analytical formulas, on the other hand, returned the results quickly without any numerical issues. Figure 3.5 depicts the transient ccdf $F(t,x)$ at $t = \{1,2,4\}$ seconds together with the stationary solution $F(x)$, while Figure 3.6 shows the mean of the message propagation distance $E(\mathcal{D}(t))$ and its squared coefficient of variation (SCV) defined by $Var(\mathcal{D}(t))/E^2(\mathcal{D}(t))$. According to the figures, by such high vehicle density, the stationary solution is achieved very fast, in just 4 seconds. The results are obtained analytically, and are verified by simulation as well, all the results are based on constant speed assumption. The investigation of message propagation is much more challenging in the multi-speed case. We developed a multi-speed simulation tool that can consider realistic parameters, and the extension of the analytical model in this direction is our future plan.

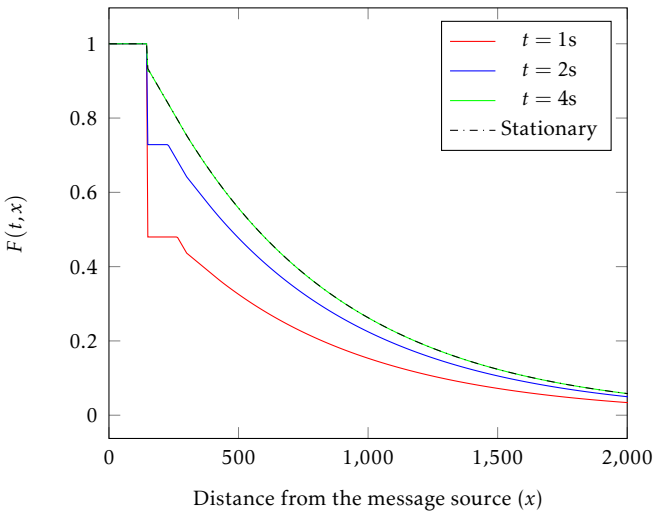


Figure 3.5: The transient distribution $F(t,x)$ at some time points, $\vartheta = 0.65$, $R = 150$m

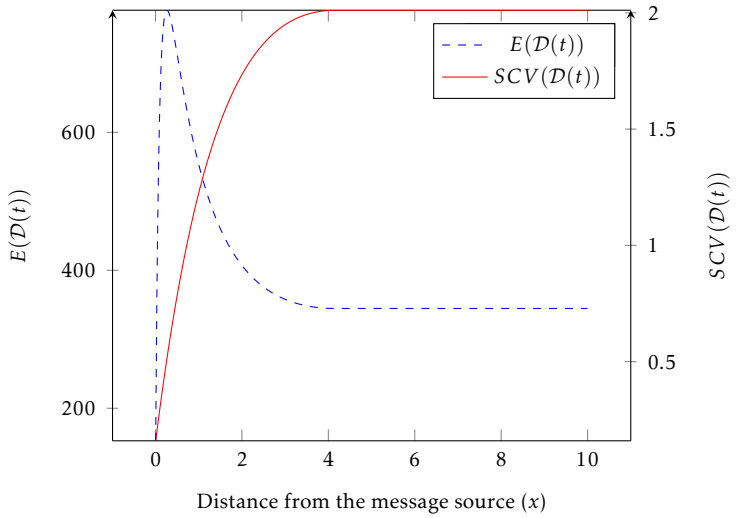


Figure 3.6: The mean and the SCV of the message propagation distance as the function of time

3.4 QUEUEING MODEL FOR THE TRAFFIC JAM

Based on the results of the previous section it is possible to calculate the distribution of the message propagation distance when there is an accident on the highway at position A. However, there is the other factor that has to be taken into consideration. Typically, the accident can be passed by at a very low speed, leading to a traffic jam. To take the effect of the traffic jam into account, we introduce a queueing model. The arrival process to this queue is a Poisson process with rate λ, and the service time is given by the time taken by a vehicle to leave the accident. If the vehicle length is L and the speed in the accident region is v_{slow}, the service time is deterministic, $\tau = L/v_{slow}$, leading to an M/D/1 (infinite capacity) queue.

3.4.1 *Stationary solution of the queue*

To obtain the stationary solution of the queue length $\mathcal{X}$, the queue length process embedded at vehicle service instants is studied. Since the arrival process is Poisson, the stationary solution of the embedded process is equal to the stationary solution seen by a random observer due to the PASTA property [35].

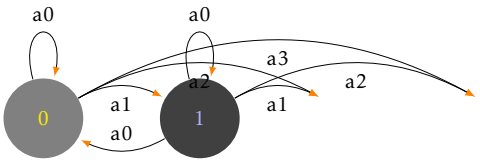


Figure 3.7: State-transition diagram for M/D/1 Markov chain

The transition probability matrix P of the embedded Markov chain has an upper-Hessenberg structure ([35], see also Figure 3.7). Hence we have

$$P = \begin{bmatrix} a_0 & a_1 & a_2 & a_3 & \dots \\ a_0 & a_1 & a_2 & a_3 & \dots \\ & a_0 & a_1 & a_2 & \dots \\ & & a_0 & a_1 & \dots \\ & & & \ddots & \vdots \end{bmatrix}. \tag{3.9}$$

In matrix P, a_i gives the probability that i vehicle arrive to the accident region while one vehicle leaves it. It can be computed as

$$a_i = \frac{(\lambda D)^i}{i!} e^{-\lambda D}, \quad i \geq 0. \tag{3.10}$$

If the queue is stable ($\lambda D < 1$ holds), the stationary queue length probabilities $\pi_i = P(\mathcal{X} = i), i \geq 0$ can be obtained by the solution of

$$\pi P = \pi, \quad \sum_{n=1}^{\infty} \pi_n = 1, \tag{3.11}$$

where π is the (infinitely long) row vector consisting of probabilities π_i. To compute the queue length distribution numerically, we have $\pi_0 = 1 - \lambda D$ and for $i > 0$ the π_i is given by the recursive formula

$$\pi_{i+1} = \frac{\pi_i - \pi_0 a_i - \sum_{j=1}^{i} \pi_j a_{i-j+1}}{a_0}. \tag{3.12}$$

For the mean queue length the Pollaczek-Khinchine formula provides a simple explicit solution

$$E(\mathcal{X}) = \rho + \frac{1}{2} \frac{\rho^2}{1-\rho}. \tag{3.13}$$

3.4.2 *Message propagation distance*

The message propagation distance in the presence of the queue, denoted by $\mathcal{B}(t)$, has two components: the queue length at time t, given by $\mathcal{X}(t)$, and the message propagation distance counted from the end of the queue given by $\mathcal{D}(t)$. These two random variables are not independent, but in this section, we approximate $\mathcal{B}(t)$ as if the two components were independent. Hence, in steady state, for the mean values we have

$$E(\mathcal{B}) = E(\mathcal{X}) + E(\mathcal{D}), \tag{3.14}$$

where $E(\mathcal{X})$ is given by (3.13) and $E(\mathcal{D}) = E(\mathcal{G})$ is given by (3.2).

For the ccdf of $\mathcal{B}$, defined by $B(x) = P(\mathcal{B} > x)$, our approximation is given by

$$B(x) = \sum_{i=0}^{\infty} \pi_i \cdot F(x - i \cdot L), \tag{3.15}$$

since the length of i cars in the traffic jam is $i \cdot L$.

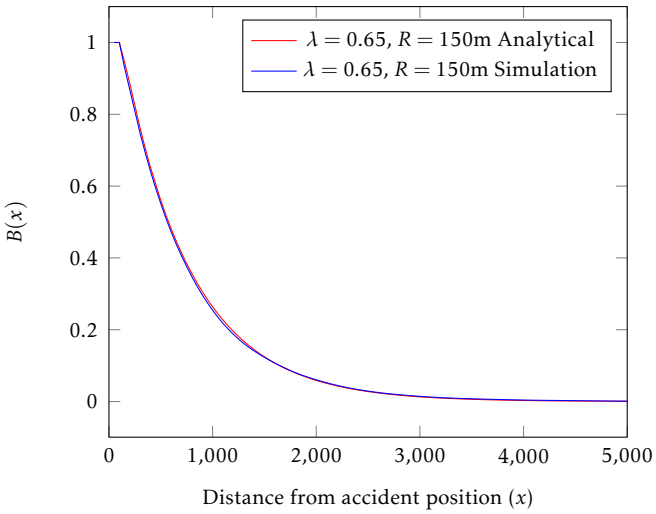


Figure 3.8: The comparison of the analytical and simulation-based results

3.4.3 *Simulation results*

In this section we prepare a simulation study to evaluate the accuracy of the results. Our SUMO/Veins/OMNeT++-based simulations are detailed, in the sense that the messages are transferred through beacons, and the behavior of the radio transmission, as well as the motion model of the vehicles, is taken into account in full details (for the parameters see Table 3.1).

Table 3.1: Parameters used in simulation

Parameter	Definition
R	The radius of the radio transmission coverage, $R = 150$m
L	Length of the vehicles, $L = 4.5$m
v	Normal vehicle speed, $v = 36$m/s
v_{slow}	Slow vehicle speed at accident, $v_{slow} = 3$m/s
λ	Car arrival rate, $\lambda = 0.65$
Center frequency	5.890 GHz
Analogue Model	Simple path loss model
Safety message duplicate period	50ms
Channel bandwidth	10MHz
Channel data rate	3Mbps
Safety message packet size	100 bytes
Simulation time	8000s

We assume that the vehicles slow down from $v = 36$m/s to $v_{slow} = 3$m/s when they reach the accident. Since the car length is set to $L = 4.5$m and the car arrival rate is $\lambda = 0.65$cars/s, this means that the utilization of the queue representing the traffic jam is $\rho = 0.975$.

The comparison between the analytical and the simulation results is depicted in Figure 3.8. According to the results, the simple model proposed to approximate the effect of the traffic jam turned out to be reasonably accurate.

3.5 SUMMARY

In this chapter, we have introduced new contributions regarding the message propagation in VANET systems. We have derived the stationary and transient solutions of the message propagation distance. We validated our analytical results with simulation (Veins and SUMO within OMNET++). we have proposed an accurate approximation to take into account the length of the traffic jam caused by an accident as well.

CHAPTER FOUR

4

ALERT MESSAGE PROPAGATION SPEED IN VANET WITH DISCONNECTED RSUS

In Chapter 3 we presented the transient analysis of the alert message propagation process in the case when there are no RSUs available on the highway. We derived the probability that the message is available beyond a certain distance from the message source after a given amount of time. In this chapter we consider a similar system, but now RSUs are also available, thus V2I we is also enabled. However, this kind of extension of the system needs a different analysis approach. In this system we have identified a Markov renewal process, which characterizes the message transfer between subsequent RSUs. By the analysis of this Markov renewal process, we derived the message propagation speed, thus, starting from the message source (possibly an accident), how fast the alert message travels back on the highway. Our results make it possible to evaluate how efficient the message propagation is by a given RSU distance, which is an important information when the optimal RSU placement is decided while planning the network. We also present the transient analysis of the message passing process among the RSUs, which is essential in determining how quickly help arrives to the accident.Since the speed of the vehicles affects the message propagation speed, our results can be useful when predicting the effect of speed restrictions, too.

4.1 RELATED WORK

The stochastic properties of the connected clusters of (close enough) vehicles were investigated in many papers in the past. In [43] the mean length of such clusters was derived, when only V2V communication is possible, while [31, 69, 70], and with a generic radio channel model, [46], study the connectivity of such networks. According to [12] the connectivity can be full in some times of the day when the vehicle density is high enough, on the other hand, it has been recognized that message propagation is not efficient when the vehicles are sparse; alert messages arrive with big delays when the connectivity between the vehicles is not good enough [48].

As we mentioned in the introduction, there are two types of communication in VANET, V2V and V2I, and in some network conditions V2V may not be enough to provide a satisfactory network performance. Therefore, to increase the quality of communication, road-side units (RSUs) can also be deployed, enabling V2I we communication, too. There are several analytical results available regarding VANETs with disconnected RSUs. In [1] the scenario investigated is similar to ours, but the performance measure studied there is different, bounds are provided for the message delivery delay with the effective bandwidth theory. In [49] again the connectivity is studied in a VANET both with interconnected and with disconnected RSUs.

One of the most important research topic in this field, which needs mathematical models, is the optimal placement of the RSUs, which has a big impact on the connectivity of the vehicles [6, 18, 38, 42, 49]. When RSUs are close to each other, the

deployment cost is high, on the other hand, long RSU distance decreases the speed of the message propagation. The RSU placement problem is considered in [30] based on a graph-theoretic approach, in a complex road network with junctions. The performance measure to optimize in that paper is the end-to-end delay of the messages. In [37] both 1-D (highway-like) and 2-D vehicular networks are considered, and an RSU placement algorithm is presented to minimize the delay and the power consumption. However, both of these papers aim to cover a certain area with radio coverage, and ignore the fact that vehicles themselves are able to forward messages. V2V and V2I communication are both allowed in [46], this paper focuses on the connectivity probability (the probability that an RSU can be reached either directly or indirectly, through several vehicles), and does not consider the message propagation speed.

The delay and the optimal placement problems of RSUs have led many researchers to develop new strategies or mathematical models to reduce the number of RSUs along the highway and to increase the communication efficiency. The scenario considered in [40] is very similar to ours, this paper studies the total delay of broadcasting alert messages in VANETs along highways. However, RSUs are connected in that paper, which makes the message passing process much easier to analyze. Both V2V and V2I communications are taken into account in [71] , where the RSU placement problem is formulated as an integer linear programming problem. This chapter considers forward message propagation, while in case of alert messages the backward direction is more important, since vehicles must be warned about the accident before they arrive there.

To the best of our knowledge, there are no papers published in the literature on the analysis of the alert message propagation speed in the presence of RSUs.

4.2 MODEL DEFINITION AND ASSUMPTIONS

In this chapter we still consider Poisson vehicle arrival process, with arrival rate λ, the vehicle speed is constant with speed v, thus the distance between two adjacent vehicles is exponentially distributed with parameter $\vartheta = \lambda/v$. The radio coverage is deterministic with radius R (which is the same for all vehicles), and messages are immediately delivered among vehicles closer than R to each other.

In this chapter we assume that road-side units (RSUs) are deployed along the highway in an equidistant manner, with the distance between any two adjacent RSUs denoted by D. The radio coverage of the RSUs is denoted by $\hat{R}$, for which $\hat{R} \geq R$ holds. The RSUs are disconnected, which means that there is no direct way of communication between them [48].

During the analysis we ignore the effect of the counter-flow traffic, but at the end of the chapter we show that with a proper RSU placement the messages are passed much faster than the vehicles of the counter-flow traffic.

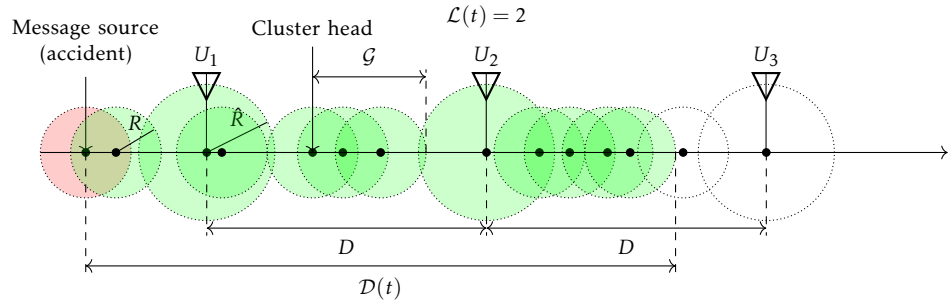


Figure 4.1: A snapshot of the highway at time t; green color denotes the informed elements

Our main goal in this chapter is the analysis of the speed of alert message propagation on the highway. The alert messages are passed among the vehicles and RSUs, when they are close enough to each other. The questions of interest are as follows:

- The asymptotic message propagation speed, C, defined by

$$C = \lim_{t\to\infty} \frac{E(\mathcal{D}(t))}{t}, \tag{4.1}$$

  thus, how fast the information propagates back on the highway. This performance measure is investigated in Section 4.4.

- The transient characterization of the message propagation process. If random variable $\mathcal{L}(t)$ represents the index of the farthest RSU having the message received at time t, then our aim is to obtain the probability $P(\mathcal{L}(t) = i)$. The corresponding results are presented in Section 4.5.

4.3 MESSAGE PROPAGATION SPEED WITHOUT RSUS

The analysis of the message propagation without the presence of RSUs has been discussed in Chapter 3. We have derived the ccdf of the cluster length $G(x) = P(\mathcal{G} > x)$ in (3.1) and the mean cluster length $E(\mathcal{G})$ in (3.2).

Since the speed is constant, and there are no RSUs, there can be only a single informed cluster on the highway. The dynamics of the system is simple: there are periods in time when only the static message source is aware of the information (the duration of these intervals is represented by exponentially distributed random variable $\mathcal{H}$). As soon as a vehicle enters the radio coverage of the message source, the cluster (of size $\mathcal{G}$) initiated by that vehicle gets informed immediately. Then, this message propagation distance starts decreasing with speed v, and when the last vehicle of the cluster leaves the message source, there will be no informed vehicles on the highway again. The trajectory of the message propagation distance, $\mathcal{D}(t)$ is depicted by Figure 4.2.

The transient analysis of the underlying stochastic process is presented by [41], characterizing $F(t,x) = P(\mathcal{D}(t) > x)$. The information propagation speed itself, however, is not considered in that paper. Nevertheless, the lack of RSUs enables the derivation of it in a simple way. For this system we have that

$$C = \lim_{t\to\infty} \frac{E(\mathcal{G})}{t} = 0, \tag{4.2}$$

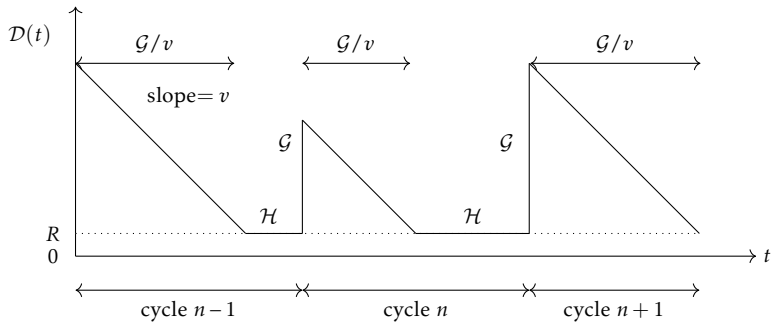


Figure 4.2: Trajectory of the information distance $\mathcal{D}(t)$

where we used the fact that $\lim_{t\to\infty} F(t,x) = \lim_{t\to\infty} P(\mathcal{D}(t) > x) = G(x)$, according to [41], hence, $\lim_{t\to\infty} E(\mathcal{D}(t)) = E(\mathcal{G})$. This result is as expected, since the current informed cluster always has to leave the accident before the next cluster can get informed. In the next sections we show that the presence of RSUs leads to essentially different results.

4.4 MESSAGE PROPAGATION SPEED IN THE PRESENCE OF RSUS

When there are RSUs on the highway as well, a long enough cluster of vehicles can bridge the distance between them, carrying the message from RSU to RSU farther and farther away from the message source. The analysis of the message propagation is based on the observation that the stochastic process $\mathcal{L}(t) \in \{0,1,\dots\}, t \geq 0$, representing the index of the farthest informed RSU at time t (see Figure 4.1), is a Markov renewal process.

4.4.1 $\mathcal{L}(t)$ *as a Markov renewal process*

In order to derive the asymptotic message propagation speed, we first identify renewal time instants in the evolution of $\mathcal{D}(t)$, where the process becomes memory-less. There are several such instants, the analysis presented in this chapter is based on the time points where the farthest informed RSU, $\mathcal{L}(t)$, changes. These time points are denoted by $\mathcal{Q}_k, k = 1,\dots$, thus $\mathcal{Q}_k = t : \mathcal{L}(t) > \mathcal{L}(t-\Delta)$ for infinitesimally small Δ. It is easy to see that time points $\mathcal{Q}_k$ are renewal instants: when an RSU receives the message, it acts as a new message source on the highway.

Figure 4.3 depicts a sample trajectory of $\mathcal{D}(t)$, where the positions of the RSUs are denoted by $U_k, k = 1,\dots$ (to make the figure more compact the distance between the RSUs is not equal here, as opposed to our modeling assumption). Whenever a vehicle enters the coverage of an RSU, a cluster is formed with length $\mathcal{G}'$. If the cluster is long enough, the message can propagate through several RSUs, marking the next renewal time instant. The distance between the RSU that was the farthest informed previously and the RSU that got the farthest informed at the kth renewal point is denoted $\mathcal{S}_k$, hence we have $\mathcal{S}_k = U_{\mathcal{L}(\mathcal{Q}_k)} - U_{\mathcal{L}(\mathcal{Q}_{k-1})}$.

Some remarks:

- Random variables $\mathcal{G}$ and $\mathcal{G}'$ are different, since $\mathcal{G}$ measures the length of clusters consisting of vehicles only, while in $\mathcal{G}'$ the RSUs themselves contribute to the clusters, too.
- Random variables $\mathcal{T}_k = \mathcal{Q}_{k+1} - \mathcal{Q}_k, k = 1,\dots$ (the time between renewal instants) and $\mathcal{S}_k, k = 1,\dots$ are independent and identically distributed, therefore their subscripts are neglected, they are referred to as $\mathcal{T}$ and $\mathcal{S}$.
- We assume equidistant RSU placement, hence $U_k = k \cdot D$ throughout the chapter.
- Following from the definition of $\mathcal{S}$ and the equidistant placement of RSUs, $\mathcal{S}$ is an integer multiple of D.

To make the derivations simpler, we assume that there is an RSU at U_0 and that the accident occurs at a renewal point. Since the message propagation speed is an asymptotic quantity, these assumptions do not affect the results. In Section 4.5, where the transient behavior is studied, the position of the accident is important, and it will be properly taken into account.

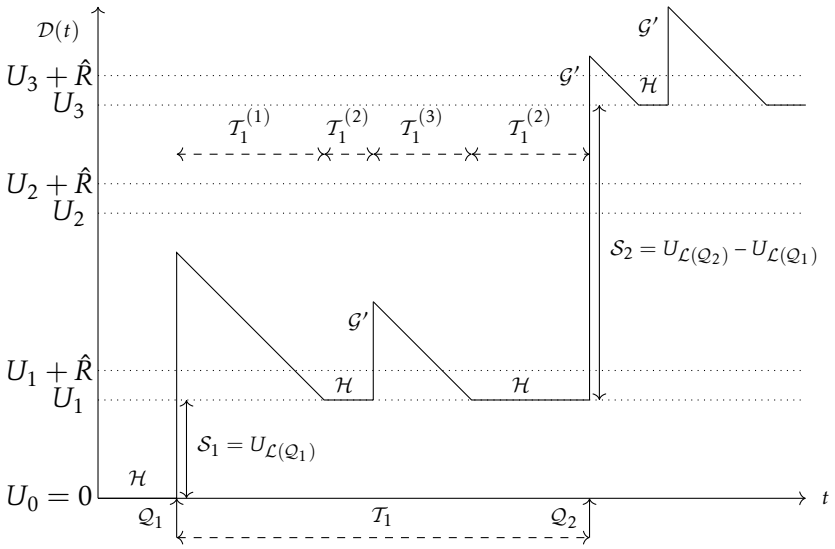


Figure 4.3: The evolution of the information distance $\mathcal{D}(t)$

The following lemma relates the information propagation speed C and the mean values of $\mathcal{T}$ and $\mathcal{S}$.

Lemma 4. *The information propagation speed C can be computed by*

$$C = \frac{E(\mathcal{S})}{E(\mathcal{T})}. \tag{4.3}$$

Proof. Let us assume that there are K renewal instants till time t, thus $K = \max\{k : \sum_{i=1}^{k} \mathcal{T}_i < t\}$. In this case $\mathcal{D}(t)/t$ can be expressed by

$$\frac{\mathcal{D}(t)}{t} = \frac{\mathcal{S}_1 + \mathcal{S}_2 + \dots + \mathcal{S}_K + c_S}{\mathcal{Q}_1 + \mathcal{T}_1 + \mathcal{T}_2 + \dots + \mathcal{T}_K + c_T}, \tag{4.4}$$

where $\mathcal{Q}_1$ is the time till the first renewal instant, and the terms c_S and c_T are the distance and the time taken since the last renewal point. Observe that c_S is bounded since $c_S < D$ and c_T and $\mathcal{Q}_1$ have finite mean values. Taking the limit leads to

$$\begin{aligned} C = \lim_{t\to\infty} \frac{\mathcal{D}(t)}{t} &= \lim_{K\to\infty} \frac{\sum_{k=1}^{K} \mathcal{S}_k + c_S}{\mathcal{Q}_1 + \sum_{k=1}^{K} \mathcal{T}_k + c_T} \\ &= \lim_{K\to\infty} \frac{\frac{1}{K}\sum_{k=1}^{K} \mathcal{S}_k + \frac{c_S}{K}}{\frac{\mathcal{Q}_1}{K} + \frac{1}{K}\sum_{k=1}^{K} \mathcal{T}_k + \frac{c_T}{K}} = \frac{E(\mathcal{S})}{E(\mathcal{T})}, \end{aligned} \tag{4.5}$$

that establishes the lemma. □

4.4.2 *Message propagation when $\hat{R} = R$*

In the rest of the section we derive $E(\mathcal{S})$ and $E(\mathcal{T})$ in order to compute C, in the case when the radio coverage of the RSUs equals the radio coverage of the vehicles. This assumption will be relaxed later in Section 4.4.3.

