Routing and performance evaluation of disruption tolerant networks

Mouhamad Ibrahim

Université de Nice - Sophia Antipolis – UFR Sciences
École Doctorale STIC

THÈSE

Présentée pour obtenir le titre de :
Docteur en Sciences de l'Université de Nice - Sophia Antipolis

Spécialité : INFORMATIQUE

par

Mouhamad IBRAHIM

Équipes d'accueil:
MAESTRO – INRIA Sophia Antipolis

Routing and performance evaluation of disruption tolerant networks

Soutenue publiquement à l'INRIA le 14 Novembre 2008 devant le jury composé de :

Président :	Walid	DABBOUS	INRIA, France
Directeur de thèse :	Philippe	NAIN	INRIA, France
Rapporteurs :	Marco	AJMONE MARSAN	Ecole polytechnique de Turin, Italie
	Marwan	KRUNZ	Université d'Arizona, USA
Examinateurs :	Jerome	GALTIER	France Telecom R&D, France
	Fabio	MARTIGNON	Université de Bergame, Italie

THÈSE

Routing and performance evaluation of disruption tolerant networks

Routage et évaluation des performances des réseaux tolérants aux perturbations

Mouhamad IBRAHIM

14 Novembre 2008

To my family

Acknowledgments

At the end of these three years of work towards my PhD, I would like to express some words of gratitude to people who contributed in some ways to its achievement.

First and foremost, I want to express my deepest gratitude to my advisor Philippe Nain for having given me the opportunity to work on this exciting research topic and the freedom to carry it out. I am also grateful to him for the time he spent listening, reading and commenting my works, and also for his many constructive criticisms that enriched me.

Next, I wish to express my big thanks to Ahamd Al Hanbali, Sara Alouf and Iacopo Carreras for their fruitful collaborations and help during several periods of my research work.

I wish to thank all the members of my PhD committee. I thank Marwan Krunz and Marco Ajmone Marsan for their patience and effort to review my thesis dissertation. Thanks go also to Walid Dabbous, Jerome Galtier and Fabio Martignon for accepting to be in the jury committee.

I would like also to thank Konstantin Avrachenkov for giving me the chance to participate in the teaching of his class last year and Eitan Altman for giving me the opportunity to get involved in BIONETs project. Personal thanks go to Giovanni Neglia for the discussions regarding my work and for the amusing time shared together. Thanks as well to all my colleagues at Maestro team for the pleasant and friendly working atmosphere : Abdulhalim for the many advices, Alonso for being what you are ! Thanks also to Natalia, Danil, Damiano, Alberto, Dinesh, and Amar. I am grateful also to Elaine Isnard for being so kind to read and correct the French resume of the thesis. I am very thankful to our cheerful assistant Ephie Deriche for her inexhaustible help, her kindness and for the enjoyable discussions during the daily lunches. My acknowledgements go also to all the staff at INRIA who made it a pleasant working place.

I would like also to address personal thanks to all the lovely friends I met during my stay at Sophia and who changed in a way or in another my way of looking to life. In random order : Ghuilaine, Olga, Malli, Hicham, Marwan, Mahmoud, Gaelle, Mokhtar, Aymen, Amine, Maya, Luigi, Aimeric, Fadi, Nancy, Inass, Kevin, Corinne, Manar, Taleb, Hassan, Salam, ... Thank you for your friendship and for the wonderful and memorable moments spent together !

Last but not least, love and thanks to my family for their endless love, support and encouragement during all my studies.

Mouhamad IBRAHIM
Novembre 2008, INRIA Sophia Antipolis

Table of Contents

Introduction

In this thesis, we are interested in the design and evaluation of routing and medium access control protocols for wireless networks. In particular, in the first part, we focus on evaluating the performance of existing and novel forwarding protocols for sparse and intermittently connected mobile ad hoc networks when these networks include fixed relays. In the second part, we deal with the design and the evaluation of an adaptive backoff mechanism for the 802.11 MAC protocol.