The next theorem provides $E(\mathcal{S})$, which is the mean distance between the farthest informed RSUs when the subsequent vehicles form a long enough cluster and transfer the message to an RSU farther away on the highway (see Figure 4.3).

Theorem 5. *The mean distance between the farthest informed RSUs at two subsequent renewal instants, $E(\mathcal{S})$, can be obtained by*

$$E(\mathcal{S}) = \frac{D}{1 - G(D)}, \tag{4.6}$$

where $G(x)$ is the ccdf of the cluster length given by (3.1).

Proof. A renewal instant occurs when, after an idle period of the farthest informed RSU, a new vehicle arrives leading with a cluster long enough to reach the next RSU. If the current RSU is located at position U_{k-1}, the head of the new cluster enters the range of the RSU at position $U_{k-1} + R$, hence the minimal cluster length reaching the next RSU is $U_k - (U_{k-1} + R) = D - R$ (see Figure 4.4), consequently, the probability of this event is $G(D - R)$.

The question is, given that the cluster is long enough, implying that a renewal instant occurs, how far does the message get, how many RSUs receive the information at this renewal instant. When this long enough cluster of vehicles reaches the next RSU, the next RSU at U_k acts as a cluster head itself. If this cluster (initiated by the RSU) is longer than D (with probability $G(D)$), the next RSU gets informed, too, initiating a further cluster, etc. Hence, the number of RSUs receiving the information is geometrically distributed, i RSUs get informed with probability $G(D)^{i-1}(1 - G(D))$. For any two subsequent renewal time instants Q_1 and Q_2 we have

$$\begin{aligned} E(\mathcal{S}) &= \sum_{i=1}^{\infty} i \cdot D \cdot P(\mathcal{L}(\mathcal{Q}_2) - \mathcal{L}(\mathcal{Q}_1) = i) \\ &= \sum_{i=1}^{\infty} i \cdot D \cdot G(D)^{i-1}(1 - G(D)) = \frac{D}{1 - G(D)}, \end{aligned}$$

that proves 4.6. □

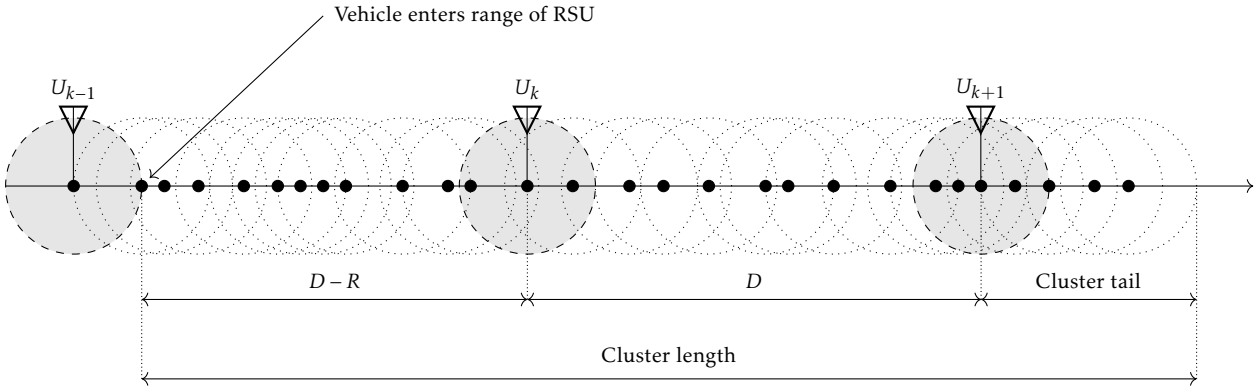


Figure 4.4: Vehicle arrival after a regenerative point, $\hat{R} = R$

Our next target towards obtaining C is the derivation of $E(\mathcal{T})$. Before the corresponding theorem we need the following lemma enabling the efficient computation of the finite integrals of $G(x)$.

Lemma 6. *The finite integral of $G(x)$, defined by $H(x) = \int_{y=0}^{x} G(y)\,dy$, can be calculated as*

$$H(x) = \frac{1}{\vartheta}\left(e^{\vartheta R} - 1 - e^{\vartheta R} G(x+R)\right), \tag{4.7}$$

for $x > R$.

Proof. Let us integrate (3.1) from 0 to x, for $x > R$. We get

$$\begin{aligned} H(x) &= \int_{z=0}^{R} 1\,dz + \int_{z=R}^{x}\int_{y=0}^{R} \vartheta e^{-\vartheta y} G(z-y)\,dy\,dz \\ &= R + \int_{y=0}^{R} \vartheta e^{-\vartheta y} \underbrace{\int_{z=y}^{x} G(z-y)\,dz}_{H(x-y)}\,dy - \int_{y=0}^{R} \vartheta e^{-\vartheta y} \underbrace{\int_{z=y}^{R} G(z-y)\,dz}_{R-y}\,dy \\ &= \frac{1}{\vartheta}(1 - e^{-\vartheta R}) + \int_{y=0}^{R} \vartheta e^{-\vartheta y} H(x-y)\,dy. \end{aligned} \tag{4.8}$$

Now we show that (4.7) satisfies this equation. Inserting (4.7) into the left-hand side (LHS) of (4.8) leads to

$$\frac{1}{\vartheta}\left(e^{\vartheta R} - 1 - e^{\vartheta R} G(x+R)\right). \tag{4.9}$$

Inserting (4.7) into the right-hand side (RHS) of (4.8) gives

$$\frac{1}{\vartheta}(1 - e^{-\vartheta R}) + \frac{1}{\vartheta}\underbrace{\int_{y=0}^{R} \vartheta e^{-\vartheta y}(e^{\vartheta R} - 1)\,dy}_{e^{-\vartheta R}+e^{\vartheta R}-2} - \frac{1}{\vartheta} e^{\vartheta R} \underbrace{\int_{y=0}^{R} \vartheta e^{-\vartheta y} G(x-y+R)\,dy}_{G(x+R)}. \tag{4.10}$$

The LHS and the RHS are matching, proving the lemma. □

Observe that from (4.7) and (3.2) we have $\lim_{x\to\infty} H(x) = E(G)$, as expected. The mean time between two renewal instants, $E(\mathcal{T})$, is provided by the following theorem.

Theorem 7. *The mean time between two renewal instants can be expressed by*

$$E(\mathcal{T}) = \frac{\frac{1}{\vartheta} + H(D-R)}{G(D-R)v} + \frac{H(D)-D}{(1-G(D))v}. \tag{4.11}$$

Proof. Let us denote the probability density function of the cluster length $\mathcal{G}$ by $g(x) = \lim_{\Delta\to 0} \frac{1}{\Delta} P(\mathcal{G} \in (x, x+\Delta))$, for $x > R$. Random variable $\mathcal{G}$ has a probability mass at point R as well (when the cluster consists of a single element only), thus $P(\mathcal{G} = R) = 1 - G(R)$.

Using these notations we can express the three main components of $\mathcal{T}$, denoted by $\mathcal{T}^{(1)}, \mathcal{T}^{(2)}$ and $\mathcal{T}^{(3)}$ (see Figure 4.3), as follows:

- The first component, $\mathcal{T}^{(1)}$ is the departure time of the cluster tail (Figure 4.4) after the renewal instant. By definition, renewal instants are those time instants when the message gets transferred to farther RSUs. Right after the renewal instant t, the length of the cluster of informed vehicles, measured from RSU $\mathcal{L}(t)$ (and including this RSU), is $\mathcal{D}(t) - U_{\mathcal{L}(t)}$, hence the time till these vehicles leave the RSU (when $\mathcal{D}(t)$ becomes $U_{\mathcal{L}(t)} + R$) is $\mathcal{T}^{(1)} = (\mathcal{D}(t) - U_{\mathcal{L}(t)} - R)/v$. The mean value of this component can be obtained by

$$E(\mathcal{T}^{(1)}) = \frac{(1-G(R))\cdot 0}{1-G(D)} + \frac{\int_{x=R}^{D} g(x)\frac{x-R}{v}\,dx}{1-G(D)}, \tag{4.12}$$

  where the first term corresponds to the case when the length of the cluster tail is exactly R (no vehicles are informed beyond the RSU after the renewal instant), implying $\mathcal{T}^{(1)} = 0$. The second term covers the case when the length of the cluster tail is $R < x < D$. When the length of the cluster tail (including the RSU) is x, the time till it leaves the RSU is $(x-R)/v$, since this is the time till the cluster length reduces to R, meaning that only the RSU is informed, and no vehicles beyond it. Note that the cluster tail can not be longer than D, since in that case the next RSU would receive the message that contradicts the definition of the cluster tail.

- The second component of $\mathcal{T}$ is the duration of the periods when no vehicles are informed beyond the RSU. These periods are exponentially distributed with parameter λ, thus

$$E(\mathcal{T}^{(2)}) = \frac{1}{\lambda}. \tag{4.13}$$

- When there are no informed vehicles beyond the RSU, and a new vehicle enters the range of the RSU, a new cluster gets informed. The position of the first vehicle in this cluster is $U_{\mathcal{L}(t)} + R$. If the length of the cluster, x, is less than $D - R$, then the message does not reach the next RSU, thus no renewal instant occurs. In this case the time till this cluster leaves the RSU, x/v, contributes to $\mathcal{T}$. For the mean value of this component, that is the mean departure time of clusters too short to reach the next RSU, we have

$$E(\mathcal{T}^{(3)}) = \frac{(1-G(R))\frac{R}{v}}{1-G(D-R)} + \frac{\int_{x=R}^{D-R} g(x)\frac{x}{v}\,dx}{1-G(D-R)}. \tag{4.14}$$

According to the first term the new cluster consists of a single vehicle only; the case when there are more vehicles in the cluster are covered by the second term.

Every renewal period $\mathcal{T}$ starts with component $\mathcal{T}^{(1)}$, followed by $\mathcal{T}^{(2)}$. Then, components $\mathcal{T}^{(3)}$ and $\mathcal{T}^{(2)}$ are repeated as many times as the arriving cluster length turned out to be too short to reach the next RSU, the probability of which is $1-G(D-R)$. When after a $\mathcal{T}^{(2)}$ component a long enough cluster arrives to the RSU, with probability $G(D-R)$, a new renewal instant occurs. We get

$$E(\mathcal{T}) = E(\mathcal{T}^{(1)}) + E(\mathcal{T}^{(2)}) + \underbrace{\sum_{i=0}^{\infty}(1-G(D-R))^i G(D-R)\cdot i}_{\frac{1-G(D-R)}{G(D-R)}} \cdot (E(\mathcal{T}^{(3)}) + E(\mathcal{T}^{(2)})). \tag{4.15}$$

Substituting (4.12), (4.13) and (4.14) into (4.15) and making use of the relations

$$\int_{x=R}^{D} g(x)\,dx = G(R) - G(D), \tag{4.16}$$

$$\int_{x=R}^{D} x\,g(x)\,dx = H(D) - D\,G(D) - R(1-G(R)), \tag{4.17}$$

we get

$$\begin{aligned} E(\mathcal{T}) &= \frac{\int_{x=R}^{D} g(x)\frac{x-R}{v}\,dx}{1-G(D)} + \frac{1}{\lambda} + \frac{(1-G(R))\frac{R}{v}}{G(D-R)} + \frac{\int_{x=R}^{D-R} g(x)\frac{x}{v}\,dx}{G(D-R)} + \frac{1-G(D-R)}{G(D-R)}\frac{1}{\lambda} \\ &= \frac{H(D)-D}{v(1-G(D))} + \frac{D-R}{v} + \frac{1}{\lambda} + \frac{1}{vG(D-R)}(H(D-R) - (D-R)G(D-R)) \\ &\quad + \frac{1-G(D-R)}{G(D-R)}\frac{1}{\lambda} \\ &= \frac{H(D)-D}{v(1-G(D))} + \frac{(D-R)G(D-R)}{vG(D-R)} + \frac{\frac{v}{\lambda}}{vG(D-R)} \\ &\quad + \frac{H(D-R)-(D-R)G(D-R)}{vG(D-R)}, \end{aligned} \tag{4.18}$$

which equals (4.12). □

Finally, the main result, the message propagation speed is formulated by the following corollary.

Corollary 1. *The message propagation speed can be calculated by*

$$C = \frac{E(\mathcal{S})}{E(\mathcal{T})} = v \cdot \frac{D}{\left(\frac{1}{\vartheta} + H(D-R)\right)\frac{1-G(D)}{G(D-R)} + H(D) - D}. \tag{4.19}$$

4.4.3 *Message propagation when $\hat{R} > R$*

In the practice, the radio coverage of the RSUs is greater than the one of the vehicles, thus $\hat{R} > R$. In this section we generalize the results obtained above to this, more realistic scenario.

Theorem 8. *When $\hat{R} > R$ holds, $E(\mathcal{S})$ is computed by*

$$E(\mathcal{S}) = \frac{D}{1 - G(D - 2\hat{R} + 2R)}. \tag{4.20}$$

Proof. The proof is similar to the one of theorem 5. When a vehicle enters the range of the farthest informed RSU (located at U_{k-1}) after the renewal instant, the minimal cluster length to reach the coverage of the next RSU is $D - 2\hat{R} + R$ (see Figure 4.5). If the cluster is longer than that (with probability $G(D - 2\hat{R} + R)$), the next RSU receives the message, a renewal instant occurs, and we have to find out how many RSUs will get the message at this instant.

At this instant, after the RSU at U_k received the message, it starts behaving as a cluster head and the cluster formed by the vehicles close enough to each other can be long enough to reach the next RSU at (U_{k+1}) as well. The end of this cluster, initiated by the RSU at U_k, must be at least $U_{k+1} - \hat{R} + R$ to pass the message to the next RSU (the position of the last vehicle of the cluster should be greater than $U_{k+1} - \hat{R}$). However, the ccdf $G(x)$ characterizing the cluster length we have introduced in (3.1) assumes that the radio coverage of all cluster elements are the same, it can not take the different radio coverage of the RSU into consideration.

To overcome this technical difficulty, we can observe that the RSU at U_k having radius $\hat{R}$ can be replaced by an RSU at $U_k + \hat{R} - R$ having radius R, the distribution of the end position of the cluster generated by them will be exactly the same. Consequently, in this equivalent system the cluster starts at $U_k + \hat{R} - R$, it must end beyond $U_{k+1} - \hat{R} + R$, implying that the length must be at least $(U_{k+1} - \hat{R} + R) - (U_k + \hat{R} - R) = D - 2\hat{R} + 2R$, hence the probability of passing the message to the next RSU is $G(D - 2\hat{R} + 2R)$.

As in the proof of theorem 5, we have

$$\begin{aligned} E(\mathcal{S}) &= \sum_{i=1}^{\infty} i \cdot D \cdot P(\mathcal{L}(\mathcal{Q}_2) - \mathcal{L}(\mathcal{Q}_1) = i) \\ &= \sum_{i=1}^{\infty} i \cdot D \cdot G(D - 2\hat{R} + 2R)^{i-1}(1 - G(D - 2\hat{R} + 2R)) = \frac{D}{1 - G(D - 2\hat{R} + 2R)}, \end{aligned}$$

that equals 4.20. □

Theorem 9. *When $\hat{R} > R$ holds, $E(\mathcal{T})$ is computed by*

$$E(\mathcal{T}) = \frac{\frac{1}{\vartheta} + H(D - 2\hat{R} + R)}{G(D - 2\hat{R} + R)v} + \frac{H(D - 2\hat{R} + 2R) - D + 2\hat{R} - 2R}{(1 - G(D - 2\hat{R} + 2R))v}. \tag{4.21}$$

Proof. The proof follows the same steps as the one of Theorem 7. The mean values of the three components of the time between two renewal instants, $\mathcal{T}$, depicted in Figure 5.4, are as follows.

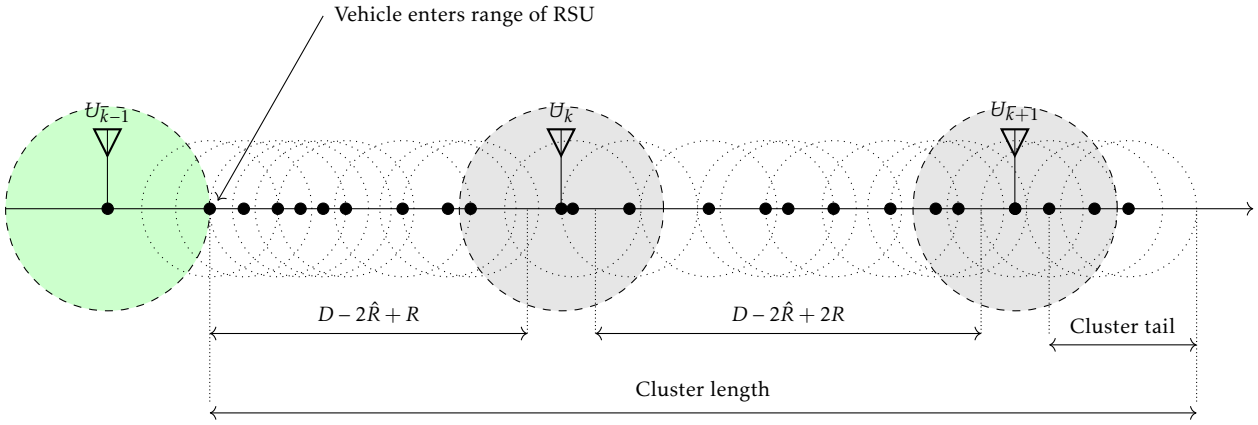


Figure 4.5: Vehicle arrival after a regenerative point, $\hat{R} > R$

- The mean cluster tail departure time can be obtained as

$$E(\mathcal{T}^{(1)}) = \frac{\int_{x=R}^{D-2\hat{R}+2R} g(x)\frac{x-R}{v}\,dx}{1-G(D-2\hat{R}+2R)}. \tag{4.22}$$

  The idea behind this formula is the same as we applied in the proof of theorem 8, hence we replace the RSU at U_k with radius $\hat{R}$ with an RSU at $U_k+\hat{R}-R$ with radius R. Since the cluster tail after a renewal instant can not reach the coverage of the next RSU, the cluster must end before $U_{k+1}-\hat{R}+R$, hence the cluster length must be less than $(U_{k+1}-\hat{R}+R)-(U_k+\hat{R}-R) = D-2\hat{R}+2R$.

- The duration of the periods when there are no informed vehicles beyond the RSU is the same as in case of $\hat{R} = R$, thus

$$E(\mathcal{T}^{(2)}) = \frac{1}{\lambda}. \tag{4.23}$$

- The departure time of the clusters that are too short to reach the next RSU is

$$E(\mathcal{T}^{(3)}) = \frac{(1-G(R))\frac{R}{v}}{1-G(D-2\hat{R}+R)} + \frac{\int_{x=R}^{D-2\hat{R}+R} g(x)\frac{x}{v}\,dx}{1-G(D-2\hat{R}+R)}. \tag{4.24}$$

In this formula, the first term corresponds to the case when the cluster consists of a single vehicle only, and the second term covers the case when the cluster is longer. The maximum cluster length that does not reach the next RSU is $D-2\hat{R}+R$ (see the proof of Theorem 8). As in the proof of Theorem 7, the time between renewal instants can be obtained by combining these components, taking $E(\mathcal{T}^{(3)})+E(\mathcal{T}^{(2)})$ as many times the cluster of the entering vehicle was too short to transfer the message to the next RSU, thus

$$\begin{aligned} E(\mathcal{T}) &= E(\mathcal{T}^{(1)}) + E(\mathcal{T}^{(2)}) \\ &+ \sum_{i=0}^{\infty} (1-G(D-2\hat{R}+R))^i G(D-2\hat{R}+R)\cdot i\cdot (E(\mathcal{T}^{(3)})+E(\mathcal{T}^{(2)})), \end{aligned} \tag{4.25}$$

that, after substitution and algebraic manipulations leads to the theorem. □

From $E(\mathcal{S})$ and $E(\mathcal{T})$, it is easy to express the speed of the information propagation for the $\hat{R} \geq R$ case as well.

Corollary 2. *For $\hat{R} \geq R$ the message propagation speed can be calculated by*

$$C = \frac{E(\mathcal{S})}{E(\mathcal{T})} = v \times \frac{D}{\left(\frac{1}{\vartheta} + H(D-2\hat{R}+R)\right)\frac{1-G(D-2\hat{R}+2R)}{G(D-2\hat{R}+R)} + H(D-2\hat{R}+2R) - D + 2\hat{R} - 2R}. \tag{4.26}$$

4.5 TRANSIENT ANALYSIS OF THE MESSAGE PROPAGATION

A potential application of VANETs is the propagation of alert messages containing emergency information related to highway accidents. To quantify how efficiently the RSUs and the vehicles disseminate the information, it is important to study the transient properties of the message propagation, thus the probability that the message is available beyond a given RSU after a certain amount of time. In this section we are going to characterize the distribution

$$L_i(t) = P(\mathcal{L}(t) = i | \mathcal{L}(0) = 0), \tag{4.27}$$

through its Laplace transform

$$L_i^*(s) = \int_{t=0}^{\infty} L_i(t) e^{-st}\, dt, \tag{4.28}$$

and its double transform

$$L^*(s,z) = \sum_{i=0}^{\infty} z^i \int_{t=0}^{\infty} L_i(t) e^{-st}\, dt. \tag{4.29}$$

In these quantities we implicitly assume that $t = 0$ is a renewal point. We will relax this restriction at the end of the section.

First we derive the Laplace-Stieltjes transform (LST) of $\mathcal{T}$, the time between two renewal intervals.

Theorem 10. *The Laplace-Stieltjes transform $f^*(s) = \int_{t=0}^{\infty} e^{-st} dP(\mathcal{T} < t)$ can be obtained by*

$$f^*(s) = \frac{\lambda G(D-2\hat{R}+2R) f_1^*(s)}{s + \lambda - \lambda(1 - G(D-2\hat{R}+2R)) f_3^*(s)}, \tag{4.30}$$

where $f_1^*(s)$ and $f_3^*(s)$ are given by (4.31) and (4.33), respectively.

Proof. As before, in the proof of Theorem 9, we are going to characterize the three components of $\mathcal{T}$ in LST domain separately, then we combine them to get $f^*(s)$.

- The LST of the cluster tail departure time is

$$f_1^*(s) = E(e^{-s\mathcal{T}^{(1)}}) = \frac{1}{1-G(D-2\hat{R}+2R)} \times \left(\int_{x=R}^{D-2\hat{R}+2R} g(x) e^{-s\frac{x-R}{v}}\, dx + 1 - G(R) \right). \tag{4.31}$$

  Note that in case the length of the cluster tail is exactly R (no vehicles are informed beyond the RSU after the renewal instant), the departure time of the cluster tail is exactly 0, whose Laplace transform is $(1-G(R)) \cdot e^{-0t} = 1 - G(R)$, providing the second term in the parenthesis.

- The duration of the "idle" periods, when there are no informed vehicles beyond the RSU, are exponentially distributed with parameter λ, hence

$$f_2^*(s) = E(e^{-s\mathcal{T}^{(2)}}) = \frac{\lambda}{s+\lambda}. \tag{4.32}$$

- The departure time of an arriving cluster that turned out to be too short to reach the next RSU is

$$f_3^*(s) = E(e^{-s\mathcal{T}^{(3)}}) = \frac{1}{1-G(D-2\hat{R}+R)} \times \left((1-G(R)) e^{-s\frac{R}{v}} + \int_{x=R}^{D-2\hat{R}+R} g(x) e^{-s\frac{x}{v}}\, dx \right). \tag{4.33}$$

The renewal intervals always start with a $\mathcal{T}^{(1)}$ and a $\mathcal{T}^{(2)}$ component, and then they have as many $\mathcal{T}^{(3)}$ and $\mathcal{T}^{(2)}$ components as many times the arriving cluster was too short to reach the next RSU. A long enough cluster arriving after a $\mathcal{T}^{(2)}$ component triggers a renewal instant, ending the renewal period. As we have shown before, the number of short clusters is geometrically distributed with parameter $G(D-2\hat{R}+2R)$, thus we have that

$$\begin{aligned} f^*(s) &= \sum_{i=0}^{\infty} (1-G(D-2\hat{R}+2R))^i G(D-2\hat{R}+2R) f_1^*(s) f_2^*(s) (f_3^*(s) f_2^*(s))^i \\ &= \frac{G(D-2\hat{R}+2R) f_1^*(s) f_2^*}{1-(1-G(D-2\hat{R}+2R)) f_3^*(s) f_2^*(s)}, \end{aligned} \tag{4.34}$$

from which the theorem follows. □

We have exploited that the probability density function (pdf) of the sum of random variables has a product form in Laplace transform domain. The following theorem and the corresponding corollary provides the Laplace transform and the double transform of the transient distribution.

Theorem 11. *The Laplace transform of the transient distribution $L_i^*(s)$ satisfies the following recursion*

$$L_i^*(s) = \begin{cases} \sum_{k=1}^{i} q_k f^*(s) L_{i-k}^*(s), & \text{for } i > 0, \\ \frac{1}{s}(1-f^*(s)), & \text{for } i = 0, \end{cases} \tag{4.35}$$

where q_k is the probability that k RSUs receive the message at a renewal instant,

$$q_k = G(D-2\hat{R}+2R)^{k-1}(1-G(D-2\hat{R}+2R)). \tag{4.36}$$

Proof. Assuming that $t=0$ is a renewal instant and conditioning on the time till the next renewal instant leads to

$$P(\mathcal{L}(t)=i|\mathcal{T}_1=\tau) = \begin{cases} \delta_{i,0}, & t<\tau, \\ \sum_{k=1}^{i} q_k L_{i-k}(t-\tau), & t>\tau. \end{cases} \tag{4.37}$$

In the first case, when t is before the renewal instant τ, only RSU 0 has the message, as represented by the Kronecker delta, for which $\delta_{i,0}=1$ holds if $i=0$ and it is $\delta_{i,0}=0$ otherwise. In the second case there was a renewal instant at time τ when k RSUs received the message, thus, in the remaining time $t-\tau$, $k-i$ further RSUs need to receive the message. The probabilities q_k follow from the proof of Theorem 8.

After deconditioning (4.37) with the pdf of $\mathcal{T}_1$, $f(t)$, we get

$$L_i(t) = \delta_{i,0}\int_{\tau=t}^{\infty} f(\tau)\,d\tau + \sum_{k=1}^{i} q_k \int_{\tau=0}^{t} f(\tau)L_{i-k}(t-\tau)\,d\tau, \tag{4.38}$$

which, after Laplace transformation, provides (4.35). □

Corollary 3. *The double tranform of the transient distribution, $L^*(s,z)$, is explicitly given by*

$$L^*(s,z) = \frac{(1-zG(D-2\hat{R}+2R))(1-f^*(s))}{s(1-zG(D-2\hat{R}+2R))-szf^*(s)(1-G(D-2\hat{R}+2R)))}. \tag{4.39}$$

Proof. Applying the definition of $L^*(s,z)$ on (4.35) yields

$$\begin{aligned} L^*(s,z) &= \sum_{i=0}^{\infty} z^i L_i^*(s) = \frac{1}{s}(1-f^*(s)) + \sum_{i=1}^{\infty} z^i \sum_{k=1}^{i} q_k f^*(s) L_{i-k}^*(s) \\ &= \frac{1}{s}(1-f^*(s)) + f^*(s)\sum_{k=1}^{\infty} z^k q_k \sum_{i=k}^{\infty} z^{i-k} L_{i-k}^*(s) \\ &= \frac{1}{s}(1-f^*(s)) + f^*(s)q(z)L^*(s,z), \end{aligned} \tag{4.40}$$

where $q(z)$ is the probability generating function of q_k,

$$\begin{aligned} q(z) &= \sum_{k=1}^{\infty} z^k G(D-2\hat{R}+2R)^{k-1}(1-G(D-2\hat{R}+2R)) \\ &= \frac{z-zG(D-2\hat{R}+2R)}{1-zG(D-2\hat{R}+2R)}. \end{aligned} \tag{4.41}$$

Expressing $L^*(s,z)$ from (4.40) gives (4.39). □

Typically, $t = 0$ is not a renewal point, for example because there are no vehicles in the range of the highway accident initially, or because the distance between the accident and the first RSU is not D. In this case we have to treat the time interval till the first renewal point differently, which leads to a slightly different result for the double transform (see (4.40)),

$$\hat{L}^*(s,z) = \frac{1}{s}(1-\hat{f}^*(s)) + q(z)\hat{f}^*(s)L^*(s,z), \tag{4.42}$$

where the quantities denoted by a hat correspond to the initial time interval that do not start with a renewal point. Here, the special nature of the first interval is captured solely by $\hat{f}^*(s)$. For instance, if there are no vehicles in the range of the accident, the distance to the first RSU is $D' > R$, and the radio coverage of the message source is R, $\hat{f}^*(s)$ can be calculated by

$$\hat{f}^*(s) = \frac{\lambda G(D')}{s + \lambda - \lambda(1-G(D'))\hat{f}_3^*(s)}, \tag{4.43}$$

where

$$\hat{f}_3^*(s) = \frac{1}{1-G(D'-R)}\left((1-G(R))e^{-s\frac{R}{v}} + \int_{x=R}^{D'-R} g(x)e^{-s\frac{x}{v}}\,dx\right), \tag{4.44}$$

based on the proof of Theorem 10. In general, the proof of Theorem 10 provides guidance to derive the time till the first renewal point $\hat{f}^*(s)$, from which the double transform of the transient distribution is given by (4.42).

4.6 NUMERICAL EXAMPLES

In this section, we investigate the alert message dissemination in VANET in the presence of RSUs in different scenarios. In the first scenario the radio range of the RSU and the vehicles are the same ($\hat{R} = R$), while in the second scenario the radio ranges are different. The implementation of the analytical method is based on Matlab. To verify the results, we also developed a simulation tool in C++ language. This simulation tool models the system at the same level of details as the analytical model. It is based on the discretization of the time axis, it updates the position of the vehicles every 0.001 seconds. We stopped the simulation after considering 2 million vehicles. Decreasing the time step or increasing the number of vehicles further did not change the results significantly. Nevertheless, we note that our results are exact, the only purpose of the simulation is to validate the correctness of our Matlab implementation. The analytical and simulation methods gave essentially the same results in all of the different scenarios, as shown by the plots in Section 4.6.1 and Section 4.6.2, where the lines correspond to the analytical, and the marks to the simulation results. We note that in many cases the simulation is very slow, while the analytical method gives prompt results.

In all of the numerical examples, the vehicle speed is assumed to be $v = 36m/s$, which is a common speed limitation for highways in several countries.

4.6.1 *RSU and vehicle have the same radio range* $\hat{R} = R$

In this subsection the transmission range of both the RSUs and the vehicles is set to $\hat{R} = R = 150m$. In each forthcoming figure in this section we present two plots. On the left the radio coverage is fixed to $R = 150m$, and different arrival rate values are considered, $\lambda \in \{0.45, 0.65, 0.95\}$. On the right the arrival rate is fixed to $\lambda = 0.65$, and two different radio coverages are investigated, $R \in \{150m, 250m\}$.

Figure 4.6 depicts the mean time between two renewal instants $E(\mathcal{T})$ (the mean time the message is passed to RSUs farther away), according to Theorem 7. As expected, increasing the RSU distance D increases $E(\mathcal{T})$ as well. At higher traffic rate the clusters are longer, thus the message can reach the next RSUs sooner. The radio range of the vehicles and RSUs, R has a big impact on $E(\mathcal{T})$, as shown on the right side of Figure 4.6.

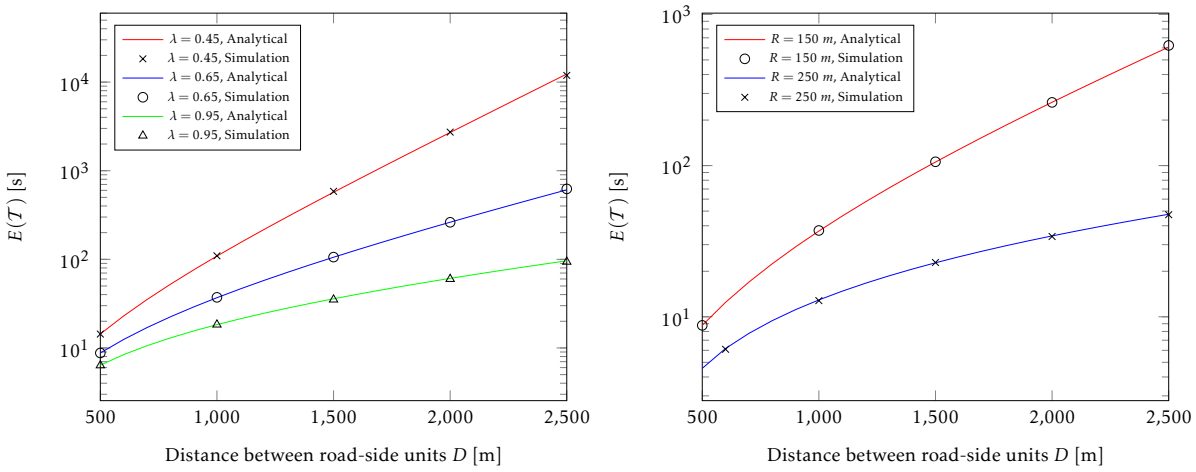


Figure 4.6: The mean time between renewal instants, $E(\mathcal{T})$

Figure 4.7 depicts $E(\mathcal{S})$, the mean distance between the RSU that was the farthest informed previously and the RSU that got the farthest after getting informed at a renewal instant. According to Figure 4.7 the relation between D and $E(\mathcal{S})$ is not monotonous. If D is small, more RSUs can get the message when a long enough cluster arrives. If D is big, a lot of time is needed to transfer the message from an RSU to an other one, but once it occurs, the distance the message moved forward is greater. Increasing transmission range R can increase the mean distance as shown on the right side of Figure 4.7. Let take the following example: if D is small, then every time the information jumps to a further RSU there is a high chance that many further RSUs will get it at the same time, since RSUs are close to each other, only a few cars are enough to bridge the gap between them. If D is big, it takes much longer to bridge the gap till the next RSU, but when it happens, the message travels at least D, which is big as shown in the right hand side of Figure 4.7

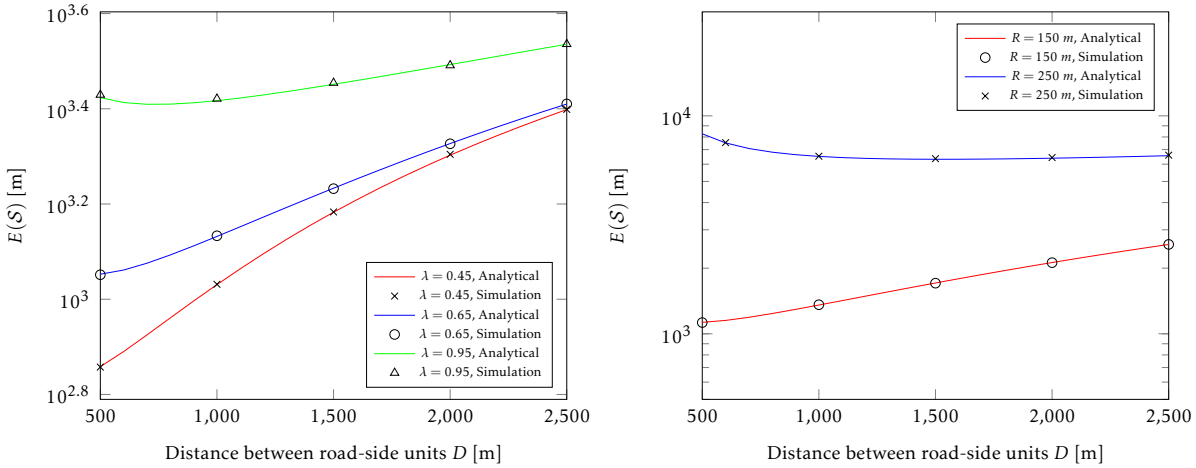


Figure 4.7: The mean distance between the farthest informed RSUs at the renewal instant, $E(\mathcal{S})$

The ratio of $E(\mathcal{S})$ and $E(\mathcal{T})$ gives the speed C, depicted in Figure 4.8. Greater radio coverage and greater traffic intensity leads to higher message propagation speed. Observe that at $R = 250$ and with appropriate RSU distance the message propagation speed can be very high, much higher than the speed of counter-flow vehicles.

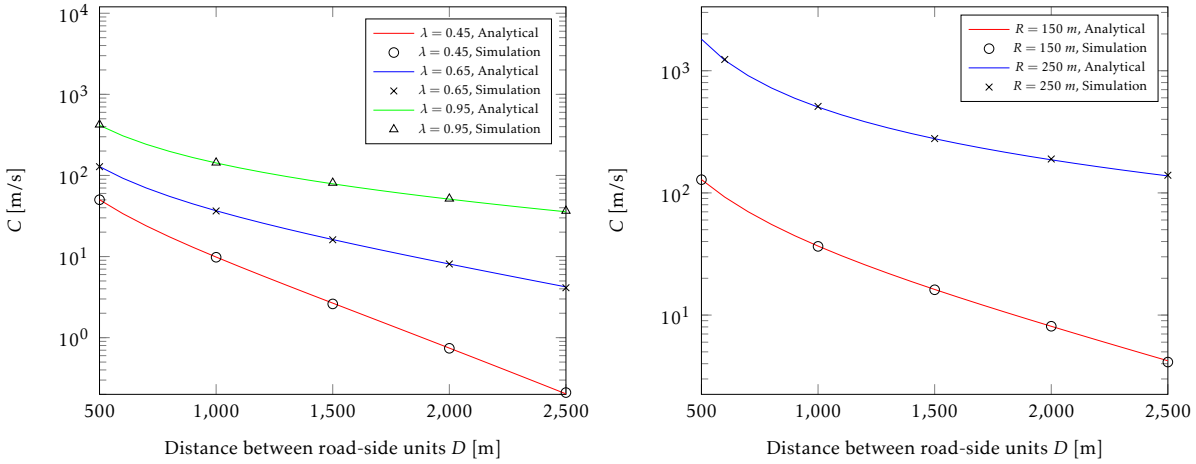


Figure 4.8: Speed of message propagation

These curves are useful to find the optimal distance when planning the RSU deployment. Due to the high cost, complexity and lack of cooperation between government and private sectors the deployment of RSUs is slow. The aforementioned factors lead our research work to find analytical model for RSU placement when we plan to place the RSUs on the highway in the proper position that can improve the connectivity between the vehicles. The optimal RSU distance as the function of the intended message propagation speed is shown in Figure 4.9.

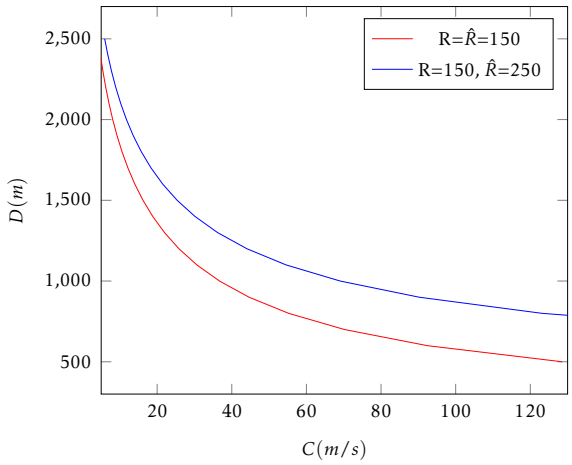


Figure 4.9: Optimal RSUs placement on the highway

4.6.2 *RSU and vehicle with different radio range $\hat{R} \neq R$*

In many published papers [1, 6, 49] regarding the RSU deployment, the radio range of vehicles and RSUs are the same. We extended our investigation to the case of different radio ranges for vehicles and RSUs, which is more realistic, since RSUs have more powerful radio equipment. In this section we study two different radio coverage values for the RSUs, $\hat{R} \in \{250m, 500m\}$ and radio coverage of vehicles equal to 150m. The radio coverage $\hat{R}$ of the RSUs on left-side of figures in this Section is 250m. Moreover, all right hand side of Figures in this Section compared with the results shown in the right hand side of Figures 4.6, 4.7 and 4.8, the comparison is visible in both cases when $\hat{R} = R$ and $\hat{R} \neq R$ as shown in right hand side of Figures in this Section.

As shown on the right side of Figure 4.10, increasing the radio range of the RSUs by a factor of 2 decreases $E(\mathcal{T})$ by at least one order of magnitude. The left side of the Figure is similar to the results with the same ranges (Figure 4.6), there is some improvement due to the greater $\hat{R}$.

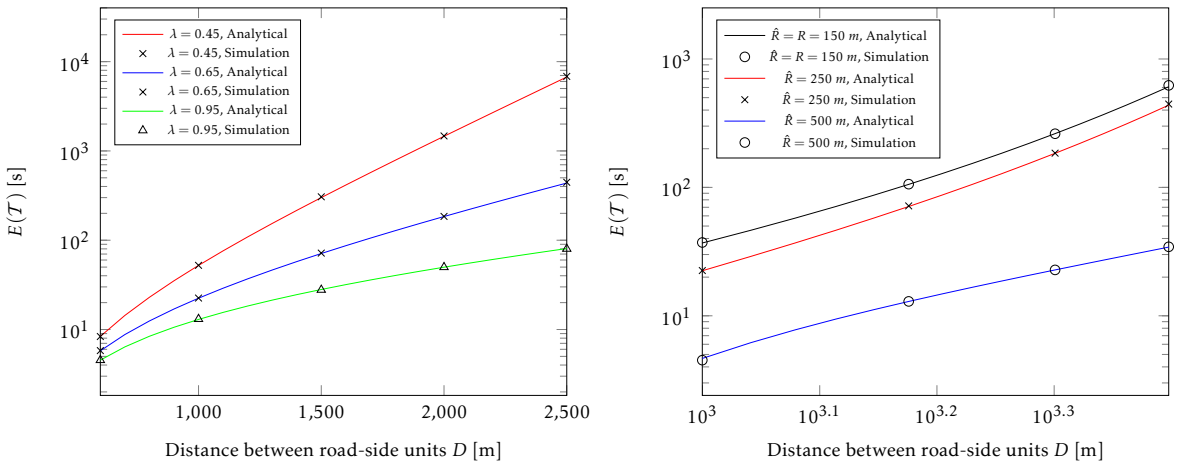


Figure 4.10: The mean time between renewal instants, $E(\mathcal{T})$

Figure 4.11 depicting $E(\mathcal{S})$ is similar to Figure 4.7, the same non-monotonous behavior can be observed here, too. Due to the greater $\hat{R}$, $E(\mathcal{S})$ is greater in this case.

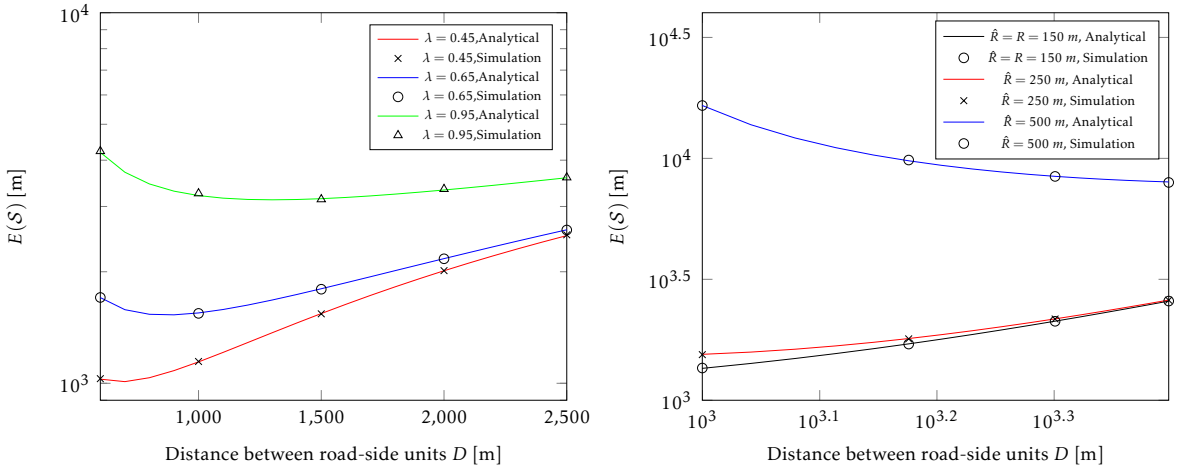


Figure 4.11: The mean distance between the farthest informed RSUs at the renewal instant, $E(\mathcal{S})$

According to the right side of Figure 4.12, the greater radio coverage of the RSU has a big impact on the message propagation speed. On the left side of the Figure, the speed values got improved compared to the $\hat{R} = R$ case.

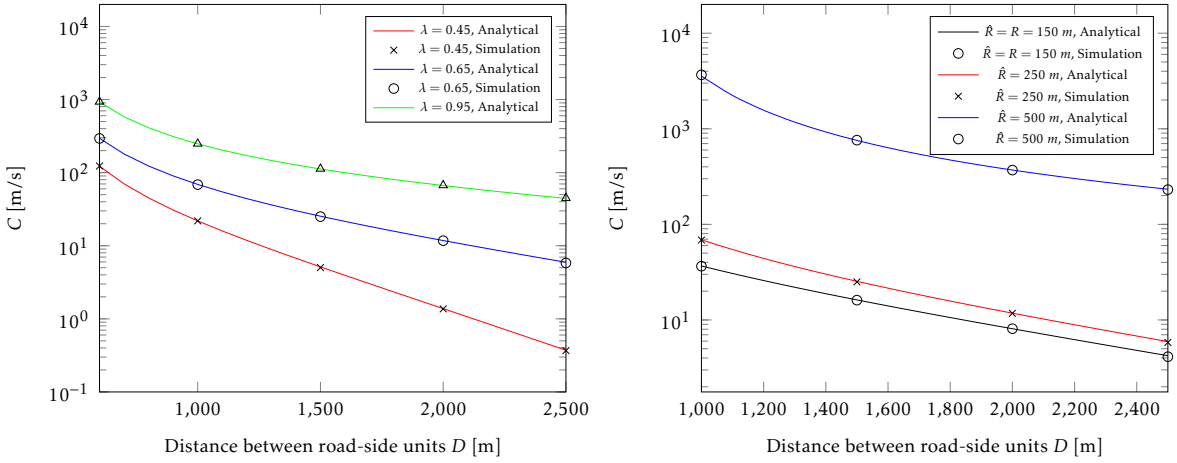


Figure 4.12: Speed of message propagation

4.6.3 *Transient distribution of the farthest informed RSU*

The results of Section 4.5 make it possible to investigate which is the farthest RSU having the message received at time t. To this end, we computed the Laplace transforms $L_i^*(s)$ for $i = 0,\ldots,30$ using (4.35) and applied the CME-based inverse Laplace transform method [29] to get the results in time domain. We have set the parameters to $\hat{R} = R = 250m, D = D' = 1500m$ in this experiment. The distributions are depicted by Figure 4.13. According to the Figure, after 60 seconds it is most likely that RSU number 4 is the farthest informed RSU, after 90 seconds it is RSU number 10 and after 120 seconds it is RSU number 15.

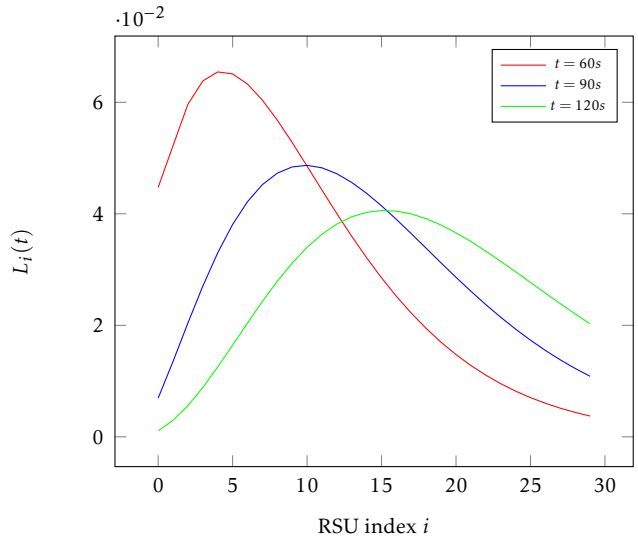


Figure 4.13: Transient distribution of the farthest informed RSU

4.6.4 *Discussion of the results*

Based on the results presented above we can conclude that the dependence of the message propagation speed on parameters D, λ and R is rather complex. In general, greater distance D, lower traffic rate λ and smaller radio coverage R lead to slower message propagation speed, although their relation is far not linear. We also demonstrated that the radio coverage of the RSU, $\hat{R}$, also has a big impact on the message propagation speed, by the appropriate setting of the D and $\hat{R}$ parameters the message propagation can be very fast, well above $100 m/s$.

In order to optimize the placement of the RSUs one has to determine the greatest distance D by which the quality of service (QoS) parameters are satisfied. Based on the results presented in this chapter the criteria can be

- that the message propagation speed exceeds a minimal acceptable value,
- or that the probability that the message reaches a given RSU in a certain amount of time exceeds a threshold.