Evolution of wireless technology

Driven by the need to support continuously evolving services and applications, the networking technologies have to address constantly novel challenges and to propose correspondingly innovative solutions. Conceived initially to enable communication among fixed stations, the wired networks have given rise to the Internet and all the web services around. Later on, wireless networks have appeared on the scene with wireless local area networks (WLANs) as the main popular technology.

WLANs allow for fast deployment where the stations only need to be equipped with wireless cards in order to connect to a base station representing their bridging point to the Internet or to other parts in the network. Moreover, wireless transmission allows nodes to move inside the network and potentially to "roam" to other neighboring networks. Consequently, WLANs have extended traditional wired networks to many new life sectors thereby providing flexible access to remote network services and ubiquitous connectivity. Examples include hot spots at airports, coffee bars, metro stations, public area, and many other outdoor applications.

In parallel to their spread, the deployment and utilisation of WLANs has brought also novel technical challenges for the traditional networking protocols. One of the main issues in WLANs is related to the design of an efficient Medium Access Control (MAC) suitable for the characteristics of the wireless medium. For instance, differently from packet transmission on a wired medium where transmitting stations can detect and accordingly react to the presence of a collision, stations transmitting on a wireless medium are not aware of all the collisions occurring on the channel. In addition, capacity of a wireless channel is in general much smaller than that of a wired medium, and hence optimal utilisation of the available bandwidth needs to be addressed.

As a consequence, several approaches to the design of MAC layer in WLANs have been proposed whether in terms of research proposals or in terms of commercial products that appeared on the market [99]. Nowadays, the most popularly deployed standard for physical and MAC layers is the IEEE 802.11 standard introduced in its initial version in 1997 [2]. The 802.11 MAC layer uses the Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) mechanism, a derived mechanism from the Carrier Sense Multiple Access/Collision Detection (CSMA/CD) used on the traditional wired Ethernet LANs.

The 802.11 standard defines two possible operation modes for the one-hop networks : *infrastructure mode* and *ad hoc mode*. In the infrastructure mode, there is a fixed base station which "bridges" the local network to other external networks, and which acts as a router for the local fixed and mobile nodes. In the ad hoc mode, the network does not require any infrastructure or centralized administration, and nodes, whether fixed or mobile, can communicate directly with each others in a peer-to-peer fashion[1].

With the proliferation and the development of new applications for wireless networks, a new class of ad hoc wireless networks based on multihop communication has emerged. Specifically, a node can communicate directly with another node located within transmission range in an ad hoc manner. To communicate with nodes located beyond that range, the node needs to use other intermediate nodes to route the messages hop by hop. Multihop wireless networks include mesh networks and Mobile Ad Hoc Networks (MANET). Wireless mesh networks, are built on a mix of fixed and mobile nodes that are interconnected via wireless links [19] in order to provide long distance wireless access for outdoor urban and rural areas. Commercial applications for wireless mesh networks include wireless sensor networks, wireless networks for public safety and transport systems, public Internet access networks and many others (see [85]). These applications have shown great potential on the wireless market, and several companies started proposing solutions to implement indoor and outdoor wireless mesh networks [32, 19].

On the other side, MANETs have been designed to provide multihop communication inside a group of mobile nodes. Conceived initially for military applications, the use of MANETs has been tailored to some specific and specialised civilian applications (e.g. emergency recovery networks) or for vehicular ad hoc networks (VANETs)[2]. Differently from wireless mesh networks, MANETs have faced until now poor commercial deployment despite of the significant amount of research work. The principal reason is due to the fact that in many application scenarios high nodes mobility makes hard to establish and maintain multihop routes, resulting in low performance. In some cases, there could even be no complete path between two given nodes at a given time, mainly because of high node mobility and/or of small coverage area of the network. This class of MANETs is referred to in the literature as disruption tolerant networks (DTNs) or intermittently connected networks. In such a

[1]Note that in the literature, the term ad hoc network is used also to refer to any self-organized network that does not necessarily use the IEEE 802.11 as MAC protocol.