It would also be important to investigate the case where some kind of connection has been established between RSUs and to examine the case of failure of RSUs and vehicle communications. All the numerical results assume that RSUs are disconnected, the extension of the proposed solution in the future work will consider the second type of RSU communication, i.e., connected RSUs. The connected RSUs have two communication mode, communication between RSUs, where there is direct communication channel, while the another mode the RSUs communicate with vehicles as well. We need to take into account deployment cost and connection type between the RSUs, too.

4.7 SUMMARY

Relying on the Markov renewal process embedded at those instants where the message gets delivered to further RSUs we were able to derive the asymptotic speed of the

message propagation, also in the case when the radio coverage of the vehicles and the RSUs are different. The corresponding formula is explicit, quite compact, and easy to evaluate. According to our numerical investigations with realistic parameters we found that with the proper placement of RSUs it is possible to achieve very high message propagation speed, much higher than the speed of the vehicles of the counter-flow traffic. The results of this chapter can be used in the network planning process of VANETs with disconnected RSUs.

CHAPTER FIVE

5

ALERT MESSAGE PROPAGATION IN VANET ASSUMING MARKOVIAN VEHICLE ARRIVAL PROCESS

Most papers published on the analysis of message propagation assume that the inter arrival times between vehicles follow a Poisson process; there are very few results available with more general traffic model. In this chapter we have shown that the Poisson process is not always suitable for modeling vehicle traffic. Instead of the Poisson process, we propose to use the more general Markovian arrival process (MAP) to model the vehicle headway times, and derive the probability that the message propagates beyond a certain distance from the accident under this traffic assumption. Through several numerical examples, we demonstrate how much the statistics of the vehicle arrival process impacts the message propagation distance.

5.1 BACKGROUND

Accurate modeling of stochastic systems the input of the system has to be described appropriately. In many cases Poisson process is used to describe the input of the system, since it is simple enough for analytical modeling. Poisson arrival process has some unique properties, it is memory-less, the time between two consecutive arrivals is exponentially distributed with parameter λ. The Poisson process is not always suitable to model vehicle arrivals. MAP is another stochastic arrival process introduced in Neuts [45].

More general traffic models are needed to describe the arrival process of the vehicles more accurately. As natural extensions of the Poisson process, Markov arrival processes have been successfully used in several areas to model complex traffic patterns. MAPs are not only more general than the Poisson process, in fact, MAPs include the Poisson process as a special case. When MAPs fail to model the realistic traffic, the Poisson process fails, too. However, if the Poisson process fails, MAPs can still approximate the real behavior reasonably well. MAPs have been used for modeling vehicular traffic previously, starting with [4] and including [16].

In this chapter, we provide several analytical results regarding the message propagation on the highway, such that the vehicle arrival process is given by a MAP. In our scenario, there is a message source, that has a fixed location (i.e., alert messages are generated at an accident). Vehicles can receive the message and get informed when they reach the radio transmission range of the message source or the transmission range of another informed vehicle. The transmission range is assumed to be fixed, and the messages are assumed to be received immediately and without being dropped. This idealistic behavior is called deterministic message passing [61] in the literature, novel broadcasting schemes [74] are getting closer and closer to it. In our model, the arrival process of the vehicles is time stationary (we do not consider seasonality), the velocity of the vehicles is constant, and the effect of the counter-flow traffic is ignored. While the constant velocity assumption does not hold in reality, we need this restriction in order

to keep the analytical model tractable. This scenario seems simple and idealistic, yet it is challenging to study with analytical tools, even assuming Poisson traffic. Our aim to study the message propagation distance, thus, how far the alert message propagates (measured from the accident) after a certain amount of time.

A similar problem has been solved in [8], with a Poisson process and the presence of channel randomness. In [77] and in [43], the mean message propagation distance has been derived for the same scenario, with Poisson vehicle traffic. In [41] the detailed transient analysis of the message propagation is provided, still under the Poisson assumption. Not surprisingly, [43, 77] and [41] arrived to the same results, but with three different analysis approaches: they express the mean message propagation distance explicitly, as the function of the radius of the radio coverage, the vehicle speed and the vehicle arrival intensity. In a more recent paper [54] the message delivery delay is investigated in case of non-constant, uniformly distributed vehicle velocity, but it still assumes Poisson vehicle arrival process.

In this chapter, we show that the Poisson process is not always suitable for modeling vehicle traffic. Instead of the Poisson process, we propose to use the more general Markovian arrival process to model the vehicle headway times, which still allows for tractable analysis. We provide analytical results on both the stationary and the transient properties of the message propagation distance assuming Markovian vehicle arrival process, and show that the more accurate modeling of the vehicle arrival process implies a better approximation of the message propagation distance.

5.2 STOCHASTIC MODELS FOR THE VEHICLE ARRIVAL PROCESS

The Poisson process has been used in the literature for a long time to model the inter-arrival times of the vehicles. The reason for this modeling choice is not its validated correctness, but its analytical simplicity; many useful performance measures can be expressed analytically when Poisson traffic is assumed.

5.2.1 *The failure of Poisson process in modeling vehicular traffic*

After examining a number of recent traces comprising of LIDAR measurements [12], it is easy to show the apparent differences between the statistics of the empirical measurements and the Poisson process.

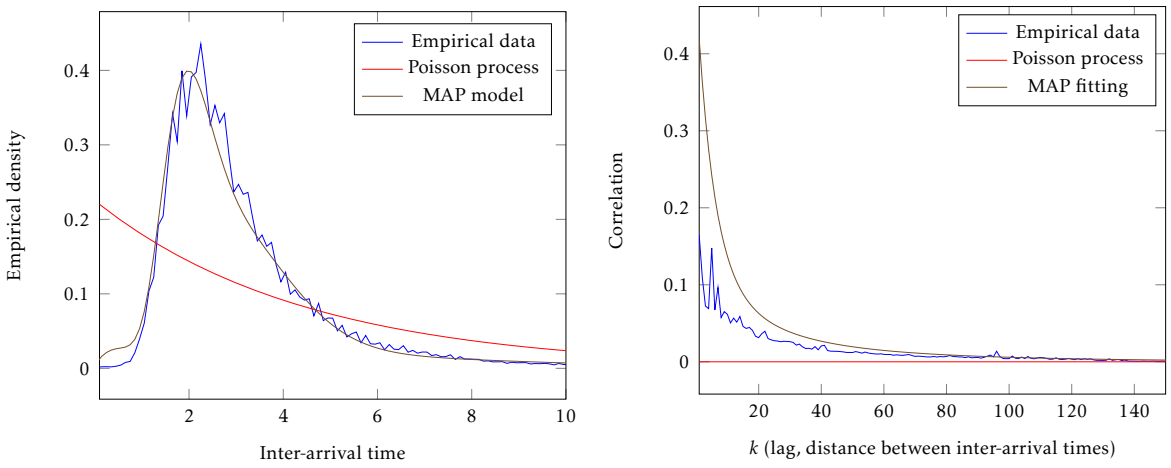


Figure 5.1: Comparison of the empirical pdfs and lag-k correlations of the vehicle inter-arrival times

The left plot in Figure 5.1 depicts the empirical probability density function (pdf) of the empirical measurements of vehicle inter-arrival times (also referred to as headway times in the literature) and the pdf of the Poisson model having the same traffic intensity. It can be seen clearly that the shapes of the pdfs are significantly different. A popular measure used to quantify "burstiness" of the traffic is the squared coefficient of variation (SCV), which is defined by the ratio of the variance and the square of the mean value. In the Poisson case (when the inter-arrival times are exponentially distributed), $SCV = 1$ holds. SCV values lower than 1 correspond to more "regular" (closer to deterministic) traffic, while $SCV > 1$ means that the traffic is highly varying. For this measurement trace, we got $SCV = 46.913$, which means that the vehicle inter-arrival times are extremely varying.

Another property of the Poisson process is that it is unable to capture the correlations between the vehicle inter-arrival times. In practice, however, the inter-arrival times are correlated, as shown by the right plot in Figure 5.1, which depicts the correlation of an inter-arrival time and the kth successive inter-arrival time.

In the subsequent sections, we show that the accurate modeling of vehicular traffic is essential to obtain accurate analytical results for the message propagation distance (and, we believe that for many other performance measures, too).

5.2.2 *Markov arrival process*

Having shown that the Poisson model is not always suitable for vehicular traffic, we propose an alternative solution in this section. The Markovian arrival process (MAP [36]) is a useful modeling tool that is capable of characterizing non-Poisson traffic. It is widely applied in many areas, including telecommunication and logistic systems. MAPs, while being rather general and flexible traffic models, are still relatively simple to work with during analytical derivations.

In a Poisson process, the inter-arrival times are exponentially distributed. In the case of MAPs, the inter-arrival times are not exponentially distributed but are given by the composition of several exponentially distributed phases. More precisely, MAPs have a background process, which is an irreducible continuous-time Markov chain

(CTMC), where two kinds of state transitions are distinguished: those that generate an arrival and those that do not. If the size $N \times N$ generator matrix of the background process is denoted by $\boldsymbol{D}$, we have that $\boldsymbol{D} = \boldsymbol{D_0} + \boldsymbol{D_1}$, where matrix $\boldsymbol{D_1}$ contains the rates of those state transitions which are accompanied by a vehicle arrival, and matrix $\boldsymbol{D_0}$ contains the rate of those transitions that do not generate arrival events, they are internal transitions only. Since $\boldsymbol{D}\mathbb{1} = \underline{0}$ ($\mathbb{1}$ is the column vector of ones and $\underline{0}$ is the column vector of zeros of appropriate size), we have that $\boldsymbol{D_0}\mathbb{1} = -\boldsymbol{D_1}\mathbb{1}$. Figure 5.2 presents an example with $N = 5$, where the dashed lines are the internal transitions, and the solid ones are accompanied by vehicle arrivals. The states of the background process are often referred to as *phases*, too.

If the state process of the MAP is denoted by $\mathcal{J}(t), t \geq 0$, and the kth inter-arrival time by $\mathcal{T}_k$, then the complementary cumulative distribution function (ccdf) and the density function (pdf) of the inter-arrival times are given by [35]

$$P(\mathcal{T}_k > t, \mathcal{J}(t) = j | \mathcal{J}(0) = i) = [e^{\boldsymbol{D_0}t}]_{i,j}, \tag{5.1}$$

$$\lim_{\Delta \to 0} \frac{1}{\Delta} P(\mathcal{T}_k \in (t, t+\Delta), \mathcal{J}(t+\Delta) = j | \mathcal{J}(0) = i) = [e^{\boldsymbol{D_0}t}\boldsymbol{D_1}]_{i,j}, \tag{5.2}$$

if the initial state is i and the state at $t + \Delta$ is j. (Note that $e^{\boldsymbol{A}}$ denotes the matrix-exponential function and $[\boldsymbol{A}]_{i,j}$ denotes the i,jth element of matrix $\boldsymbol{A}$. Matrices are denoted by bold letters throughout the chapter). Since the pdf of the inter-arrival times in a Poisson process is $e^{-\lambda t}\lambda$, MAPs can be considered as the matrix generalizations of the Poisson process.

The stochastic matrix $\boldsymbol{P}$ containing the state transition probabilities between two consecutive inter-arrival events is obtained by

$$\boldsymbol{P} = \int_0^\infty e^{\boldsymbol{D_0}t}\boldsymbol{D_1}\, dt = (-\boldsymbol{D_0})^{-1}\boldsymbol{D_1}, \tag{5.3}$$

and will be used frequently in the sequel. The stationary row vector of matrix $\boldsymbol{P}$, denoted by $\underline{\pi}$, is the solution of $\underline{\pi}\boldsymbol{P} = \underline{\pi}, \underline{\pi}\mathbb{1} = 1$. With this vector we can express the scalar (state independent) pdf of the inter arrival times from (5.2) as [35]

$$h(t) = \lim_{\Delta \to 0} \frac{1}{\Delta} P(\mathcal{T}_k \in (t, t+\Delta)) = \underline{\pi} e^{\boldsymbol{D_0}t}\boldsymbol{D_1}\mathbb{1}. \tag{5.4}$$

Based on this result, it is easy to derive the following simple quantities characterizing the inter-arrival times [35]:

$$H(t) = P(\mathcal{T}_k < t) = 1 - \underline{\pi} e^{\boldsymbol{D_0}t}\mathbb{1}, \tag{5.5}$$

$$E(\mathcal{T}_k^n) = n! \cdot \underline{\pi}(-\boldsymbol{D_0})^{-n}\mathbb{1}. \tag{5.6}$$

The inter-arrival times generated by a MAP can be correlated, too. The lag-k correlation, thus the correlation between $\mathcal{T}_1$ and $\mathcal{T}_{k+1}$ can be calculated by [35]

$$\mathrm{Corr}(\mathcal{T}_1, \mathcal{T}_{k+1}) = \frac{E(\mathcal{T}_1\mathcal{T}_{k+1}) - E(\mathcal{T}_1)E(\mathcal{T}_{k+1})}{\sqrt{E(\mathcal{T}_1^2) - E^2(\mathcal{T}_1)}\sqrt{E(\mathcal{T}_{k+1}^2) - E^2(\mathcal{T}_{k+1})}} = \frac{\lambda \underline{\alpha}\boldsymbol{P}^k(-\boldsymbol{D_0})^{-1}\mathbb{1} - 1}{2\lambda \underline{\alpha}(-\boldsymbol{D_0})^{-1}\mathbb{1} - 1}, \tag{5.7}$$

where $\lambda = 1/E(\mathcal{T}_k)$ is the mean arrival rate and $\underline{\alpha}$ is the stationary probability vector of the CTMC given by matrix $\boldsymbol{D}$, satisfying $\underline{\alpha}\boldsymbol{D} = \underline{0}, \underline{\alpha}\mathbb{1} = 1$.

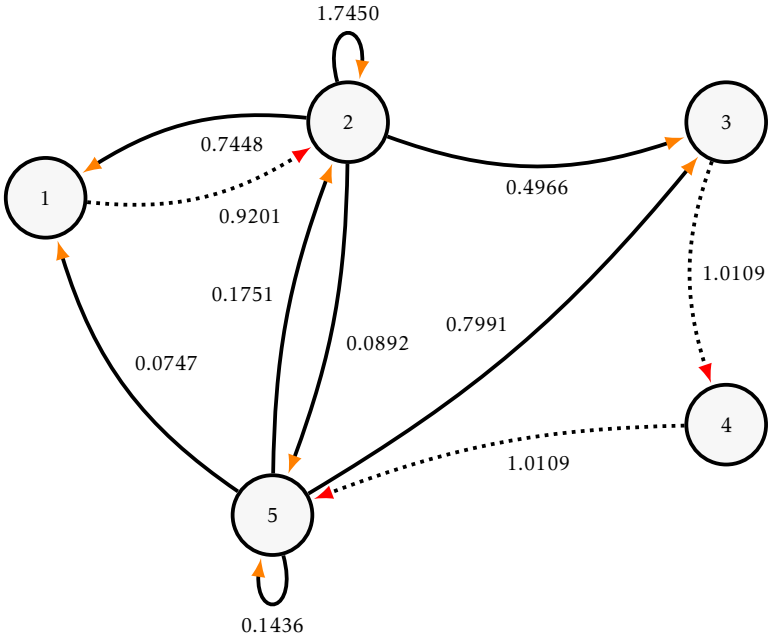


Figure 5.2: Markov Arrival Process transition diagram

$$\boldsymbol{D} = \underbrace{\begin{pmatrix} -0.9201 & 0.9201 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & -3.0756 & 0.0000 & 0.0 & 0.0000 \\ 0.0000 & 0.0000 & -1.0109 & 1.0109 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & -1.0109 & 1.0109 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & -1.1925 \end{pmatrix}}_{\boldsymbol{D}_0} + \underbrace{\begin{pmatrix} 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.7448 & 1.7450 & 0.4966 & 0.0000 & 0.0892 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0747 & 0.1751 & 0.7991 & 0.0000 & 0.1436 \end{pmatrix}}_{\boldsymbol{D}_1}$$

5.2.3 *Obtaining the MAP from empirical measurement data*

MAPs form a dense class in the set of point processes [7], which means that every point process can be approximated arbitrary well with a MAP having appropriately many phases. The critical question is, how to obtain the matrix parameters of the MAP, $\boldsymbol{D_0}, \boldsymbol{D_1}$, to approximate the real traffic behavior accurately.

One possibility is to build the background Markov chain of the MAP based on the intuitive understanding of the traffic. For instance, in Figure 5.3, the arrival process is Poisson, but the arrival rate is time-dependent, where the periods determining the arrival rate are exponentially distributed and follow each other in a circular way. The solid lines represent transitions of $\boldsymbol{D_1}$, and the dashed lines the ones of $\boldsymbol{D_0}$ (here, vehicle arrivals are not accompanied by phase transitions).

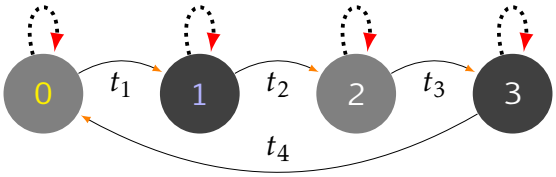


Figure 5.3: State-transition diagram for MAP based-on VANET

When empirical measurement data is available, an automatic approach can also be used. The process of obtaining the $\boldsymbol{D_0}, \boldsymbol{D_1}$ matrix parameters of a MAP from empirical measurements is called *MAP fitting*. There are several MAP fitting algorithms available. Some of them compute statistical quantities (marginal moments, auto-correlations, joint moments, etc.), and create the $\boldsymbol{D_0}, \boldsymbol{D_1}$ matrices such that the resulting MAP exhibits the same statistics [11, 26, 59]. Some other algorithms try to find the MAP that is most likely to generate the measurement data with the expectation-maximization algorithm [27, 47]. Nevertheless, MAP fitting is an inherently difficult problem, the corresponding algorithms are continuously improving, but there is no single perfect tool available yet.

For our VANET traffic trace we have applied the EM algorithm in [27] and the KPC-toolbox [11]. From the two, the first one provided better results, and it will be used for numerical investigations in the rest of the chapter. As depicted by Figure 5.1, the pdf of the vehicle inter-arrival times is accurately captured by the MAP model, as opposed to the Poisson model. We got somewhat worse matching for the correlations, which can be a consequence of the too low number of data to fit.

5.3 CLUSTERS OF INFORMED VEHICLES

We are going to study the stochastic behavior of the *cluster length* $\mathcal{G}$ (see Figure 3.2 as describe in Section 3.3.1) , as it is an essential ingredient in the analysis of the message propagation distance. Since we have a MAP vehicle arrival process, to describe the system completely, we have to consider the joint behavior of $\mathcal{G}$ and the phase of the MAP. Namely, we have to keep track of the phase of the MAP at two specific time points: at the time when the first vehicle of the cluster was generated, and at $\mathcal{G}/v$ later (at the end of the cluster).

If the phase of the MAP at time t is denoted by $\mathcal{J}(t)$, and the first vehicle of the cluster was generated at time $t = 0$ (without loss of generality), the complementary cumulative distribution function (ccdf) of $\mathcal{G}$ and the phase at the end of the cluster is defined by

$$G_{ij}(x) = P(\mathcal{G} > x, \mathcal{J}(\mathcal{G}/v) = j | \mathcal{J}(0) = i), \tag{5.8}$$

and the corresponding matrix $\boldsymbol{G}(x)$ is defined by $\boldsymbol{G}(x) = [G_{ij}(x)]$.

Before expressing $\boldsymbol{G}(x)$, let us introduce matrix $\boldsymbol{Z} = [Z_{ij}]$ with the transition probabilities of the MAP phases between the beginning and the end of the cluster as

$$Z_{ij} = P(\mathcal{J}(\mathcal{G}/v) = j | \mathcal{J}(0) = i). \tag{5.9}$$

Thus, Z_{ij} is the probability that the phase of the MAP is j at the end of the cluster, given that it was j at the beginning of it. For matrix $\boldsymbol{Z}$ an explicit solution is provided by the next theorem.

Theorem 12. *Matrix $\boldsymbol{Z}$ can be expressed by*

$$\boldsymbol{Z} = (\boldsymbol{I} - \boldsymbol{P} + e^{\boldsymbol{D_0} R/v}\boldsymbol{P})^{-1} e^{\boldsymbol{D_0} R/v}. \tag{5.10}$$

Proof. By conditioning on the time between two vehicles, there are two possibilities: either the headway time is greater than R/v, or it is less than or equal to R/v. In the former case the two vehicles do not belong to the same cluster. In the latter case they are in the same cluster, and the phase transition probabilities can be calculated recursively, based on the cluster initiated by the second vehicle. Using (5.1) and (5.2) we get

$$Z_{ij} = [e^{\boldsymbol{D_0} R/v}]_{ij} + \sum_{k=1}^{N} \int_{y=0}^{R/v} [e^{\boldsymbol{D_0} y}\boldsymbol{D_1}]_{ik} Z_{kj}\, dy, \tag{5.11}$$

which, switching to matrix notation, translates to

$$\boldsymbol{Z} = e^{\boldsymbol{D_0} R/v} + \int_{y=0}^{R/v} e^{\boldsymbol{D_0} y}\boldsymbol{D_1}\boldsymbol{Z}\, dy,$$

since at speed v the transmission range R, in distance, translates to R/v, in time. Moreover, the second term of $\boldsymbol{Z}$ can be simplified as

$$\begin{aligned}\int_{y=0}^{R/v} e^{\boldsymbol{D_0} y}\boldsymbol{D_1}\boldsymbol{Z}\, dy &= \int_{y=0}^{\infty} e^{\boldsymbol{D_0} y}\boldsymbol{D_1}\boldsymbol{Z}\, dy - e^{\boldsymbol{D_0} R/v}\int_{y=R/v}^{\infty} e^{\boldsymbol{D_0}(y-R/v)} dy \boldsymbol{D_1}\boldsymbol{Z} \\ &= \underbrace{(-\boldsymbol{D_0})^{-1}\boldsymbol{D_1}}_{\boldsymbol{P}}\boldsymbol{Z} - e^{\boldsymbol{D_0} R/v}\underbrace{(-\boldsymbol{D_0})^{-1}\boldsymbol{D_1}}_{\boldsymbol{P}}\boldsymbol{Z},\end{aligned}$$

which, after simple transformations, establishes the theorem. More proof details can be found in Appendix A.1 □

Similar to the results available for the Poisson case [41], by the stochastic interpretation of the system, $\boldsymbol{G}(x)$ can be expressed recursively as

$$\boldsymbol{G}(x) = \begin{cases} \boldsymbol{Z}, & \text{if } x \le R, \\ \int_{y=0}^{R/v} e^{\boldsymbol{D_0} y}\boldsymbol{D_1}\boldsymbol{G}(x - yv)\, dy, & \text{if } x > R. \end{cases} \tag{5.12}$$

The first term of (5.12) corresponds to the case when x falls into the transmission range of the first vehicle of the cluster. In this case $G_{ij}(x) = Z_{ij}$, since $\mathcal{G} \ge R$ holds, we have to take care only about the phase transitions (see (5.8) and (5.9)).

In the second case of (5.12), x is out of range for the first vehicle, thus to cover x, the cluster has to consist of more vehicles (Figure 3.2). Conditioning on the inter-arrival time between the first and the second vehicle of the cluster (having pdf $e^{\boldsymbol{D_0} y}\boldsymbol{D_1}$) leads to (5.12).

In order to compute the mean cluster length $E(\mathcal{G})$, we define the phase-dependent mean cluster length that contains information on the MAP phase when the first car was generated $\mathcal{J}(0)$ and the MAP phase at the end of the cluster $\mathcal{J}(\mathcal{G}/v)$. The i,jth element of the phase-dependent mean cluster length matrix is defined by $[\boldsymbol{E}(\mathcal{G})]_{i,j} = E(\mathcal{G} \cdot \mathcal{I}_{\{\mathcal{J}(\mathcal{G}/v)=j\}} | \mathcal{J}(0) = i)$, where $\mathcal{I}_{\{\cdot\}}$ is the indicator variable.

Theorem 13. *The phase-dependent mean value of $\mathcal{G}$ can be calculated by*

$$\boldsymbol{E}(\mathcal{G}) = (\boldsymbol{I}-\boldsymbol{P}+e^{\boldsymbol{D_0}R/v}\boldsymbol{P})^{-1}\Big[(\boldsymbol{I}-\boldsymbol{P})R\boldsymbol{Z}-\boldsymbol{D_0}^{-1}\boldsymbol{Z}v\Big]+\boldsymbol{Z}\boldsymbol{D_0}^{-1}v. \tag{5.13}$$

Proof. To obtain the mean value, the integral of the ccdf (5.12) is calculated:

$$\begin{aligned}
\boldsymbol{E}(\mathcal{G}) &= \int_{x=0}^{\infty}\boldsymbol{G}(x)dx = \int_{x=0}^{R}\boldsymbol{G}(x)dx+\int_{x=R}^{\infty}\boldsymbol{G}(x)dx \\
&= R\boldsymbol{Z}+\int_{y=0}^{R/v}e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\int_{x=R}^{\infty}\boldsymbol{G}(x-yv)dx\,dy \\
&= R\boldsymbol{Z}+\int_{y=0}^{R/v}e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\underbrace{\int_{x=yv}^{\infty}\boldsymbol{G}(x-yv)dx\,dy}_{\boldsymbol{E}(\mathcal{G})}-\int_{y=0}^{R/v}e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\underbrace{\int_{x=yv}^{R}\boldsymbol{G}(x-yv)dx\,dy}_{(R-yv)\boldsymbol{Z}} \\
&= R\boldsymbol{Z}+\underbrace{\int_{y=0}^{R/v}e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\boldsymbol{E}(\mathcal{G})dy}_{(\boldsymbol{I}-e^{\boldsymbol{D_0}R/v})\boldsymbol{P}\boldsymbol{E}(\mathcal{G})}-\underbrace{\int_{y=0}^{R/v}(R-yv)e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\boldsymbol{Z}dy}_{R\boldsymbol{P}\boldsymbol{Z}+v\boldsymbol{D_0}^{-1}\boldsymbol{P}\boldsymbol{Z}-ve^{\boldsymbol{D_0}R/v}\boldsymbol{D_0}^{-1}\boldsymbol{P}\boldsymbol{Z}},
\end{aligned}$$

where, in the last term, we have exploited that $(\boldsymbol{I}-e^{\boldsymbol{D_0}R/v})\boldsymbol{P}\boldsymbol{Z} = \boldsymbol{Z}-e^{\boldsymbol{D_0}R/v}$ and that $(\boldsymbol{I}-\boldsymbol{P}+e^{\boldsymbol{D_0}R/v}\boldsymbol{P})^{-1}e^{\boldsymbol{D_0}R/v} = \boldsymbol{Z}$ from (5.10). Collecting the $\boldsymbol{E}(\mathcal{G})$ terms together provides the theorem. More proof details can found in Appendix A.1.2. □

In the next step, we compute the phase-independent mean cluster length $E(\mathcal{G})$ from matrix $\boldsymbol{E}(\mathcal{G})$. To do so, we have to obtain the stationary phase distribution of the MAP at the beginning of the clusters, denoted by row vector $\underline{\gamma}$.

Lemma 14. *The stationary phase distribution vector of the MAP at the beginning of the cluster, vector $\underline{\gamma}$, satisfies the linear set of equations $\underline{\gamma}\boldsymbol{Z}\boldsymbol{P} = \underline{\gamma}$, $\underline{\gamma}\mathbb{1} = 1$.*

Proof. Let us define a discrete-time Markov chain characterizing the evolution of the phases at the beginning of the clusters. The phase transitions between the beginning and the end of the clusters are governed by stochastic matrix $\boldsymbol{Z}$. The phase transitions between the beginning and the end of the intra-cluster periods are governed by stochastic $\boldsymbol{P} = (-\boldsymbol{D_0})^{-1}\boldsymbol{D_1}$. Hence, the phase transitions over the beginning of two consecutive clusters are given by matrix $\boldsymbol{Z}\boldsymbol{P}$ (that is also stochastic), whose stationary distribution provides $\underline{\gamma}$. □

Corollary 4. *The phase-independent mean cluster length, $E(\mathcal{G})$ is*

$$E(\mathcal{G}) = \underline{\gamma}\boldsymbol{E}(\mathcal{G})\mathbb{1} = v\underline{\gamma}\Big((\boldsymbol{I}-\boldsymbol{P}+e^{\boldsymbol{D_0}R/v}\boldsymbol{P})^{-1}(-\boldsymbol{D_0})^{-1}\boldsymbol{Z}-\boldsymbol{Z}(-\boldsymbol{D_0})^{-1}\Big)\mathbb{1}. \tag{5.14}$$

After the description of the first moment by (5.14), the next theorem and the corresponding corollary express the second moment. Based on the second moment, it is possible to study the variance of the cluster length, but it is an important ingredient necessary to compute the mean information distance in the next section, too.

Theorem 15. *The phase-dependent second moment of the cluster length can be expressed by*

$$\begin{aligned}\boldsymbol{E}(\mathcal{G}^2) &= \boldsymbol{Z}(v\boldsymbol{D_0}^{-1}-R\boldsymbol{I})(2\boldsymbol{P}\boldsymbol{E}(\mathcal{G})-2v\boldsymbol{D_0}^{-1}\boldsymbol{P}\boldsymbol{Z}) \\ &\quad +(\boldsymbol{I}-\boldsymbol{P}+e^{\boldsymbol{D_0}R/v}\boldsymbol{P})^{-1}\left[2v(-\boldsymbol{D_0})^{-1}\boldsymbol{P}\boldsymbol{E}(\mathcal{G})+2v^2(-\boldsymbol{D_0})^{-2}\boldsymbol{P}\boldsymbol{Z}\right].\end{aligned} \tag{5.15}$$

Proof. To obtain $\boldsymbol{E}(\mathcal{G}^2)$, the integral of the ccdf (5.12) is calculated by

$$\begin{aligned}\boldsymbol{E}(\mathcal{G}^2) &= 2\int_{x=0}^{\infty} x\boldsymbol{G}(x)dx = 2\int_{x=0}^{R} x\boldsymbol{G}(x)dx+2\int_{x=R}^{\infty} x\boldsymbol{G}(x)dx \\ &= \underbrace{2\int_{x=0}^{R} x\boldsymbol{Z}dx}_{R^2\boldsymbol{Z}}+2\int_{y=0}^{R/v} e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\underbrace{\int_{x=R}^{\infty} x\boldsymbol{G}(x-yv)dx}_{(*)}dy,\end{aligned} \tag{5.16}$$

where the term marked with $(*)$ can be simplified as

$$\underbrace{\int_{x=yv}^{\infty}(x-yv+yv)\boldsymbol{G}(x-yv)dx}_{\boldsymbol{E}(\mathcal{G}^2)/2+yv\boldsymbol{E}(\mathcal{G})}-\underbrace{\int_{x=yv}^{R}(x-yv+yv)\boldsymbol{G}(x-yv)dx}_{(R^2-(yv)^2)\boldsymbol{Z}/2}.$$

Putting this result back to (5.16), for all terms of $\boldsymbol{E}(\mathcal{G}^2)$ we get:

$$\begin{aligned}\boldsymbol{E}(\mathcal{G}^2) &= R^2\boldsymbol{Z}+\underbrace{\int_{y=0}^{R/v} e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\boldsymbol{E}(\mathcal{G}^2)dy}_{(\boldsymbol{I}-e^{\boldsymbol{D_0}R/v})\boldsymbol{P}\boldsymbol{E}(\mathcal{G}^2)}+2\int_{y=0}^{R/v} e^{\boldsymbol{D_0}y}\boldsymbol{D_1}yv\boldsymbol{E}(\mathcal{G})dy \\ &\quad -\int_{y=0}^{R/v} e^{\boldsymbol{D_0}y}\boldsymbol{D_1}(R^2-(yv)^2)\boldsymbol{Z}dy.\end{aligned} \tag{5.17}$$

To simplify the third term of (5.17), we apply integration by parts, leading to

$$2\int_{y=0}^{R/v} e^{\boldsymbol{D_0}y}\boldsymbol{D_1}yv\boldsymbol{E}(\mathcal{G})dy = \left[-2Re^{\boldsymbol{D_0}R/v}-2v\boldsymbol{D_0}^{-1}+2v\boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}R/v}\right]\boldsymbol{P}\boldsymbol{E}(\mathcal{G}).$$

In the same way we can also simplify the the last term of (5.17) as

$$\begin{aligned}&\int_{y=0}^{R/v} e^{\boldsymbol{D_0}y}\boldsymbol{D_1}(yv)^2\boldsymbol{Z}dy \\ &\quad = \left[-R^2+2vR\boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}R/v}+2v^2(\boldsymbol{D_0})^{-2}-2v^2e^{\boldsymbol{D_0}R/v}(\boldsymbol{D_0})^{-2}\right]\boldsymbol{P}\boldsymbol{Z}\end{aligned}$$

Finally, putting all terms together gives

$$\begin{aligned}\boldsymbol{E}(\mathcal{G}^2) &= (\boldsymbol{I}-\boldsymbol{P}+e^{\boldsymbol{D_0}R/v}\boldsymbol{P})^{-1}\Big[R^2(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{Z}-2Re^{\boldsymbol{D_0}R/v}\boldsymbol{P}\boldsymbol{E}(\mathcal{G}) \\ &\quad -2v\boldsymbol{D_0}^{-1}\boldsymbol{P}\boldsymbol{E}(\mathcal{G})+2v\boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}R/v}\boldsymbol{P}\boldsymbol{E}(\mathcal{G})+2vR\boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}R/v}\boldsymbol{P}\boldsymbol{Z} \\ &\quad +2v^2\boldsymbol{D_0}^{-2}\boldsymbol{P}\boldsymbol{Z}-2v^2e^{\boldsymbol{D_0}R/v}(-\boldsymbol{D_0})^{-2}\boldsymbol{P}\boldsymbol{Z}\Big].\end{aligned}$$

Exploiting $(\boldsymbol{I}-\boldsymbol{P}+e^{\boldsymbol{D_0}R/v}\boldsymbol{P})^{-1}e^{\boldsymbol{D_0}R/v}=\boldsymbol{Z}$ from (5.10) and applying simple transformations establishes the theorem. More proof details can found in Appendix A.1.3. □

As we did with the mean cluster length, we can also obtain the phase-independent (scalar) second moment of the cluster length by multiplying $\boldsymbol{E}(\mathcal{G}^2)$ by $\underline{\pi}$ from the left and by $\mathbb{1}$ from the right.

Corollary 5. *The phase-independent second moment of the cluster length, $E(\mathcal{G}^2)$ is*

$$E(\mathcal{G}^2)=2v\underline{\pi}(\boldsymbol{I}-\boldsymbol{P}+e^{\boldsymbol{D_0}R/v}\boldsymbol{P})^{-1}(-\boldsymbol{D_0})^{-1}\boldsymbol{E}(\mathcal{G})\mathbb{1}-2RE(\mathcal{G})-2vR\underline{\pi}\boldsymbol{Z}(-\boldsymbol{D_0})^{-1}\mathbb{1}. \tag{5.18}$$

Finally, we provide a differential equation for the ccdf $\boldsymbol{G}(x)$ itself. The motivation is that it is usually easier to solve a differential equation than the integral equation of (5.12).

Theorem 16. *The phase-dependent ccdf of $\mathcal{G}$ is the solution of the delayed differential equation (DDE)*

$$\frac{d}{dx}\boldsymbol{G}(x)=\boldsymbol{D}\boldsymbol{G}(x)-e^{\boldsymbol{D_0}R/v}\boldsymbol{D_1}\boldsymbol{G}(x-R),\quad x>R, \tag{5.19}$$

with boundary condition $\boldsymbol{G}(x)=\boldsymbol{Z},x\le R$.

Proof. Let us express and manipulate $\boldsymbol{G}(x+R)$. For $x>0$, we need to use the second case of (5.12) only. Changing variable in the integral gives

$$\boldsymbol{G}(x+R)=\int_{y=0}^{R/v}e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\boldsymbol{G}(x+R-yv)\,dy=\int_{u=x}^{x+R}e^{\boldsymbol{D_0}(x+R-u)/v}\boldsymbol{D_1}\boldsymbol{G}(u)\,du.$$

Taking the derivative and using the Leibniz integral rule leads to

$$\begin{aligned}\frac{d}{dx}\boldsymbol{G}(x+R)&=\boldsymbol{D_1}\boldsymbol{G}(x+R)-e^{\boldsymbol{D_0}R/v}\boldsymbol{D_1}\boldsymbol{G}(x)+\boldsymbol{D_0}\underbrace{\int_{u=x}^{x+R}e^{\boldsymbol{D_0}(x+R-u)/v}\boldsymbol{D_1}\boldsymbol{G}(u)du}_{\boldsymbol{G}(x+R)}\\&=\underbrace{(\boldsymbol{D_0}+\boldsymbol{D_1})}_{\boldsymbol{D}}\boldsymbol{G}(x+R)-e^{\boldsymbol{D_0}R/v}\boldsymbol{D_1}\boldsymbol{G}(x),\end{aligned}$$

which establishes the theorem. □

5.4 ANALYSIS OF THE MESSAGE PROPAGATION

Assume that an event occurs on the highway at position A. In the rest of the chapter, this position is assumed to be fixed, and messages advertising the event are generated continuously for a long time. Here we study the right-continuous stochastic process $\{\mathcal{D}(t),t>0\}$, where $\mathcal{D}(t)$ is the information propagation distance, that is the position

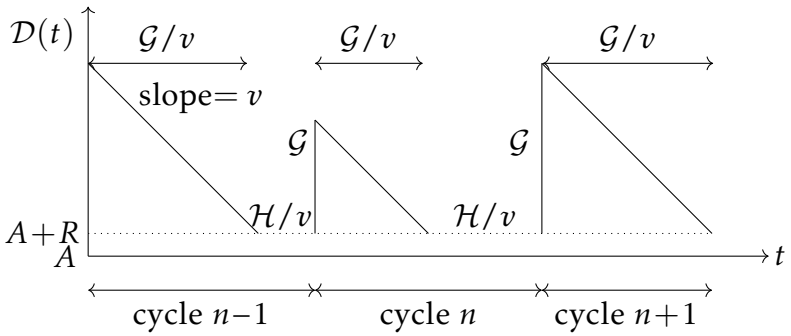


Figure 5.4: The evolution of the information distance $\mathcal{D}(t)$

of the last car measured from A having the message received at time t, plus R (the radius of its radio coverage). In [41] we have studied the same scenario with the Poisson arrival process, where we have shown that due to the PASTA property $E(\mathcal{D})$ and $E(\mathcal{G})$ are equal. In the case of MAP, the $E(\mathcal{D})$ and $E(\mathcal{G})$ are not the same, and we are going to investigate their relation in this section.

The time evolution of $\mathcal{D}(t)$ is shown in Figure 5.4. The trajectory of $\mathcal{D}(t)$ consists of alternating intervals. There are intervals where no vehicles hold the information; the length of these intervals (in distance) is denoted by $\mathcal{H}$. Then, a vehicle enters the range of the accident and gets informed, informing its cluster of length $\mathcal{G}$ as well. This informed cluster will leave the accident in time $\mathcal{G}/v$, followed by another uninformed interval, etc.

The mean length of the uninformed intervals is easy to derive. The phase of the MAP at the beginning of a cluster is given by vector $\underline{\chi}$; on the other hand, the phase at the end of the cluster is $\underline{\chi}\boldsymbol{Z}$. With this initial phase, the mean time till the MAP generates an arrival is given by $\underline{\chi}\boldsymbol{Z}(-\boldsymbol{D_0})^{-1}\mathbb{1}$, thus we have

$$E(\mathcal{H}) = v\underline{\chi}\boldsymbol{Z}(-\boldsymbol{D_0})^{-1}\mathbb{1}. \tag{5.20}$$

From the properties of $\mathcal{H}$ and $\mathcal{G}$ derived above, we can characterize the properties of $\mathcal{D}(t)$.

5.4.1 *Mean information distance*

Let us first derive the mean value of $\mathcal{D} = \lim_{t\to\infty}\mathcal{D}(t)$, denoted by $E(\mathcal{D})$.

Theorem 17. *The mean information distance is expressed by*

$$E(\mathcal{D}) = \frac{E(\mathcal{G}^2)/2}{E(\mathcal{G})+E(\mathcal{H})} + R. \tag{5.21}$$

Proof. The process $\{\mathcal{D}(t), \mathcal{J}(t)\}$ forms a Markov renewal process at the moments when the current informed cluster leaves the accident. This means that at these time instants the phase of the MAP characterizes the future of the process completely.

The theorem can be proven by the renewal reward theorem [14] as

$$E(\mathcal{D}) - R = \lim_{t\to\infty}\frac{1}{t}\int_{u=0}^{t}\mathcal{D}(u)du = \frac{\int_0^{\mathcal{C}}(\mathcal{D}(u)-R)du}{E(\mathcal{C})}, \tag{5.22}$$

where $\mathcal{C}$ denotes the time duration of a cycle, which is the sum of the informed and the subsequent uninformed intervals. Hence, (5.22) is the integral of the information distance in a stationary cycle, divided by the mean cycle time.

If the joint density of $\mathcal{G}$ and the corresponding phase transition is defined by matrix $\boldsymbol{g}(x) = -\frac{d}{dx}\boldsymbol{G}(x)$, $E(\mathcal{C})$ is calculated as

$$\begin{aligned} E(\mathcal{C}) &= \int_{x=0}^{\infty}\int_{y=0}^{\infty}\left(\frac{x}{v}+y\right)\underline{\gamma}\boldsymbol{g}(x)e^{\boldsymbol{D}_0 y}\boldsymbol{D}_1\mathbb{1}\,dy\,dx \\ &= \int_{x=0}^{\infty}\int_{y=0}^{\infty}\frac{x}{v}\underline{\gamma}\boldsymbol{g}(x)e^{\boldsymbol{D}_0 y}\boldsymbol{D}_1\mathbb{1}\,dy\,dx + \int_{x=0}^{\infty}\int_{y=0}^{\infty} y\underline{\gamma}\boldsymbol{g}(x)e^{\boldsymbol{D}_0 y}\boldsymbol{D}_1\mathbb{1}\,dy\,dx \\ &= \frac{1}{v}\underbrace{\int_{x=0}^{\infty} x\underline{\gamma}\boldsymbol{g}(x)\boldsymbol{P}\mathbb{1}\,dx}_{\boldsymbol{E}(\mathcal{G})} + \underline{\gamma}\underbrace{\int_{x=0}^{\infty}\boldsymbol{g}(x)\,dx}_{\boldsymbol{Z}}\underbrace{\int_{y=0}^{\infty} y\,e^{\boldsymbol{D}_0 y}\boldsymbol{D}_1\,dy}_{(-\boldsymbol{D}_0)^{-1}(-\boldsymbol{D}_0)^{-1}\boldsymbol{D}_1}\mathbb{1} \\ &= E(\mathcal{G})/v + \underline{\gamma}\boldsymbol{Z}(-\boldsymbol{D}_0)^{-1}\mathbb{1} = E(\mathcal{G})/v + E(\mathcal{H})/v, \end{aligned} \tag{5.23}$$

where y represents the duration of the uninformed interval and x is the cluster length in a specific cycle, and $\underline{\gamma}\boldsymbol{g}(x)e^{\boldsymbol{D}_0 y}\boldsymbol{D}_1\mathbb{1}$ is their joint pdf.

The numerator of (5.22) can be calculated as

$$\int_0^{\mathcal{C}}(\mathcal{D}(u)-R)du = \int_{x=0}^{\infty}\underline{\gamma}\boldsymbol{g}(x)\mathbb{1}\frac{x^2}{2}\frac{1}{v}dx = \frac{1}{2v}E(\mathcal{G}^2), \tag{5.24}$$

since $x^2/2v$ is the area below one "triangle" in Figure 5.4. Substituting (5.23) and (5.24) into (5.22) we get

$$E(\mathcal{D}) = \frac{\int_0^{\mathcal{C}}(\mathcal{D}(u)-R)du}{E(\mathcal{C})}+R = \frac{\frac{1}{2v}E(\mathcal{G}^2)}{E(\mathcal{G})/v+E(\mathcal{H})/v}+R,$$

which equals (5.21). □

5.4.1.1 *The transient analysis of $\mathcal{D}(t)$*

After the analysis of the mean information distance $E(\mathcal{D})$, we are going to study the transient behavior of the process $\mathcal{D}(t)$. This is one of the most interesting measures when the effect of an event/accident is analyzed since it is important to know how far the alert message gets t time after the accident.

Since we have a MAP vehicle arrival process, the analysis of the joint behavior of $\{\mathcal{D}(t),\mathcal{J}(t)\}$ is easier than analyzing $\mathcal{D}(t)$ alone. More precisely, we are going to introduce the random variable $\widehat{\mathcal{J}}(t)$ as the phase of the MAP at the moment when the cluster present at time t will leave the accident, and define the joint ccdf $F_i(t,x) = P(\mathcal{D}(t) > x, \widehat{\mathcal{J}}(t) = i)$, and the corresponding row vector $\underline{F}(t,x) = [F_i(t,x)]$. The joint probability of being in the uninformed interval and in a certain phase at time t is given by row vector $\underline{\beta}(t) = [\beta_i(t)]$, with elements $\beta_i(t) = P(\mathcal{D}(t) = R, \mathcal{J}(t) = i)$. Note that

for the latter quantity we use $\mathcal{J}(t)$ instead of $\widehat{\mathcal{J}}(t)$. Hence, in the uninformed intervals $\underline{\beta}(t)$ follows the evolution of the background Markov chain of the MAP, and when a vehicle enters the range of the accident, we let the MAP generate all the vehicles of the cluster immediately and freeze its phase at the end of the cluster. Hence, as long as the cluster has not left the accident yet, $\underline{F}(t,x)$ will characterize the information distance and the MAP phase at the end of the current cluster.

Theorem 18. *The transient ccdf $\underline{F}(t,x), x > R$ and the probability of an uninformed interval $\underline{\beta}(t)$ satisfy the partial differential equations (PDEs)*

$$\frac{\partial}{\partial t}\underline{F}(t,x) - v\frac{\partial}{\partial x}\underline{F}(t,x) = \underline{\beta}(t)\boldsymbol{D_1}\boldsymbol{G}(x-R), \tag{5.25}$$

$$\frac{\partial}{\partial t}\underline{\beta}(t) = -\frac{\partial}{\partial t}\underline{F}(t,R) + \underline{\beta}(t)(\boldsymbol{D_0}+\boldsymbol{D_1}\boldsymbol{Z}). \tag{5.26}$$

Proof. To prove the theorem, we describe the evolution of $\mathcal{D}(t)$ in an infinitesimally small time period $(t, t+\Delta)$. Since we have a Markovian arrival process, the events to consider must include the phase transitions, too, both those that generate an arrival and those that do not generate any. There are two possibilities leading to $\mathcal{D}(t+\Delta) > x$ and $\widehat{\mathcal{J}}(t+\Delta) = i$, for $x > R$:

- At time t, there was an informed cluster on the highway already, which moved towards the accident, by a distance of $v\Delta$. Since $\widehat{\mathcal{J}}(t)$ is the phase at the end of the current informed cluster, we have that $\widehat{\mathcal{J}}(t+\Delta) = \widehat{\mathcal{J}}(t)$ in this case (as the current cluster remained the same).

- Or, at time t the system was in an uninformed interval ($\mathcal{D}(t) = R$) in phase $\mathcal{J}(t) = j$, and the MAP has generated an arrival (a vehicle has entered the range of the accident) in $(t, t+\Delta)$, forming a new cluster of informed vehicles. The probability of an arrival in $(t, t+\Delta)$ accompanied by a phase transition from j to k is $[\boldsymbol{D_1}]_{jk}\Delta + o(\Delta)$, and the length of the cluster ending in phase i is given by ccdf $G_{ki}(x-R)$.

The probability of having multiple events in $(t, t+\Delta)$ is $o(\Delta)$, for which $\lim_{\Delta\to 0} o(\Delta)/\Delta = 0$ holds. Based on these two possibilities we have that

$$F_i(t+\Delta, x) = F_i(t, x+v\Delta) + \sum_{k=1}^{N}\sum_{j=1}^{N}\beta_j(t)[\boldsymbol{D_1}]_{jk}\Delta G_{ki}(x-R) + o(\Delta).$$

With simple algebraic manipulations we get

$$\frac{F_i(t+\Delta,x)-F_i(t,x)}{\Delta} - v\frac{F_i(t,x+v\Delta)-F_i(t,x)}{v\Delta} = \sum_{k=1}^{N}\sum_{j=1}^{N}\beta_j(t)[\boldsymbol{D_1}]_{jk}G_{ki}(x-R) + o(\Delta)/\Delta,$$

which, tending Δ to 0 and switching to matrix notations, is equal to (5.25).

To derive (5.26), we have to investigate how it is possible to be in an uninformed interval and phase i at time $t+\Delta$. There are three cases:

- The system was in an uninformed interval and phase i at time t already, and no phase transitions occurred in the MAP. The probability of the latter event is $1+[\boldsymbol{D_0}]_{ii}\Delta+o(\Delta)$.
- The system was in an uninformed interval and phase j at time t, and there was an internal phase transition in the MAP from phase j to i, meaning that no vehicles arrived to the range of the accident, with probability $[\boldsymbol{D_0}]_{ji}\Delta+o(\Delta)$.
- At time t, there was a cluster of informed vehicles on the highway (ending with phase i), which has left the accident in $(t,t+\Delta)$. Hence, a new uninformed interval begins at time $t+\Delta$. The probability of this event is $F_i(t,R)-F_i(t,R+v\Delta)$.