[2]VANETs are a form of MANETs ; they are used for communication among vehicles and between vehicles and roadside equipment.

case, traditional MANETs routing protocols (e.g. Destination-Sequenced Distance-Vector (DSDV) [89], Dynamic Source Routing (DSR) [57], Ad hoc On-Demand Distance Vector (AODV) [88], etc...) necessarily fail in their attempt to discover a path.

In order to deal with such challenging scenarios, a novel approach has emerged which exploits the mobility of nodes to perform routing. This approach, termed as store-carry-and-forward [61, 55, 21], takes advantage of intermediate nodes mobility to relay the packets to other nodes. Specifically, upon contacts, intermediate nodes store the data of other nodes and then forward them at later future contacts to their respective destinations or to other intermediate nodes. This approach has been adopted in *ZebraNet* [61] which is a DTN designed to support wildlife tracking of zebras. The network is constituted by sensor collars that are attached to a group of zebras in order to study their movement patterns, and by researchers base stations that are mounted on cars which move around sporadically. The objective of the network is to vehicle traces collected at the zebras to researchers base stations. Due to the lack of connectivity, the network makes use of zebras mobility to route the collected traces. In particular, when two zebras meet, the corresponding collar sensors exchange their data, hence increasing the chance of data to be delivered to one of the moving base stations (please refer to Section 1.1 for other potential and realistic implementations of these networks).

Besides being often an unavoidable constraint, lack of global connectivity can be sometimes a desirable feature, in fact mobility has shown to increase the throughput when the wireless network is disconnected. This observation was outlined in [42] by Grossglauser and Tse which showed that source nodes can take advantage of the mobility of other relay nodes in order to reach a constant per-node throughput capacity. More specifically, the authors have considered a simple relaying protocol where a packet can be delivered either directly from the source to the destination upon contact or indirectly where the source forwards its packet to another relay node that forwards it later to the destination. Their observation was in sharp disagreement with the previously established "pessimistic" observation outlined by Gupta and Kumar in [44] where the authors showed that the per-node throughput capacity in fixed wireless networks decreases approximately like $1/\sqrt{n}$, with n the number of nodes by unit area.

Along with this approach, there have been numerous studies that appeared lately which rely on the mobility of nodes in the network to assist in packet routing (see Chapter 1). Since these protocols are mobility dependent, their performances are greatly related to the characteristics of contact opportunities among the relay nodes such as frequency, duration and sequence of such contacts. In addition, these protocols need to support the special features and requirements of DTNs. For instance, Although applications designed for disruption tolerant networks would tolerate relatively large delivery delay (emails exchange, information dissemination, etc...), they would clearly also benefit from low delivery delays. Moreover, nodes belonging to these networks are generally characterized by scarce resources in terms of power and storage capacities (case of pocket switched networks [49], mobile sensors networks [61, 94], etc...), and consequently forwarding protocols are expected to present low overhead. Therefore, the various routing studies proposed for DTNs aim as performance criteria either to maximize the packet delivery ratio, or to minimize the delivery delay and resources consumption, or to trade off among these different metrics.

Motivation and contribution of the thesis

Most of the existing literature on routing protocols for intermittently connected mobile ad hoc networks assumes that all relay nodes are mobile. In this thesis we are interested to evaluate the potential improvement deriving from adding throwboxes. Throwboxes are stationary wireless devices that act as fixed relays with enhanced power and storage capabilities [109, 84]. They can be easily deployed in an ad hoc manner into the network using batteries for short-term use or solar panel for longer-term use. For instance, they can be thrown via airplanes in case of difficult-to-reach areas (battlefield, forests,...) or can be simply installed on the roofs of buildings in case of urban networks. Furthermore, due to their low cost, throwboxes can be deployed in a large number in the network to increase the number of contacts among the mobile nodes.