Putting all parts together leads to

$$\beta_i(t+\Delta)=\beta_i(t)(1+[\boldsymbol{D_0}]_{ii}\Delta)+F_i(t,R)-F_i(t,R+v\Delta)+\sum_{j=1,j\neq i}^{N}\beta_j(t)[\boldsymbol{D_0}]_{ji}\Delta,$$

that can be transformed to

$$\frac{\beta_i(t+\Delta)-\beta_i(t)}{\Delta}=v\frac{F_i(t,R)-F_i(t,R+v\Delta)}{v\Delta}+\sum_{j=1}^{N}\beta_j(t)[\boldsymbol{D_0}]_{ji},$$

which, after taking the limit $\Delta\to 0$ and using matrices yields

$$\frac{\partial}{\partial t}\underline{\beta}(t)=-v\frac{\partial}{\partial x}\underline{F}(t,x)|_{x=R+}+\underline{\beta}(t)\boldsymbol{D_0}.$$

Observing that the derivative with regards to x on the right-hand side can be expressed using the time derivative based on (5.25) establishes the theorem. More proof details can found in Appendix A.1.4. □

5.4.2 *Stationary analysis*

The stationary behavior can be obtained by taking the limit $\mathcal{D}=\lim_{t\to\infty}\mathcal{D}(t)$. The following theorem provides the stationary solution of the phase dependent ccdf of the information distance, denoted by $\underline{F}(x)=\lim_{t\to\infty}\underline{F}(t,x)$, using the stationary phase dependent ccdf of the cluster length $\boldsymbol{G}(x)$.

Theorem 19. *For $x>R$, the stationary phase dependent ccdf of the information distance is given by*

$$\frac{d}{dx}\underline{F}(x)=-\frac{1}{E(\mathcal{G})+E(\mathcal{H})}\underline{\gamma}\boldsymbol{G}(x-R), \tag{5.27}$$

and, for $x=R$, the stationary probability vector of an uninformed interval is the solution of the system of linear equations

$$\underline{\beta}(\boldsymbol{D_0}+\boldsymbol{D_1}\boldsymbol{Z})=\underline{0},\quad \underline{\beta}\mathbb{1}=\frac{E(\mathcal{H})}{E(\mathcal{G})+E(\mathcal{H})}. \tag{5.28}$$

Proof. Taking the limit $t \to \infty$ in (5.25) gives

$$-v\frac{d}{dx}\underline{F}(x) = \underline{\beta}\boldsymbol{D_1}\boldsymbol{G}(x-R), \tag{5.29}$$

and taking the limit in (5.26) leads to

$$\underline{0} = \underline{\beta}(\boldsymbol{D_0}+\boldsymbol{D_1}\boldsymbol{Z}).$$

In the latter equation, observe that matrix $\boldsymbol{D_0}+\boldsymbol{D_1}\boldsymbol{Z}$ defines a continuous-time Markov chain, that characterizes the phase process of the MAP in uninformed intervals. The MAP evolves according to matrix $\boldsymbol{D_0}$, and whenever a vehicle arrives to the accident (matrix $\boldsymbol{D_1}$), the phases are immediately adjusted by matrix $\boldsymbol{Z}$ to reflect the end-of-cluster phases. The stationary solution of this Markov chain provides the phase distribution in the uninformed intervals. From the definition of $E(\mathcal{H})$ and $E(\mathcal{G})$, the stationary (phase-independent) probability of the uninformed intervals, $\underline{\beta}\mathbb{1}$, is $E(\mathcal{H})/(E(\mathcal{G})+E(\mathcal{H}))$, proving (5.28).

On the other hand, we can observe that vector $\underline{\beta}\boldsymbol{D_1}$ is proportional to the phase distribution right at the end of the uninformed intervals when a new cluster gets formed. The same phase distribution is given by vector $\underline{\gamma}$, too, the only difficulty is to find the scaling factor between them, that normalizes $\underline{\beta}\boldsymbol{D_1}$. Since exactly one cluster is formed at the end of each uninformed period (that lasts for $E(\mathcal{H})/v$ in time), we have that the scaling factor is $v/E(\mathcal{H})$, thus, (5.29) becomes

$$\frac{d}{dx}\underline{F}(x) = -\underline{\beta}\boldsymbol{D_1}\boldsymbol{G}(x-R)/v = \frac{E(\mathcal{H})}{E(\mathcal{G})+E(\mathcal{H})}\frac{1}{E(\mathcal{H})}\underline{\gamma}\boldsymbol{G}(x-R),$$

that proves (5.27). □

Theorem 20. *The phase independent ccdf of the information distance, $F(x) = \underline{F}(x)\mathbb{1}$, can be expressed by*

$$F(x) = -\frac{\int_0^{x-R} G(y)dy}{E(\mathcal{G})+E(\mathcal{H})} + \frac{E(\mathcal{G})}{E(\mathcal{G})+E(\mathcal{H})}, \tag{5.30}$$

for $x > R$. For $x \le R$ we have that $F(x) = 1$.

Proof. Multiplying (5.27) by $\mathbb{1}$ from the right leads to scalar ODE

$$\frac{d}{dx}F(x) = -\frac{G(x-R)}{E(\mathcal{G})+E(\mathcal{H})}.$$

Taking the integral of both sides from $y = 0$ to ∞ gives

$$F(y) = -\frac{\int_{x=0}^{y} G(x-R)dx}{E(\mathcal{G})+E(\mathcal{H})} + c,$$

where, from $F(R) = P(\mathcal{D} > R)$, for the constant we have that $c = E(\mathcal{G})/(E(\mathcal{G})+E(\mathcal{H}))$ (see Figure 5.4), yielding (5.30). □

Theorem 17 provides a way to calculate the mean information distance $E(\mathcal{D})$. The same quantity can be derived from Theorem 20 as well, if we take the integral of (5.30):

$$\underbrace{\int_{x=R}^{\infty} F(x)\,dx}_{E(\mathcal{D})-R} = \frac{1}{E(\mathcal{G})+E(\mathcal{H})} \underbrace{\left(\int_R^{\infty}\left(-\int_0^{x-R} G(y)\,dy + E(\mathcal{G})\right)dx\right)}_{(*)}, \tag{5.31}$$

where the integral term in the parenthesis simplifies to

$$\begin{aligned}(*) &= \int_R^{\infty}\left(-\int_0^{x-R} G(y)\,dy + \int_0^{\infty} G(y)\,dy\right)dx = \int_{x=R}^{\infty}\int_{y=x-R}^{\infty} G(y)\,dy\,dx \\ &= \int_{y=0}^{\infty}\int_{x=R}^{y+R} G(y)\,dx\,dy = \int_{y=0}^{\infty} yG(y)\,dy = E(\mathcal{G}^2)/2,\end{aligned}$$

which is in line with (5.21).

5.4.3 *Speed of the information propagation*

Making use of the transient distribution, another interesting study can be carried out: the analysis of the speed of the information propagation. From the ccdf $F(t,x)$, the mean information distance at time t can be obtained as

$$E(\mathcal{D}(t)) = R + \int_R^{\infty} F(t,x)\,dx. \tag{5.32}$$

Hence, taking the integral of both sides of (5.25), and multiplying it by $\mathbb{1}$ from the right, we get that

$$\frac{d}{dt}(E(\mathcal{D}(t))-R) - v\underbrace{[F(t,x)]_R^{\infty}}_{-1} = \underline{\beta}(t)\boldsymbol{D_1}\boldsymbol{E}(\mathcal{G})\mathbb{1},$$

which leads to

$$\frac{d}{dt}E(\mathcal{D}(t)) = \underline{\beta}(t)\boldsymbol{D_1}\boldsymbol{E}(\mathcal{G})\mathbb{1} - v. \tag{5.33}$$

Equation (5.33) is easy to interpret and is completely reasonable. The term $-v$ means that the informed cluster moves towards the accident at speed v; hence, the information distance decreases at this speed. At the other hand, $\underline{\beta}(t)\boldsymbol{D_1}$ is the rate of new cluster formation (the rate of a vehicle arrival in an uninformed period), when the information distance jumps up to $\boldsymbol{E}(\mathcal{G})\mathbb{1}$.

The negativity of $\frac{d}{dt}E(\mathcal{D}(t))$ is not obvious from the formula. Still, it follows from the behavior of the system, namely that in case of constant speed, only a single informed cluster can exist, and a new one can be formed only after the previous one has left the accident.

5.5 NUMERICAL EXAMPLES

In this section, we investigate the impact of the statistics of the vehicle inter-arrival times on the cluster size $\mathcal{G}$ and the information distance $\mathcal{D}(t)$. Our implementation is based on Matlab, and we use the BuTools library [28] for obtaining the MAPs. To validate the correctness of the presented analytical methods, we have developed a custom simulation tool that gives the exact same results in all of the studied cases.

In all of the numerical examples the vehicle speed is assumed to be $v = 36m/s$ (which is the typical speed limitation on highways in many countries), and the radio coverage of the communication is $R = 150m$.

In Sections 5.5.1 and 5.5.2 we demonstrate the importance of using appropriate traffic models by showing the impact of the squared coefficient of variation and the lag-1 correlation coefficient on the mean cluster length and the mean information distance. Section 5.5.3 studies the transient distribution of the message propagation distance, and Section 5.5.4 presents some results with real traffic data as well.

5.5.1 *Analysis of the cluster length*

We have used the procedure of [25] to obtain matrices $\boldsymbol{D_0}$ and $\boldsymbol{D_1}$ from the mean arrival rate, the SCV and the lag-1 correlation, $\hat{\rho}$, of the headway times. With several combinations of these statistical parameters, the mean and the SCV of the cluster length $\mathcal{G}$ were computed based on Corollaries 4 and 5.

Setting the mean arrival rate to a low value leads to very short clusters, that mostly consist of isolated vehicles. On the other hand, by setting it too high, almost all vehicles will be connected into a single cluster. We found that $\lambda = 0.65$ (vehicles/second) represents a car density that falls between the two extreme cases, through which the cluster length is worth examining. The SCV of the headway times has been varied in the range of 0.5 (close to deterministic) to 5 (very bursty), where $SCV = 1$ corresponds to the case of the Poisson vehicle arrival process. As for the lag-1 correlation parameter ρ, we studied the effect of no correlation ($\hat{\rho} = 0$), negative correlation (strong negative correlation $\hat{\rho} = -0.3$ and moderate negative correlation $\hat{\rho} = -0.19$) and positive correlation (strong positive correlation $\hat{\rho} = 0.85$ and moderate positive correlation $\hat{\rho} = 0.19$) as well. Of course, our procedure can handle the full range of these parameters, i.e. $SCV > 0$ and $-1 < \hat{\rho} < 1$, the only limitation being that $\hat{\rho} > -1/SCV$ must be respected[25]. High negative correlation is more difficult to achieve with the Markovian arrival process (MAP), so in the forthcoming examples we could not set negative correlation with the same absolute value as in case of positive correlation.

The results are shown in Figure 5.5. As visible from the Figure, the statistics of the vehicle arrival process has a significant impact on the cluster length statistics. In general, the lower the SCV is, the longer the clusters are, since the variability of the headway times is lower. The combination of high correlation and low SCV leads to the longest clusters, in this case the headway times of many subsequent vehicles are almost identical. When the SCV is higher, the difference between the curves is smaller, but still significant (observe that the scale of the y axis is logarithmic). According to the plot on

the right side of Figure 5.5, the squared coefficient of variation of the cluster length also depends on the SCV of the headway times.

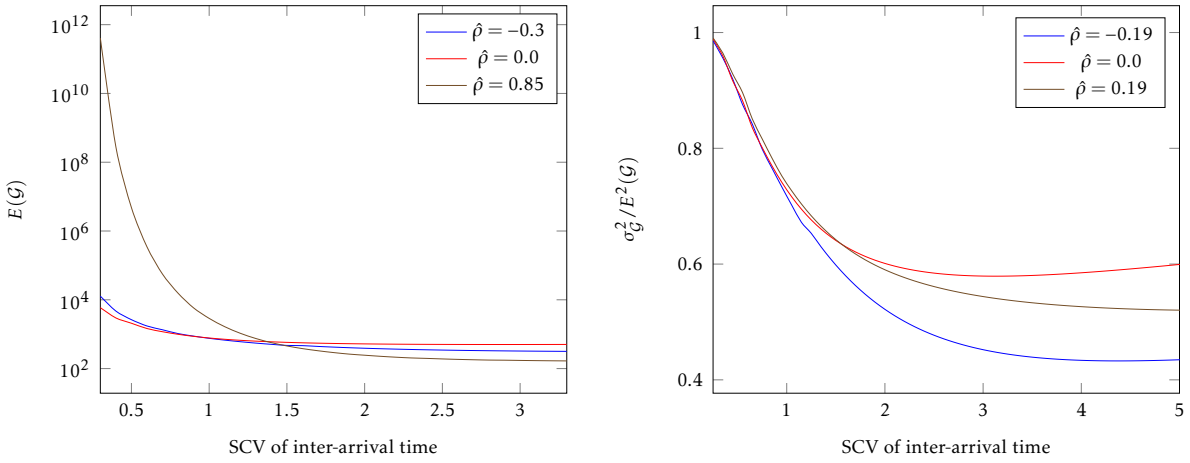


Figure 5.5: The mean and the SCV of cluster length $\mathcal{G}$

We note that there are some short intervals in Figure 5.5 where the plots are not perfectly smooth. The reason for this effect is that the procedure we used for creating the MAPs [25] has changed the size of the MAPs around these points.

The numerical solution of the DDE defined by Theorem 16 makes it is possible to investigate the ccdf of the cluster length as the function of the SCV as well. According to Figure 5.6, the cluster length is always greater than or equal to R, due to the definition of the system. It is clearly visible that the burstiness of the vehicle flow has a significant impact on the length of clusters of informed vehicles. When SCV is small (more regular headway times), the informed clusters are longer, while higher SCV leads to shorter clusters.

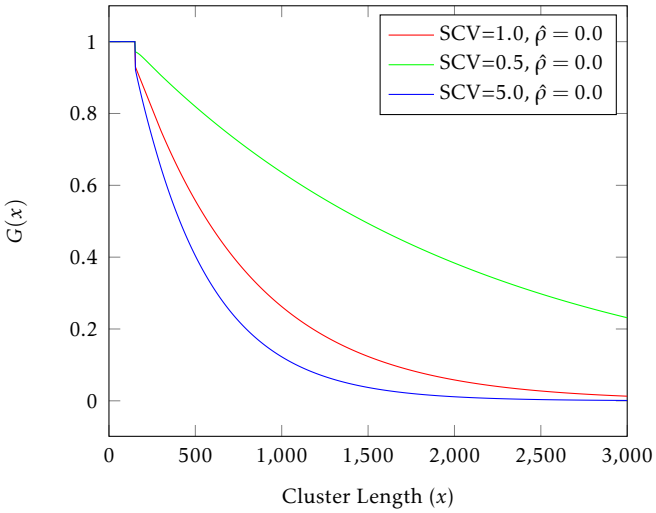


Figure 5.6: The ccdf of the cluster length distribution $G(x)$ with various SCV parameters

5.5.2 *Analysis of the information distance*

The main target of the chapter is the analysis of the information distance, that is, how far a message (e.g., an alert message) can propagate in the stationary state. On the left side of Figure 5.7, the mean information distance is depicted assuming different SCV and correlation values, based on Theorem 17. As known from earlier results [41], the mean information distance $E(\mathcal{D})$ and the mean cluster length $E(\mathcal{G})$ are equal for the Poisson case, when $SCV = 1$ and $\hat{\rho} = 0$. In all the other cases, $E(\mathcal{D}) = E(\mathcal{G})$ does not hold. Nevertheless, comparing Figure 5.5 and Figure 5.7 it is clear that the difference between $E(\mathcal{D})$ and $E(\mathcal{G})$ is small.

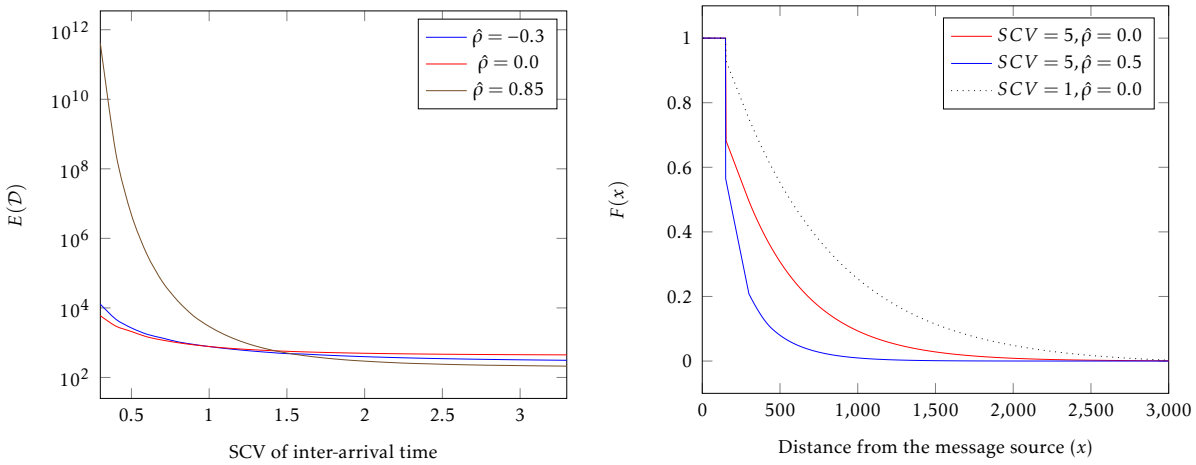


Figure 5.7: The mean and the ccdf of the stationary message propagation distance $\mathcal{D}$

According to the right side of Figure 5.7 higher SCV of the headway time leads to lower message propagation distance, and high correlation decreases the message propagation distance even more. The plots were obtained by the numerical solution of the ODE defined in Theorem 20.

5.5.3 *Transient analysis*

The results of Section 4.5 make it possible to investigate the time-evolution of the message propagation distance, $\mathcal{D}(t)$. As a demonstration, we study a case when an accident occurs at time $t = 0$, when the state of the MAP is stationary. We assume that no vehicles are informed about the accident initially, hence we have $\underline{\beta}(0) = \underline{\alpha}$ and $\underline{F}(0,x) = \underline{0}, x > R$. From this starting point the ccdf $F(t,x) + \beta(t)$ can be obtained by the numerical solution of the ODE given in Theorem 18. (For the solution the step size along the time axis was 1, and along the distance axis it was $1/v$).

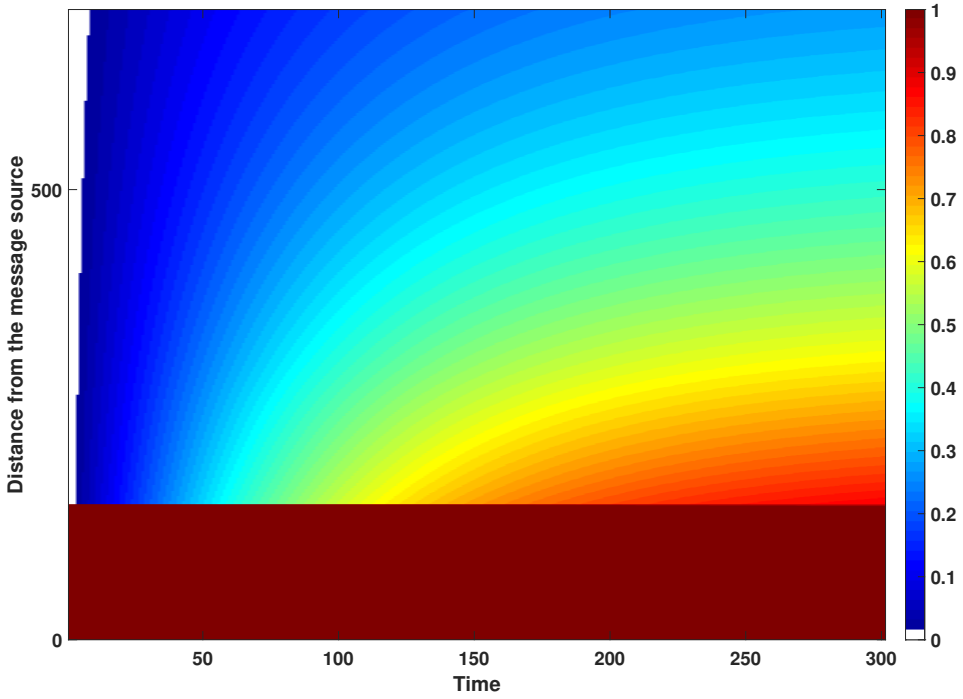


Figure 5.8: Transient distribution, $F(t,x)$

The results are visualized in Figure 5.8 as a heat map. In line with the expectations, the information distance is always at least R, and the more time elapses since the accident, the higher the probability is that vehicles farther away from the accident receive the message about it. After $t = 300$ (that is, 5 minutes) the stationary state is almost reached, the distribution of $\mathcal{D}(t)$ does not change significantly.

5.5.4 *Experiments with real data*

We did experiments with real data as well. Two factors make such a study difficult.

The first difficulty is that there are very few high-quality headway data traces available publicly. The majority of traffic data contains counts over a certain period of time (e.g., number of vehicles detected in an hour), which is not suitable for MAP fitting. For our procedure, we need the exact arrival times of the vehicles (or equivalently, all headway times). The only relevant data set we found was [12], which was based on LIDAR measurements. While this is a fairly large data set, it is still not long enough, since the treatment of such seasonal traffic measurements needs a lot of data. Hence, we decided to ignore the seasonal nature of the traffic and cut out a part of the data consisting of around ≈ 629000 samples where the traffic was approximately stationary.

The second difficulty is that fitting MAPs needs a lot of data. Capturing the characteristics of the density function of the marginal distribution can be accurate with less data, too, but for correlation fitting, especially for higher lags, much more data is needed. This observation is reflected by Figure 5.1, too, where the density is fitted well, and the lag-correlations are not matching as well. As more headway time data become publicly available, it will be possible to fit MAPs that better represent the real traffic, making our analytical model more accurate.

Table 5.1: Experiments with real data

Method \ Metric	$E(\mathcal{G})$	$E(\mathcal{G}^2)$	$E(\mathcal{H})$	$E(\mathcal{D})$
Simulation	495	8.19134+6	290	5375
Poisson Model	248.6306	8.0594e+04	159.9331	248.6306
MAP Model	489.4438	5.9635e+06	282.2731	4013

Based on the ≈ 629000 samples extracted from the data set [12], we executed the EM-algorithm published in [27] to create a MAP with 800 states. Even with that many states, the formulas presented in this chapter give instant results. Table 5.1 compares these analytical results with the simulation results driven by the original measurement data. According to the results, the mean cluster length $E(\mathcal{G})$ is obtained very accurately; the error (the deviation from the simulation results) is below 2%. The Poisson assumption, commonly used in the literature [41, 43], gives almost 100% error. The same holds for the mean non-informed periods $E(\mathcal{H})$ as well. However, for the second moment of the cluster length $E(\mathcal{G}^2)$, our method has a significant error, due to the imperfect MAP fitting caused by the overly small data set. Still, the MAP model-based results are not far away from the simulation results, as opposed to the Poisson model-based results, where there is a 100-times difference. The inaccuracy in $E(\mathcal{G}^2)$ implies inaccuracy in $E(\mathcal{D})$, too. Our method gives 4013 meters for the mean message propagation distance instead of 5375 meters, but it is still much better than the Poisson result with 248.6 meters.

We believe that, as more data gets available and more mature MAP fitting methods get developed, the practical relevance of our procedure is going to improve in the future.

5.6 SUMMARY

The behavior of the message propagation in VANET systems is affected by the traffic model of the vehicles. In this chapter, we consider the Markovian Arrival Process as a traffic model and derive various results related to message propagation. We have derived the moments and the ccdf of the stationary cluster length, and the stationary and transient distribution of the information distance. We validated our analytical results with simulation. We conclude that, when the Poisson assumption fails to model the traffic, the Markov arrival process can be a remedy to approximate the message propagation distance in practice.

CHAPTER SIX

6

SUMMARY OF RESULTS AND FUTURE WORK

6.1 SUMMARY OF RESULTS

This dissertation focused on the self-organized network technology in two main communication system areas: Queueing systems for network devices and Vehicular Ad-hoc Networks. Both of the fields, as mentioned above, have many modeling challenges. My proposed methods include analytical algorithms and numerical methods, whose accuracy is verified by simulation. The new results can be summarized as follows:

- Explicit approximation formula for the mean response times in the two-class weighted fair queueing system. We developed the analytical method, compared the results with simulation, and proved that the proposed solution is better than the published method found in the literature.
- We introduced new contributions regarding the alert message propagation in VANET systems. Stationary and transient solutions are present in this solution as new results. We validated the analytical results with simulation as well.
- We derive the asymptotic speed of the alert message propagation in VANETs with disconnected RSUs. Different new results are proposed as the speed of message propagation and the transient analysis of message propagation as well. The proposed results can be used in network planning of VANETs.
- The stochastic properties of the arrivals process play important roles on the message propagation in VANET systems. The new results consider the Markovian Arrival Process as a traffic model and we derived different results related to message propagation: the moments and the ccdf of the stationary cluster length, and the stationary and transient distribution of the information distance.

Moreover, we proposed using an M/D/1 queueing model for the traffic jam at the accident to get an improved approximation for the message propagation distance.

On the other hand, we made experiments with simulation both with synthetic and real data to prove that the MAP based vehicle arrival process leads to more accurate results.

6.2 FUTURE WORK

The telecommunication systems and Vehicular Ad-hoc Networks are rapidly changing to meet user requirements. Many open issues should be considered before implementing a telecommunication system in the practice. My dissertation deals with some areas of these challenges. We can briefly describe the future work of my research as follows:

- *Adapt the algorithm or develop new algorithms* for more complex packet schedulers and more complex traffic patterns for the schedulers of telecommunication devices (i.e. routers).
- Consider the possibility of *dynamic class switching* in WFQ. Investigate the feasibility of the algorithms in real systems such as 5G devices.
- *Consider the counter-flow* traffic for the alert message propagation model in VANET.
- Extend the research work to the *urban environment* within VANET, not only in highways.
- Collect *more realistic vehicle arrival data* based on a detector loop or video camera. Fitting model based on the large data can give more accurate results.
- *Build a tool based on mathematical result* to support the decisions of road operators.

The new protocols and algorithms are important parts of the alert message propagation in VANET technology, in the future we adapt these protocols to meet the quality of service in intelligent transportation system. As shown in many research work [15], there are different types of services provided through VANET communication systems. Some of these services have low priority compared to emergency message, so in the further research we can consider the *Delay Tolerant Network (DTN)* paradigm to deal with other services such as comfort application (i.e. journey time estimation).

6.3 PUBLICATION

6.3.1 *International Journals and conferences (Peer-reviewed)*

- International Journals

[J1] **Mahmood, Dhari Ali**, and Gábor Horváth. "Analysis of the Message Propagation on the Highway in VANET." Arabian Journal for Science and Engineering 44.4 (2019): 3405-3413. (WOS, IF=1.711, Q2)

[J2] **Mahmood, Dhari Ali**, and Gábor Horváth. Analysis of the Message Propagation Speed in VANET with Disconnected RSUs. Mathematics 2020, 8, 782. (WOS, IF=1.747, Q3)

[J3] **Mahmood, Dhari Ali**, and Gábor Horváth. Analysis of alert message propagation on the highway in VANET assuming Markovian vehicle arrival process. International Journal of Communication Systems 2020. (WOS, IF=1.319, Q2)

- **International Conferences**

[C1] **Mahmood, Dhari Ali**, and Gábor Horváth. "A simple approximation for the response times in the two-class weighted fair queueing system." International Conference on Analytical and Stochastic Modeling Techniques and Applications. Springer, Cham, 2017. (WOS, Scopus, CiteScore=1.9)

6.3.2 *Other own publication (Peer-reviewed)*

[J4] **MAHMOOD, DHARI ALI**, and RAHUL JOHARI. "Application of Routing Metrics in Wireless Network." International Journal of Engineering Science and Innovative Technology (IJESIT) Volume 2, Issue 4, July (2013).

[C2] Johari, Rahul, and **Dhari Ali Mahmood**. "MeNDARIN: Mobile Education Network Using DTN Approach in North IRAQ." Proceedings of the International Conference on Internet of things and Cloud Computing. 2016.

[C3] Johari, Rahul, and **Dhari A. Mahmood**. "GA-LORD: Genetic Algorithm and LTPCL-Oriented Routing Protocol in Delay Tolerant Network." Wireless Communications, Networking and Applications. Springer, New Delhi, 2016. 141-154.

[C4] Johari, Rahul, and **Dhari Ali Mahmood**. "GAACO: Metaheuristic driven approach for routing in OppNet." 2014 Global Summit on Computer and Information Technology (GSCIT). IEEE, 2014.

[C5] **Mahmood, Dhari Ali**, and Rahul Johari. "Routing in MANET using cluster based approach (RIMCA)." 2014 International Conference on Computing for Sustainable Global Development (INDIACom). IEEE, 2014.

- **Other publication**

[C6] Reja, Ahmed Hameed, Syed Naseem Ahmad, and **Dhari Ali Mahmood**. "Study the effect of adding new components on conventional microstrip LPF design." 2014 International Conference on Computing for Sustainable Global Development (INDIACom). IEEE, 2014.

BIBLIOGRAPHY

[1] Atef Abdrabou and Weihua Zhuang. "Probabilistic delay control and road side unit placement for vehicular ad hoc networks with disrupted connectivity." In: *IEEE Journal on Selected Areas in Communications* 29.1 (2010), pp. 129–139.

[2] Amina Al-Sawaai, Irfan Awan, and Rod Fretwell. "Analysis of the weighted fair queuing system with two classes of customers with finite buffer." In: *Advanced Information Networking and Applications Workshops, 2009. WAINA'09. International Conference on*. IEEE. 2009, pp. 218–223.

[3] Attahiru Sule Alfa. *Queueing theory for telecommunications: discrete time modelling of a single node system*. Springer Science & Business Media, 2010.

[4] Attahiru Sule Alfa and Marcel F Neuts. "Modelling vehicular traffic using the discrete time Markovian arrival process." In: *Transportation Science* 29.2 (1995), pp. 109–117.

[5] GG Md Nawaz Ali, Peter Han Joo Chong, Syeda Khairunnesa Samantha, and Edward Chan. "Efficient data dissemination in cooperative multi-RSU vehicular ad hoc networks (VANETs)." In: *Journal of Systems and Software* 117 (2016), pp. 508–527.

[6] Baber Aslam and Cliff C Zou. "Optimal roadside units placement along highways." In: *2011 IEEE Consumer Communications and Networking Conference (CCNC)*. IEEE. 2011, pp. 814–815.

[7] Søren Asmussen and Ger Koole. "Marked point processes as limits of Markovian arrival streams." In: *Journal of Applied Probability* 30.2 (1993), pp. 365–372.

[8] AV Babu and VK Muhammed Ajeer. "Analytical model for connectivity of vehicular ad hoc networks in the presence of channel randomness." In: *International Journal of Communication Systems* 26.7 (2013), pp. 927–946.

[9] Yuanguo Bi, Hangguan Shan, Xuemin Sherman Shen, Ning Wang, and Hai Zhao. "A multi-hop broadcast protocol for emergency message dissemination in urban vehicular ad hoc networks." In: *IEEE Transactions on Intelligent Transportation Systems* 17.3 (2016), pp. 736–750.

[10] Gunter Bolch, Stefan Greiner, Hermann de Meer, and Kishor S Trivedi. *Queueing networks and Markov chains: modeling and performance evaluation with computer science applications*. John Wiley & Sons, 2006.

[11] Giuliano Casale, Eddy Z Zhang, and Evgenia Smirni. "KPC-toolbox: Simple yet effective trace fitting using Markovian arrival processes." In: *2008 Fifth International Conference on Quantitative Evaluation of Systems*. IEEE. 2008, pp. 83–92.

[12] Benjamin Coifman and Lizhe Li. "A critical evaluation of the Next Generation Simulation (NGSIM) vehicle trajectory dataset." In: *Transportation Research Part B: Methodological* 105 (2017), pp. 362–377.

[13] S Corson and Joseph Macker. *RFC2501: Mobile ad hoc networking (MANET): Routing protocol performance issues and evaluation considerations*. 1999.

[14] David Roxbee Cox. *Renewal theory*. Methuen, 1962.

[15] Felipe Cunha, Leandro Villas, Azzedine Boukerche, Guilherme Maia, Aline Viana, Raquel AF Mini, and Antonio AF Loureiro. "Data communication in VANETs: Protocols, applications and challenges." In: *Ad Hoc Networks* 44 (2016), pp. 90–103.

[16] Csaba Farkas and Miklós Telek. "Capacity planning of electric car charging station based on discrete time observations and MAP(2)/G/c queue." In: *Periodica Polytechnica Electrical Engineering and Computer Science* 62.3 (2018), pp. 82–89.

[17] Guy Fayolle and Roudolf Iasnogorodski. "Two coupled processors: the reduction to a Riemann-Hilbert problem." In: *Probability Theory and Related Fields* 47.3 (1979), pp. 325–351.

[18] Manuel Fogue, Julio A Sanguesa, Francisco J Martinez, and Johann M Marquez-Barja. "Improving Roadside Unit deployment in vehicular networks by exploiting genetic algorithms." In: *Applied Sciences* 8.1 (2018), p. 86.

[19] Harald T. Friis. "A Note on a Simple Transmission Formula." In: *Proceedings of the IRE* 34.5 (1946), pp. 254–256. ISSN: 00968390.

[20] S Jamaloddin Golestani. "A self-clocked fair queueing scheme for broadband applications." In: *INFOCOM'94. Networking for Global Communications., 13th Proceedings IEEE*. IEEE. 1994, pp. 636–646.

[21] Sebastian Gräfling, Petri Mähönen, and Janne Riihijärvi. "Performance evaluation of IEEE 1609 WAVE and IEEE 802.11p for vehicular communications." In: *Ubiquitous and Future Networks (ICUFN), 2010 Second International Conference on*. IEEE. 2010, pp. 344–348.

[22] F Guillemin and D Pinchon. "Analysis of the weighted fair queuing system with two classes of customers with exponential service times." In: *Journal of Applied Probability* (2004).

[23] Jerome Harri, Fethi Filali, and Christian Bonnet. "Mobility models for vehicular ad hoc networks: a survey and taxonomy." In: *IEEE Communications Surveys & Tutorials* 11.4 (2009), pp. 19–41.

[24] G. Horváth and M. Telek. "An approximate analysis of two class WFQ systems." In: *Workshop on Preformability Modeling of Computer and Communication Systems-PMCCS*. Citeseer. 2003, pp. 43–46.

[25] Gábor Horváth. " Marginal Moments and Lag Autocorrelations with MAPs." In: *Proceedings of the 7th International Conference on Performance Evaluation Methodologies and Tools* (2013), pp. 59–68.

[26] Gábor Horváth. "Matching marginal moments and lag autocorrelations with MAPs." In: *Proceedings of the 7th International Conference on Performance Evaluation Methodologies and Tools*. 2013, pp. 59–68.

[27] Gábor Horváth and Hiroyuki Okamura. "A fast EM algorithm for fitting marked Markovian arrival processes with a new special structure." In: *European Workshop on Performance Engineering*. Springer. 2013, pp. 119–133.

[28] Gábor Horváth and Miklós Telek. "BuTools 2: A rich toolbox for Markovian performance evaluation." In: *proceedings of the 10th EAI International Conference on Performance Evaluation Methodologies and Tools*. 2017, pp. 137–142.

[29] Gábor Horváth, Illés Horváth, Salah Al-Deen Almousa, and Miklós Telek. "Numerical inverse Laplace transformation using concentrated matrix exponential distributions." In: *Performance Evaluation* 137 (2020), p. 102067.

[30] Amine Kchiche and Farouk Kamoun. "Centrality-based access-points deployment for vehicular networks." In: *2010 17th International Conference on Telecommunications*. IEEE. 2010, pp. 700–706.

[31] Mehdi Khabazian and Mustafa K Mehmet Ali. "A performance modeling of connectivity in vehicular ad hoc networks." In: *IEEE Transactions on Vehicular Technology* 57.4 (2008), pp. 2440–2450.

[32] Caglar Kosun and Serhan Ozdemir. "A superstatistical model of vehicular traffic flow." In: *Physica A: statistical mechanics and its applications* 444 (2016), pp. 466–475.

[33] Wolfgang Kraemer and Manfred Langenbach-Belz. "Approximate formulae for the delay in the queueing system GI/G/1." In: *Proceedings of the 8th International Teletraffic Congress*. 1976, pp. 235–1.

[34] Daniel Krajzewicz, Jakob Erdmann, Michael Behrisch, and Laura Bieker. "Recent Development and Applications of SUMO - Simulation of Urban MObility." In: *International Journal On Advances in Systems and Measurements* 5.3&4 (2012), pp. 128–138.

[35] Laszlo Lakatos, Laszlo Szeidl, and Miklos Telek. *Introduction to queueing systems with telecommunication applications*. Springer, 2019.

[36] Guy Latouche and Vaidyanathan Ramaswami. *Introduction to matrix analytic methods in stochastic modeling*. Vol. 5. Siam, 1999.

[37] Pan Li, Xiaoxia Huang, Yuguang Fang, and Phone Lin. "Optimal placement of gateways in vehicular networks." In: *IEEE Transactions on Vehicular Technology* 56.6 (2007), pp. 3421–3430.

[38] Peng Li, Qin Liu, Chuanhe Huang, Jinhai Wang, and Xiaohua Jia. "Delay-bounded minimal cost placement of roadside units in vehicular ad hoc networks." In: *2015 IEEE International Conference on Communications (ICC)*. IEEE. 2015, pp. 6589–6594.

[39] Thomas DC Little, Ashish Agarwal, et al. "An information propagation scheme for VANETs." In: *Proc. IEEE Intelligent Transportation Systems*. 2005, pp. 155–160.

[40] Chunyan Liu, Hejiao Huang, and Hongwei Du. "Optimal RSUs deployment with delay bound along highways in VANET." In: *Journal of Combinatorial Optimization* 33.4 (2017), pp. 1168–1182.

[41] Dhari Ali Mahmood and Gábor Horváth. "Analysis of the Message Propagation on the Highway in VANET." In: *Arabian Journal for Science and Engineering* 44.4 (2019), pp. 3405–3413. ISSN: 21914281.

[42] Khaleel Mershad, Hassan Artail, and Mario Gerla. "ROAMER: Roadside Units as message routers in VANETs." In: *Ad Hoc Networks* 10.3 (2012), pp. 479–496.

[43] Daniele Miorandi and Eitan Altman. "Connectivity in one-dimensional ad hoc networks: a queueing theoretical approach." In: *Wireless Networks* 12.5 (2006), pp. 573–587.

[44] Nazmus Shaker Nafi, M. K. Hasan, and Aisha H. Abdallah. "Traffic flow model for vehicular network." In: *2012 International Conference on Computer and Communication Engineering, ICCCE 2012* July (2012), pp. 738–743.

[45] Marcel F Neuts. "A versatile Markovian point process." In: *Journal of Applied Probability* 16.4 (1979), pp. 764–779.

[46] Seh Chun Ng, Wuxiong Zhang, Yang Yang, and Guoqiang Mao. "Analysis of access and connectivity probabilities in infrastructure-based vehicular relay networks." In: *2010 IEEE Wireless Communication and Networking Conference*. IEEE. 2010, pp. 1–6.

[47] Hiroyuki Okamura and Tadashi Dohi. "Faster maximum likelihood estimation algorithms for Markovian arrival processes." In: *2009 Sixth international conference on the quantitative evaluation of systems*. IEEE. 2009, pp. 73–82.

[48] Andre B Reis, Susana Sargento, and Ozan K Tonguz. "On the performance of sparse vehicular networks with road side units." In: *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*. IEEE. 2011, pp. 1–5.

[49] Andre B Reis, Susana Sargento, Filipe Neves, and Ozan K Tonguz. "Deploying roadside units in sparse vehicular networks: What really works and what does not." In: *IEEE transactions on vehicular technology* 63.6 (2013), pp. 2794–2806.

[50] Jacques Resing and Lerzan ÖRmeci. "A tandem queueing model with coupled processors." In: *Operations Research Letters* 31.5 (2003), pp. 383–389.

[51] Giovanni Resta, Paolo Santi, and Janos Simon. "Analysis of multi-hop emergency message propagation in vehicular ad hoc networks." In: *Proceedings of the 8th ACM international symposium on Mobile ad hoc networking and computing*. ACM. 2007, pp. 140–149.

[52] Karim Rostamzadeh and Sathish Gopalakrishnan. "Analysis of Message Delivery Delay in Vehicular Networks." In: *IEEE Transactions on Vehicular Technology* 64.10 (2015), pp. 4770–4779. ISSN: 00189545.

[53] Karim Rostamzadeh and Sathish Gopalakrishnan. "Analysis of message delivery delay in vehicular networks." In: *IEEE Transactions on Vehicular Technology* 64.10 (2015), pp. 4770–4779.

[54] Hafez Seliem, Reza Shahidi, Mohamed Hossam Ahmed, and Mohamed S Shehata. "On the End-to-End Delay in a One-Way VANET." In: *IEEE Transactions on Vehicular Technology* 68.9 (2019), pp. 8336–8346.

[55] John F Shortle and Martin J Fischer. "Approximation for a two-class weighted fair queueing discipline." In: *Performance Evaluation* 67.10 (2010), pp. 946–958.

[56] Madhavapeddi Shreedhar and George Varghese. "Efficient fair queuing using deficit round-robin." In: *IEEE/ACM Transactions on networking* 4.3 (1996), pp. 375–385.

[57] Christoph Sommer, Reinhard German, and Falko Dressler. "Bidirectionally Coupled Network and Road Traffic Simulation for Improved IVC Analysis." In: *IEEE Transactions on Mobile Computing* 10.1 (2011), pp. 3–15.

[58] Sok-Ian Sou and Ozan K Tonguz. "Enhancing VANET connectivity through roadside units on highways." In: *IEEE transactions on vehicular technology* 60.8 (2011), pp. 3586–3602.

[59] Miklós Telek and Gábor Horváth. "A minimal representation of Markov arrival processes and a moments matching method." In: *Performance Evaluation* 64.9-12 (2007), pp. 1153–1168.

[60] Ozan K Tonguz and Wantanee Viriyasitavat. "Cars as roadside units: a self-organizing network solution." In: *IEEE Communications Magazine* 51.12 (2013), pp. 112–120.

[61] EM Van Eenennaam. "A survey of propagation models used in vehicular ad hoc network (vanet) research." In: *Paper written for course Mobile Radio Communication, University of Twente* 46 (2008).

[62] András Varga. "Using the OMNeT++ discrete event simulation system in education." In: *IEEE Transactions on Education* 42.4 (1999), 11–pp.

[63] András Varga. "Discrete event simulation system." In: *Proc. of the European Simulation Multiconference (ESM'2001)*. 2001, pp. 1–7.

[64] András Varga and Rudolf Hornig. "An overview of the OMNeT++ simulation environment." In: *Proceedings of the 1st international conference on Simulation tools and techniques for communications, networks and systems & workshops*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering). 2008, p. 60.

[65] Leandro Aparecido Villas, Azzedine Boukerche, Guilherme Maia, Richard Werner Pazzi, and Antonio AF Loureiro. "Drive: An efficient and robust data dissemination protocol for highway and urban vehicular ad hoc networks." In: *Computer Networks* 75 (2014), pp. 381–394.

[66] Christian Vitale, Gianluca Rizzo, Balaji Rengarajan, and Vincenzo Mancuso. "An Analytical Approach to Performance Analysis of Coupled Processor Systems." In: *Teletraffic Congress (ITC 27), 2015 27th International*. IEEE. 2015, pp. 89–97.

[67] K. T. Waldeer. "A Vehicular Traffic Flow Model Based on a Stochastic Acceleration Process." In: *Transport Theory and Statistical Physics* 33.1 (2004), pp. 7–30.

[68] Ward Whitt. "The queueing network analyzer." In: *Bell Labs Technical Journal* 62.9 (1983), pp. 2779–2815.

[69] Nawaporn Wisitpongphan, Fan Bai, Priyantha Mudalige, Varsha Sadekar, and Ozan Tonguz. "Routing in sparse vehicular ad hoc wireless networks." In: *IEEE journal on Selected Areas in Communications* 25.8 (2007), pp. 1538–1556.

[70] Jingxian Wu. "Connectivity analysis of a mobile vehicular ad hoc network with dynamic node population." In: *2008 IEEE Globecom Workshops*. IEEE. 2008, pp. 1–8.

[71] Tsung-Jung Wu, Wanjiun Liao, and Chung-Ju Chang. "A cost-effective strategy for road-side unit placement in vehicular networks." In: *IEEE Transactions on Communications* 60.8 (2012), pp. 2295–2303.

[72] Yuan Yao, Lei Rao, and Xue Liu. "Performance and reliability analysis of IEEE 802.11p safety communication in a highway environment." In: *IEEE transactions on vehicular technology* 62.9 (2013), pp. 4198–4212.

[73] Lixia Zhang. "Virtual clock: A new traffic control algorithm for packet switching networks." In: *ACM SIGCOMM Computer Communication Review*. Vol. 20. 4. ACM. 1990, pp. 19–29.

[74] Xin Ming Zhang, Long Yan, Kai Heng Chen, and Dan Keun Sung. "Fast, Efficient Broadcast Schemes Based on the Prediction of Dynamics in Vehicular Ad Hoc Networks." In: *IEEE Transactions on Intelligent Transportation Systems* (2019).

[75] Ting Zhe, Li Qin Huang, Qiang Wu, Jianjia Zhang, Chen Hao Pei, and Liangyu Li. "Inter-vehicle distance estimation method based on monocular vision using 3D detection." In: *IEEE Transactions on Vehicular Technology* (2020).

[76] Huan Zhou, Shouzhi Xu, Dong Ren, Chungming Huang, and Heng Zhang. "Analysis of event-driven warning message propagation in vehicular ad hoc networks." In: *Ad Hoc Networks* 55 (2017), pp. 87–96.

[77] Yanyan Zhuang, Jianping Pan, and Lin Cai. "A probabilistic model for message propagation in two-dimensional vehicular ad-hoc networks." In: *Proceedings of the seventh ACM international workshop on VehiculAr InterNETworking*. ACM. 2010, pp. 31–40.

APPENDIX

APPENDIX

A.1 DETAILS OF EQUATION (5.12) SECTION 5.3

The first term of (5.12) corresponds to the case when x falls into the transmission range of the first vehicle of the cluster. In this case $\boldsymbol{G}_{ij}(x) = Z_{ij}$ since $\mathcal{G} \geq R$ holds.