Using throwboxes in this class of networks is a novel approach and it is promising in terms of enhancements that can be achieved on the performance of forwarding protocols and consequently on the related applications. For instance, Banerjee et al. in [7] report experimental results collected on the *UmassDieselNet* testbed when it is extended with throwboxes [3]. The results showed that adding one throwbox to the network improves the packet delivery by 37% and reduces the message delivery latency by more than 10%.

As we just mentioned formerly, the thesis seeks in the first part to address the issues of routing in intermittently connected MANETs that are augmented with throwboxes. In line with that, we consider opportunistic contact networks where the mobile nodes move randomly inside the network according to some random mobility model and where throwboxes are deployed according to some general distribution. By opportunistic contact we mean that nodes have zero knowledge regarding their future or past contacts with other nodes. The choice of the random mobility of nodes is dictated by our lack of real mobility traces from one side, and the ability of random mobility models to be general and scenario independent. Given a forwarding protocol, our target is to evaluate the expected delivery delay of one packet and the expected number of copies generated at the mobile nodes and throwboxes. For homogeneous nodes, the expected delivery delay depends on the expected meeting rates of the nodes, but also on the forwarding mechanism of the protocol. The number of generated copies reflects the protocol overhead in terms of power dissipated for copy transmissions and storage capacity required at the nodes.

In accordance with network model, we show through extensive simulations that inter-meeting times between a throwbox and a mobile node moving according to some random mobility models, namely the random waypoint or the random direction, can be accurately modeled by an exponential distribution. We derive an accurate approximation of its parameter, and we show that this parameter depends on the network and nodes parameters, as well as on the stationary spatial distribution of the mobility model and the spatial distribution of throwboxes within the area (see Section 2.4). For either the random waypoint or the random direction models, we derive also accurate approximations for the mean contact

[3] The *UmassDieselNet* is an experimental DTN developed by people at the University of Massachusetts (UMass) and made of 40 buses serving the UMass campus. Each bus is equipped with wireless cards and GPS device ; Buses scan constantly for other buses within transmission ranges and transfer data upon contacts [84].

duration between a mobile and a throwbox (see Section 2.4.1).

Based on the previous exponential property and a similar property for mobile/mobile inter-meeting times (see [41]), we introduce a Markovian model to evaluate the performance of two common routing protocols, namely the multicopy two-hop routing protocol and the epidemic routing protocol, in a throwbox-augmented MANETs. For either protocols, we derive closed-form expressions for the distribution of the delivery delay of a packet and for the distribution of the total number of copies generated at the various nodes till the delivery time (see Section 3.2). The previous analyses have been carried out for the cases when throwboxes are fully disconnected or mesh connected. Through a mean-field approach, we also provided asymptotic results when the number of mobile nodes and/or throwboxes is large.

We then introduce and evaluate five new routing schemes to route packets in a throwbox-augmented MANETs. We propose a stochastic framework based on recursive equations to compute analytically and to compare the performance of the five schemes. The proposed framework is general enough and can be easily extended to compute other performance metrics under different routing protocols. The obtained results highlight the various trade-offs that are left to network designers and provide several findings to guide the design and operation of throwbox-augmented DTNs.

Along the work on routing in MANETs, we investigate the expected length of the delivery path in terms of number of hops in networks running the epidemic protocol. The latter parameter, termed in the literature as the diameter of the epidemic protocol, has been studied empirically up to now. Under the assumption of exponentially distributed inter-meeting times, we provided a closed-form expression of the diameter which is in accordance with the previously reported empirical results (see Section 5.2).

In the second part of the thesis, we propose a novel mechanism that enhances the performance of the Distributed Coordination Function (DCF) of the IEEE 802.11 backoff algorithm. The proposed mechanism is simple enough and can be easily integrated with the standard. It presents two main and novel features, and showed better performance when compared with the standard or with other existing approaches. According to the proposed algorithm, the wireless stations will adaptively select the minimum contention window CW_{min} they should use after a successful transmission taking into account the number of active stations in their vicinities. In order to estimate the number of active stations around, the mechanism detects their presence by counting their emitted signals over the air, then it tracks their variations across time using a weighted moving average filter.

Organisation of the thesis

The thesis is divided into two parts. The first part, composed of Chapters 1-5, addresses evaluation and design issues of forwarding protocols in intermittently connected mobile ad hoc networks when these networks are augmented with throwboxes. The second part made up of Chapter 6 deals with the design of an efficient backoff algorithm to enhance the performance of the IEEE 802.11 DCF mechanism in congested networks.

The chapters are organized as follows. Chapter 1 reviews the current state of art of routing approaches in delay tolerant networks. A classification of the various techniques in this domain is presented with a discussion on the subtleties of each one of them. In Chapter 2, we examine the statistical distribution of inter-meeting time between a mobile node and a throwbox, and we derive an approximated expression of its parameter. Chapter 3 addresses the delay analysis and resource consumption in throwbox-embedded opportunistic DTNs under the multicopy two-hop routing protocol and the epidemic routing protocol. In Chapter 4, we introduce a stochastic framework which allows to evaluate analytically the performance of a large set of routing techniques in opportunistic DTNs augmented with throwboxes. We propose five routing strategies, and using the proposed framework, we evaluate and compare analytically the performance of each one of them. The average length of forwarding path in opportunistic DTNs is examined in Chapter 5 when packet relaying is carried out according to the epidemic protocol. In Chapter 6, we present and analyse an adaptive backoff algorithm for the IEEE 802.11 DCF mechanism.

Performance Analysis of Disruption Tolerant Networks

Chapitre 1

A Survey of Routing Approaches for Disruption Tolerant Networks

Contents

Disruption tolerant networks, also referred to as intermittently connected networks, are mobile wireless networks where most of the time there is no end-to-end communication path from the source to the destination. This situation arises when the network is sparse and/or when it spans over large distances. Traditional mobile ad hoc routing protocols are inefficient for these networks since they require the existence of connected end-to-end paths to be able to route any data. To overcome this issue, mobility-assisted routing schemes have been proposed which make use of the mobility of intermediate nodes to store-carry and forward the packets to their destinations. The aim of this chapter is to review the current state of the art in DTN routing approaches. We present a classification of the various studies in this domain and discuss the subtleties of each technique.

1.1 Disruption tolerant networks (DTNs)

Delay and disruption tolerant networks (DTNs) are a new class of wireless networks that seek to address the networking issues in mobile or challenging environments that lack continuous network connectivity. DTNs have emerged recently and are continuing to gain extensive efforts from the networking research community [1],[29],[93]. In the literature, these networks are found under different terminologies such as sparse mobile ad hoc networks, extreme wireless networks, or under another commonly used term intermittently connected networks. Basically, DTNs appear in areas where the network spans over large distances with low node density and/or with high node mobility. DTNs might appear also due to short radio range, power saving mechanism at the nodes, or nodes failure.

Examples of such networking scenarios include, but are not limited to :