A.1.1 *Proof the first term of* (5.12)

The matrix $\boldsymbol{Z} = [Z_{ij}]$ with the transition probabilities of the MAP phases between the beginning and the end of the cluster that shown in Section 5.3 can express as

$$
\begin{aligned}
&\boldsymbol{Z} = \underbrace{e^{\boldsymbol{D_0}R/v}}_{A} + \underbrace{\int_{y=0}^{R/v} e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\boldsymbol{Z}\,dy}_{\boldsymbol{BZ}} \\
&\boldsymbol{B} = \int_{y=0}^{R/v} e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\boldsymbol{Z}\,dy = \int_{y=0}^{\infty} e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\boldsymbol{Z}\,dy - \int_{y=R/v}^{\infty} e^{\boldsymbol{D_0}(y-R)}\boldsymbol{D_1}\boldsymbol{Z}\,dy \\
&\boldsymbol{B} = (-\boldsymbol{D_0})^{-1}\boldsymbol{D_1}\boldsymbol{Z} - e^{\boldsymbol{D_0}R/v}\int_{y=R/v}^{\infty} e^{\boldsymbol{D_0}(y-R/v)}\,dy\,\boldsymbol{D_1}\boldsymbol{Z} \\
&\boldsymbol{B} = (-\boldsymbol{D_0})^{-1}\boldsymbol{D_1}\boldsymbol{Z} - e^{\boldsymbol{D_0}R/v} - e^{\boldsymbol{D_0}R/v}(-\boldsymbol{D_0})^{-1}\boldsymbol{D_1}\boldsymbol{Z} \Rightarrow \boldsymbol{P} = (-\boldsymbol{D_0})^{-1}\boldsymbol{D_1} \\
&\boldsymbol{Z} = e^{\boldsymbol{D_0}R/v} + \boldsymbol{P}\boldsymbol{Z} - e^{\boldsymbol{D_0}R/v}\boldsymbol{P}\boldsymbol{Z} \\
&\boldsymbol{Z} - \boldsymbol{P}\boldsymbol{Z} + e^{\boldsymbol{D_0}R/v}\boldsymbol{P}\boldsymbol{Z} = e^{\boldsymbol{D_0}R/v} \\
&(\boldsymbol{I} - \boldsymbol{P} + e^{\boldsymbol{D_0}R/v}\boldsymbol{P})\boldsymbol{Z} = e^{\boldsymbol{D_0}R/v} \\
&(\boldsymbol{I} - \boldsymbol{P} + e^{\boldsymbol{D_0}R/v}\boldsymbol{P})^{-1}(\boldsymbol{I} - \boldsymbol{P} + e^{\boldsymbol{D_0}R/v}\boldsymbol{P})\boldsymbol{Z} = (\boldsymbol{I} - \boldsymbol{P} + e^{\boldsymbol{D_0}R/v}\boldsymbol{P})^{-1}e^{\boldsymbol{D_0}R/v} \\
&\text{Final } \boldsymbol{Z}: \\
&\boldsymbol{Z} = (\boldsymbol{I} - \boldsymbol{P} + e^{\boldsymbol{D_0}R/v}\boldsymbol{P})^{-1}e^{\boldsymbol{D_0}R/v}
\end{aligned}
\tag{A.1}
$$

A.1.2 *Details of equation* (5.13) *Section* 5.3

In order to gives more proof details of mean cluster length $\boldsymbol{E}(\mathcal{G})$ (The second term of (5.12)), it is calculated as:

$$
\begin{aligned}
&\boldsymbol{E}(\mathcal{G}) = \int_{x=0}^{\infty} \boldsymbol{G}(x)dx = \int_{x=0}^{R} \boldsymbol{G}(x)dx + \int_{x=R}^{\infty} \boldsymbol{G}(x)dx \\
&= \int_{x=0}^{R} \boldsymbol{Z}dx + \int_{x=R}^{\infty}\int_{y=0}^{R/v} e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\boldsymbol{G}(x-y)dydx \\
&= \int_{x=0}^{R} \boldsymbol{Z}dx + \int_{y=0}^{R/v} e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\underbrace{\int_{x=R}^{\infty} \boldsymbol{G}(x-y)dxdy}_{\boldsymbol{S}} \\
&\Rightarrow \boldsymbol{S} = \int_{x=y}^{\infty} \boldsymbol{G}(x-y)dx - \int_{x=y}^{R} \boldsymbol{G}(x-y)dx \\
&= R\boldsymbol{Z} + \int_{y=0}^{R} e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\int_{x=y}^{\infty} \boldsymbol{G}(x-y)dxdy - \int_{y=0}^{R/v} e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\underbrace{\int_{x=y}^{R} \boldsymbol{G}(x-y)dxdy}_{\boldsymbol{Z}} \\
&= R\boldsymbol{Z} + \int_{y=0}^{R/v} e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\underbrace{\int_{x=y}^{\infty} \boldsymbol{G}(x-y)dxdy}_{\boldsymbol{E}(\mathcal{G})} - \int_{y=0}^{R} e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\underbrace{\int_{x=y}^{R} \boldsymbol{Z}dxdy}_{(R-y)\boldsymbol{Z}} \\
&= R\boldsymbol{Z} + \int_{y=0}^{R/v} e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\boldsymbol{E}(\mathcal{G})dy - \int_{y=0}^{R/v} (R-y)e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\boldsymbol{Z}dy \\
&= R\boldsymbol{Z} + \boldsymbol{P}\boldsymbol{E}(\mathcal{G}) - e^{\boldsymbol{D_0}R/v}\boldsymbol{P}\boldsymbol{E}(\mathcal{G}) - \underbrace{\int_{y=0}^{R/v} (R-y)e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\boldsymbol{Z}dy}_{\text{subsection Intergral by parts}} \\
&= R\boldsymbol{Z} + (\boldsymbol{P} - e^{\boldsymbol{D_0}R/v}\boldsymbol{P})\boldsymbol{E}(\mathcal{G}) - R\boldsymbol{P}\boldsymbol{Z} + \boldsymbol{D_0}^{-1}(\boldsymbol{Z} - e^{\boldsymbol{D_0}R/v}) \\
&= (\boldsymbol{I} - e^{\boldsymbol{D_0}R/v})\boldsymbol{P}\boldsymbol{E}(\mathcal{G}) + R\boldsymbol{Z} - R\boldsymbol{P}\boldsymbol{Z} - \boldsymbol{D_0}^{-1}\boldsymbol{Z} + \boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}R/v} \\
&(\boldsymbol{I} - \boldsymbol{P} + e^{\boldsymbol{D_0}R/v}\boldsymbol{P})\boldsymbol{E}(\mathcal{G}) = R\boldsymbol{Z} - R\boldsymbol{P}\boldsymbol{Z} - \boldsymbol{D_0}^{-1}\boldsymbol{Z} + \boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}R/v} \\
&\boldsymbol{E}(\mathcal{G}) = (\boldsymbol{I} - \boldsymbol{P} + e^{\boldsymbol{D_0}R/v}\boldsymbol{P})^{-1}\left[R\boldsymbol{Z} - R\boldsymbol{P}\boldsymbol{Z} - \boldsymbol{D_0}^{-1}\boldsymbol{Z} + \boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}R/v}\right] \\
&= (\boldsymbol{I} - \boldsymbol{P} + e^{\boldsymbol{D_0}R/v}\boldsymbol{P})^{-1}\left[R\boldsymbol{Z} - R\boldsymbol{P}\boldsymbol{Z} - \boldsymbol{D_0}^{-1}\boldsymbol{Z}\right] + \underbrace{(\boldsymbol{I} - \boldsymbol{P} + e^{\boldsymbol{D_0}R/v}\boldsymbol{P})^{-1}e^{\boldsymbol{D_0}R/v}}_{\boldsymbol{Z}}\boldsymbol{D_0}^{-1}
\end{aligned}
\tag{A.2}
$$

The Final matrix of $\boldsymbol{E}(\mathcal{G})$:

$$\boldsymbol{E}(\mathcal{G}) = (\boldsymbol{I} - \boldsymbol{P} + e^{\boldsymbol{D_0}R/v}\boldsymbol{P})^{-1}(\boldsymbol{I} - \boldsymbol{P})R\boldsymbol{Z} - (\boldsymbol{I} - \boldsymbol{P} + e^{\boldsymbol{D_0}R/v}\boldsymbol{P})^{-1}\boldsymbol{D_0}^{-1}\boldsymbol{Z} + \boldsymbol{Z}\boldsymbol{D_0}^{-1}$$

A.1.3 *Details of equation* (5.15) *Section 5.3*

More details of the phase-dependent second moment $\boldsymbol{E}(G^2)$ of the cluster length as

$$
\begin{aligned}
&\boldsymbol{E}(\mathcal{G}^2) = 2\int_{x=0}^{\infty} x\boldsymbol{G}(x)dx = 2\int_{x=0}^{R} x\boldsymbol{G}(x)dx + 2\int_{x=R}^{\infty} x\boldsymbol{G}(x)dx \\
&= 2\int_{x=0}^{R} x\boldsymbol{Z}dx + 2\int_{x=R}^{\infty}\int_{y=0}^{R}(x-y+y)e^{\boldsymbol{D}_0 y}\boldsymbol{D}_1\boldsymbol{G}(x-y)dydx \\
&= 2\int_{x=0}^{R} x\boldsymbol{Z}dx + 2\int_{y=0}^{R} e^{\boldsymbol{D}_0 y}\boldsymbol{D}_1 \underbrace{\int_{x=R}^{\infty}(x-y+y)\boldsymbol{G}(x-y)dx\,dy}_{\boldsymbol{S}} \\
&\Rightarrow \boldsymbol{S} = \underbrace{\int_{x=y}^{\infty}(x-y+y)\boldsymbol{G}(x-y)dx}_{\boldsymbol{S}_1} - \underbrace{\int_{x=y}^{R}(x-y+y)\boldsymbol{G}(x-y)dx}_{\boldsymbol{S}_2} \\
&\rightarrow \boldsymbol{S}_1 = \underbrace{\int_{x=y}^{\infty}(x-y)\boldsymbol{G}(x-y)dx}_{\boldsymbol{E}(\mathcal{G}^2)/2} + y\underbrace{\int_{x=y}^{\infty}\boldsymbol{G}(x-y)dx}_{\boldsymbol{E}(\mathcal{G})} \rightarrow \boldsymbol{E}(\mathcal{G}^2)/2 + y\boldsymbol{E}(\mathcal{G}) \\
&\rightarrow \boldsymbol{S}_2 = \int_{x=y}^{R} x\underbrace{\boldsymbol{G}(x-y)}_{\boldsymbol{Z}}dx \rightarrow \frac{R^2-y^2}{2}\boldsymbol{Z} \qquad \text{(A.3)} \\
&\boldsymbol{E}(\mathcal{G}^2) = \underbrace{\int_{x=0}^{R} xdx}_{\frac{R2}{2}}\boldsymbol{Z} + \underbrace{\int_{y=0}^{R} e^{\boldsymbol{D}_0 y}\boldsymbol{D}_1\underbrace{\int_{x=R}^{\infty}(x-y+y)\boldsymbol{G}(x-y)dx\,dy}_{\boldsymbol{S}}}_{\boldsymbol{K}} \\
&\boldsymbol{S} = \boldsymbol{S}_1 - \boldsymbol{S}_2 \Rightarrow \boldsymbol{S}_1 = \boldsymbol{E}(\mathcal{G}^2)/2 + y\boldsymbol{E}(\mathcal{G}), \boldsymbol{S}_2 = \frac{R^2-y^2}{2}\boldsymbol{Z} \Rightarrow \boldsymbol{K} = \int_{y=0}^{R} e^{\boldsymbol{D}_0 y}\boldsymbol{D}_1\boldsymbol{S}dy
\end{aligned}
$$

$$
\begin{aligned}
&\boldsymbol{K} = \int_{y=0}^{R} e^{\boldsymbol{D}_0 y}\boldsymbol{D}_1\left[\boldsymbol{E}(\mathcal{G}^2)/2 + y\boldsymbol{E}(\mathcal{G}) - \frac{R^2-y^2}{2}\boldsymbol{Z}\right]dy \\
&= \underbrace{\int_{y=0}^{R} e^{\boldsymbol{D}_0 y}\boldsymbol{D}_1\boldsymbol{E}(\mathcal{G}^2)/2dy}_{(\boldsymbol{P}-e^{\boldsymbol{D}_0 R}\boldsymbol{P})\boldsymbol{E}(\mathcal{G}^2)/2} + \underbrace{\int_{y=0}^{R} e^{\boldsymbol{D}_0 y}\boldsymbol{D}_1 y\boldsymbol{E}(\mathcal{G})dy}_{\text{integral by part1}} - \underbrace{\int_{y=0}^{R} e^{\boldsymbol{D}_0 y}\boldsymbol{D}_1\frac{R^2-y^2}{2}\boldsymbol{Z}dy}_{\text{integral by part2}}
\end{aligned}
$$

$$= (\boldsymbol{P}-e^{\boldsymbol{D_0}R}\boldsymbol{P})\boldsymbol{E}(\mathcal{G}^2)/2-Re^{\boldsymbol{D_0}R}\boldsymbol{P}\boldsymbol{E}(\mathcal{G})-\boldsymbol{D_0}^{-1}\boldsymbol{P}\boldsymbol{E}(\mathcal{G})+\boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}R}\boldsymbol{P}\boldsymbol{E}(\mathcal{G})-$$

$$\frac{R^2}{2}+R\boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}R}\boldsymbol{P}\boldsymbol{Z}+\boldsymbol{D_0}^{-2}\boldsymbol{P}\boldsymbol{Z}-e^{\boldsymbol{D_0}R}(\boldsymbol{D_0})^{-2}\boldsymbol{P}\boldsymbol{Z}$$

$$\boldsymbol{E}(\mathcal{G}^2)/2=\frac{R^2}{2}\boldsymbol{Z}+(\boldsymbol{P}-e^{\boldsymbol{D_0}R}\boldsymbol{P})\boldsymbol{E}(\mathcal{G}^2)/2-Re^{\boldsymbol{D_0}R}\boldsymbol{P}\boldsymbol{E}(\mathcal{G})-\boldsymbol{D_0}^{-1}\boldsymbol{P}\boldsymbol{E}(\mathcal{G})+\boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}R}\boldsymbol{P}\boldsymbol{E}(\mathcal{G})-$$

$$\frac{R^2}{2}\boldsymbol{P}\boldsymbol{Z}+R\boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}R}\boldsymbol{P}\boldsymbol{Z}+\boldsymbol{D_0}^{-2}\boldsymbol{P}\boldsymbol{Z}-e^{\boldsymbol{D_0}R}(\boldsymbol{D_0})^{-2}\boldsymbol{P}\boldsymbol{Z}$$

$$\boldsymbol{E}(\mathcal{G}^2)/2=\boldsymbol{Z}(\boldsymbol{D_0}^{-1}-R\boldsymbol{I})\boldsymbol{P}\left[\boldsymbol{E}(\mathcal{G})-\boldsymbol{D_0}^{-1}\boldsymbol{Z}\right]+(\boldsymbol{I}-\boldsymbol{P}+e^{\boldsymbol{D_0}R}\boldsymbol{P})^{-1}(-\boldsymbol{D_0}^{-1})\left[\boldsymbol{P}(\boldsymbol{E}(\mathcal{G})+e^{\boldsymbol{D_0}R}\right]$$

$\Rightarrow$ integral by part1

$$\int_{y=0}^{R} ye^{\boldsymbol{D_0}y}dy\boldsymbol{D_1}\boldsymbol{E}(\mathcal{G})=\left[y\boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}y}\right]_0^R\boldsymbol{D_1}\boldsymbol{E}(\mathcal{G})-\int_{y=0}^{R}\boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}y}1dy\boldsymbol{D_1}\boldsymbol{E}(\mathcal{G})$$

$$\left[Re^{\boldsymbol{D_0}R}\underbrace{\boldsymbol{D_0}^{-1}\boldsymbol{D_1}}_{-\boldsymbol{P}}\right]\boldsymbol{E}(\mathcal{G})-\boldsymbol{D_0}^{-1}\boldsymbol{P}\boldsymbol{E}(\mathcal{G})+\boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}R}\boldsymbol{P}\boldsymbol{E}(\mathcal{G})$$

$$\rightarrow -Re^{\boldsymbol{D_0}R}\boldsymbol{P}\boldsymbol{E}(\mathcal{G})-\boldsymbol{D_0}^{-1}\boldsymbol{P}\boldsymbol{E}(\mathcal{G})+\boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}R}\boldsymbol{P}\boldsymbol{E}(\mathcal{G})$$

$\Rightarrow$ integral by part2

$$\int_{y=0}^{R}e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\frac{R^2-y^2}{2}\boldsymbol{Z}dy=\underbrace{\int_{y=0}^{R}e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\frac{R^2}{2}\boldsymbol{Z}dy}_{part2-1}-\underbrace{\int_{y=0}^{R}e^{\boldsymbol{D_0}y}\boldsymbol{D_1}\frac{y^2}{2}\boldsymbol{Z}dy}_{part2-2} \tag{A.4}$$

$$\int_{y=0}^{R}\frac{y^2}{2}e^{\boldsymbol{D_0}y}dy\boldsymbol{D_1}\boldsymbol{Z}=\left[\frac{y^2}{2}\boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}y}\right]_0^R\boldsymbol{D_1}\boldsymbol{Z}-\underbrace{\int_{y=0}^{R}y\boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}y}dy\boldsymbol{D_1}\boldsymbol{Z}}_{\text{integral by part3}}$$

$$\left[\frac{R^2}{2}e^{\boldsymbol{D_0}R}\underbrace{\boldsymbol{D_0}^{-1}\boldsymbol{D_1}}_{-\boldsymbol{P}}\right]\boldsymbol{Z}-\left[-R\boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}R}-\boldsymbol{D_0}^{-2}+e^{\boldsymbol{D_0}R}(\boldsymbol{D_0})^{-2}\right]\boldsymbol{P}\boldsymbol{Z}$$

$$\left[\frac{-R^2}{2}e^{\boldsymbol{D_0}R}\right]\boldsymbol{P}\boldsymbol{Z}-\left[-R\boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}R}-\boldsymbol{D_0}^{-2}+e^{\boldsymbol{D_0}R}(-\boldsymbol{D_0})^{-2}\right]\boldsymbol{P}\boldsymbol{Z}$$

$$\text{part2-2}\rightarrow\left[\frac{-R^2}{2}e^{\boldsymbol{D_0}R}+R\boldsymbol{D_0}^{-1}e^{\boldsymbol{D_0}R}+\boldsymbol{D_0}^{-2}-e^{\boldsymbol{D_0}R}(\boldsymbol{D_0})^{-2}\right]\boldsymbol{P}\boldsymbol{Z}$$

integral by part2

$$\underbrace{\int_{y=0}^{R} e^{\boldsymbol{D}_0 y}\boldsymbol{D}_1\frac{R^2}{2}\boldsymbol{Z}\,dy}_{part2-1}$$

$$\Rightarrow \boldsymbol{P}\frac{R^2}{2}\boldsymbol{Z}-e^{\boldsymbol{D}_0 R}\boldsymbol{P}\frac{R^2}{2}\boldsymbol{Z}$$

$$\int_{y=0}^{R} e^{\boldsymbol{D}_0 y}\boldsymbol{D}_1\frac{R^2-y^2}{2}\boldsymbol{Z}\,dy=(part2-1)-(part2-2)$$

$$=\boldsymbol{P}\frac{R^2}{2}\boldsymbol{Z}-e^{\boldsymbol{D}_0 R}\boldsymbol{P}\frac{R^2}{2}\boldsymbol{Z}-\left[\frac{-R^2}{2}e^{\boldsymbol{D}_0 R}+R\boldsymbol{D_0}^{-1}e^{\boldsymbol{D}_0 R}+\boldsymbol{D_0}^{-2}-e^{\boldsymbol{D}_0 R}(\boldsymbol{D_0})^{-2}\right]\boldsymbol{PZ}$$

$$=\left[\frac{R^2}{2}-\cancel{e^{\boldsymbol{D}_0 R}\frac{R^2}{2}+\frac{R^2}{2}e^{\boldsymbol{D}_0 R}}-R\boldsymbol{D_0}^{-1}e^{\boldsymbol{D}_0 R}-\boldsymbol{D_0}^{-2}+e^{\boldsymbol{D}_0 R}(-\boldsymbol{D_0})^{-2}\right]\boldsymbol{PZ}$$

$$=\left[\frac{R^2}{2}-R\boldsymbol{D_0}^{-1}e^{\boldsymbol{D}_0 R}-\boldsymbol{D_0}^{-2}+e^{\boldsymbol{D}_0 R}(\boldsymbol{D_0})^{-2}\right]\boldsymbol{PZ} \tag{A.5}$$

$\Rightarrow$ integral by part3

$$\int_{y=0}^{R} ye^{\boldsymbol{D}_0 y}dy\underbrace{\boldsymbol{D_0}^{-1}\boldsymbol{D_1}}_{-\boldsymbol{P}}\boldsymbol{Z}\Rightarrow\int_{y=0}^{R}-ye^{\boldsymbol{D}_0 y}dy\boldsymbol{PZ}$$

$$\int_{y=0}^{R}-ye^{\boldsymbol{D}_0 y}dy\boldsymbol{PZ}=\left[-y\boldsymbol{D_0}^{-1}e^{\boldsymbol{D}_0 y}\right]_0^R\boldsymbol{PZ}-\int_{y=0}^{R}\boldsymbol{D_0}^{-1}e^{\boldsymbol{D}_0 y}-1\,dy\boldsymbol{PZ}$$

$$\left[-R\boldsymbol{D_0}^{-1}e^{\boldsymbol{D}_0 R}\right]\boldsymbol{PZ}-\left[-\boldsymbol{D_0}^{-1}(-\boldsymbol{D_0})^{-1}+\boldsymbol{D_0}^{-1}e^{\boldsymbol{D}_0 R}(\boldsymbol{D_0})^{-1}\right]\boldsymbol{PZ}$$

$$\left[-R\boldsymbol{D_0}^{-1}e^{\boldsymbol{D}_0 R}\right]\boldsymbol{PZ}-\left[\boldsymbol{D_0}^{-2}+(\boldsymbol{D_0})^{-2}e^{\boldsymbol{D}_0 R}\right]\boldsymbol{PZ}$$

$$\left[-R\boldsymbol{D_0}^{-1}e^{\boldsymbol{D}_0 R}-\boldsymbol{D_0}^{-2}-(\boldsymbol{D_0})^{-2}e^{\boldsymbol{D}_0 R}\right]\boldsymbol{PZ}$$

A.1.4 *Details of Theorem 18 Section 5.4.1.1*

The transient analysis in the MAP have three types as:

- transition without arrival: $\left[\boldsymbol{D_0}\right]_{ij}\Delta$
- transition with arrival: $\left[\boldsymbol{D_1}\right]_{ij}\Delta$
- no transition no arrival : $1+\left[\boldsymbol{D_0}\right]_{ii}\Delta$

The $\underline{F}(t,x)$ will characterize the information distance and the MAP phase at the end of the current cluster, where $F_i(t,x)=P(D(t)>x,\mathcal{J}(t)=i)$

$$F_i(t+\Delta,x) = \begin{cases} \text{Movement , no Phase transition,} & i \to i, \\ \text{it was not informed vehicle at } t \text{ and phase=j , new cluster,} & j \to i, \end{cases} \tag{A.6}$$

$$\begin{aligned}
&F_i(t+\Delta,x) = F_i(t,x+v\Delta) + \sum_{k=1}^{N}\sum_{j=1}^{N} \beta_j(t)[\boldsymbol{D_1}]_{jk}\Delta \boldsymbol{G}_{ki}(x-R) \\
&\left[F_i(t+\Delta,x) = F_i(t,x+v\Delta) + \sum_{k=1}^{N}\sum_{j=1}^{N} \beta_j(t)[\boldsymbol{D_1}]_{jk}\Delta \boldsymbol{G}_{ki}(x-R)\right]/\Delta \\
&\frac{F_i(t+\Delta,x)}{\Delta} - \frac{F_i(t,x+v\Delta)}{\Delta} = \sum_{k=1}^{N}\sum_{j=1}^{N} \beta_j(t)[\boldsymbol{D_1}]_{jk}\Delta \boldsymbol{G}_{ki}(x-R) \\
&\underbrace{\frac{F_i(t+\Delta,x)-F_i(t,x)}{\Delta}}_{\frac{d}{dt}F_i(t,x)} + \frac{F_i(t,x)}{\Delta} - \frac{F_i(t,x+v\Delta)}{\Delta} \\
&\frac{d}{dt}F_i(t,x) - \underbrace{\frac{F_i(t,x+v\Delta)-F_i(t,x)}{v\Delta}}_{-\frac{d}{dx}F_i(t,x)}.v \\
&\frac{d}{dt}F_i(t,x) - v\frac{d}{dx}F_i(t,x) = \sum_{k=1}^{N}\sum_{j=1}^{N} \beta_j(t)[\boldsymbol{D_1}]_{jk}\Delta \boldsymbol{G}_{ki}(x-R)
\end{aligned} \tag{A.7}$$

Finally:

$$\frac{\partial}{\partial t}\underline{F}(t,x) - v\frac{\partial}{\partial x}\underline{F}(t,x) = \underline{\beta}(t)\boldsymbol{D_1}\boldsymbol{G}(x-R)$$

The uninformed interval $\beta_i(t+\Delta)$ begins at time $t+\Delta$,
where $\beta_i(t) = P(\text{No informed vehicle at t}, \mathcal{J}(t) = i)$

$$\beta_i(t+\Delta) = \begin{cases} P(\text{No informed vehicle at } t\ ) \cdot P(\text{No arrival in } (t,\ t+\Delta)\ ), & i \to i, \\ P(\mathcal{D}(t) \in (R, R+v\Delta)) & \text{Cluster has left} \end{cases} \tag{A.8}$$

$$\beta_i(t+\Delta) = \beta_i(t)\Big[1+[\boldsymbol{D_0}]_{ii}\Delta\Big]+P(\mathcal{D}(t)>R)-P(\mathcal{D}(t)>R+v\Delta)+\sum_{j=1,j\neq i}^{N}\beta_j(t)[\boldsymbol{D_0}]_{ji}\Delta$$

$$\Big[\beta_i(t+\Delta) = \beta_i(t)\Big[1+[\boldsymbol{D_0}]_{ii}\Delta\Big]+F_i(t,R)-F_i(t,R+v\Delta)+\sum_{j=1,j\neq i}^{N}\beta_j(t)[\boldsymbol{D_0}]_{ji}\Delta\Big]/\Delta$$

$$\frac{\beta_i(t+\Delta)-\beta_i(t)}{\Delta} = \beta_i(t)[\boldsymbol{D_0}]_{ii}+\frac{F_i(t,R)-F_i(t,R+v\Delta)}{v\Delta}v+\sum_{j=1,j\neq i}^{N}\beta_j(t)[\boldsymbol{D_0}]_{ji}$$

$$\frac{d}{dt}\beta_i(t) = -v\frac{d}{dx}F_i(t,R)|_{x=R}+\sum_{j=1}^{N}\beta_j(t)[\boldsymbol{D_0}]_{ji}$$

$$\frac{d}{dt}\underline{\beta}(t) = -v\frac{d}{dx}\underline{F}(t,x)|_{x=R}+\underline{\beta}(t)\boldsymbol{D_0}$$

$$\frac{\partial}{\partial t}\underline{\beta}(t) = -\frac{\partial}{\partial t}\underline{F}(t,R)+\underline{\beta}(t)(\boldsymbol{D_0}+\boldsymbol{D_1}\boldsymbol{Z}) \quad \square \tag{A.9}$$