- Vehicular networks, e.g. [86],[21]. In [86], the authors propose the *Drive-thru Internet* architecture where the objective is to provide network and Internet connectivity to mobile users in vehicles. The network is constituted by hot spots that are placed along the roads providing thus intermittent connectivity to the users that can connect within proximity. In [21], Burgess *et al.* introduce *UMass DieselNet* which is a network made of 30 buses equipped with 802.11b wireless interfaces and GPS devices. The objective of the network is to provide real DTN testbed for experimental and research studies. The buses move on regular trajectories inside the UMass Amherst campuses and surrounding areas. When two buses pass nearby, they transfer data to each other. Additionally, buses can connect to open wireless access points along the roads.
- Mobile sensor networks for environmental monitoring, e.g. [61],[94]. *Zebranet* [61] is a wireless networking architecture designed to support wildlife tracking for biology research. In ZebraNet, the network is constituted by sensor collars that are attached to zebras, which log movement patterns of the zebras, and by researchers base stations that are mounted on cars which move around sporadically. When two zebras meet, the corresponding sensors exchange collected data for a potential data delivery back to researchers base-stations. Another similar biological acquisition system has been proposed in [94], where the network is made of a set of sensors attached to whales and a set of fixed infostations that act as collecting nodes.
- Communication between rural zones in toward development countries, e.g. [17]. Examples include DakNet [17] which is a wireless ad hoc network that has the capacity to provide asynchronous Internet access to remote rural residents using motorcycles and buses to carry users email and web search messages.
- Deep space networks such as the Inter-planetary network (IPN) [22]. The inter-planetary network is a network of regional Internets. A *region* is an area where the characteristics of communication are the same. An example of regions includes the terrestrial Internet as a region or a ground-to-orbit region. IPN aims to achieve end-to-end communication through multiple regions in a disconnected, variable-delay environments.
- Challenged networks such as disaster healing networks after natural disaster, travel information and advertisements dissemination systems in large cities using local

transport systems, military ad hoc networks where disconnection occurs because of the war or for security reasons where some links need to be shut down from time to time.

Generally speaking, DTNs are wireless networks that do not conform to Internet or to traditional multihop and ad hoc wireless networks underlying structures and assumptions. In particular, they are characterized mainly by the following specific features [35],[102] :

- Intermittent connectivity where an end-to-end path between a given source-destination pair does not exist most of the time. Path disconnections are frequent and arise from two main factors, namely motion and/or limited power at the nodes. Disconnection due to motion can arise when one or both nodes at the end of a communication link move, or due to some intervening object or signal that obstruct the communication. These disconnections can be predicted, for instance when the nodes move away according to a predetermined schedule or, an opportunistic for instance according to random walk of the nodes. Disconnections that are due to power outage result commonly from some power saving mechanisms at the wireless devices, e.g. case of sensor networks. The latter disconnections are often predictable.
- Nodes have low power capabilities and limited resources. In many DTNs, nodes are generally battery powered and/or deployed in areas lacking power infrastructure. In some other situations, nodes have limited memory and/or processing capabilities.
- Large delays which are basically due to long queueing times resulting from frequent disconnections, or from low data rate at the devices.

1.2 Routing in DTNs

Due to frequent disconnections in DTNs, instantaneous end-to-end routes do not exist, and hence most of the traditional Internet and/or mobile ad hoc routing protocols fail [55]. However, end-to-end routes may exist over time if the nodes can take advantage of their mobility by exchanging and carrying other node messages upon meetings, and by delivering them afterward to their corresponding destinations. The latter concept has given rise to a novel routing paradigm in these networks called the store-carry-and-forward approach, in which intermediate nodes serve as relays for each other. Thus, the term "mobility-assisted routing approach" that is used in conjunction to describe these schemes.

Unfortunately, these techniques result in high latency, since packets need to be carried for long time periods before being delivered. When the delivery latency is not critical, as the case of delay-tolerant networks, the store-carry-and-forward paradigm can prove to be adequate. For instance, this is the case when the delivery of the messages is very important, possibly more important than the delay. Basically, with the store-carry-and-forward approach, the delivery delays of packets depend on the rate at which contact opportunities[1] are created in the network, as well as the availability of network resources, such as storage space and energy. The various studies that considered routing techniques in DTNs have examined the tradeoffs between optimizing the delivery ratio and delivery delay from one side, and reducing node resource consumptions in terms of storage and battery usage from

[1]Two nodes are said to be in contact if they are within transmission range of one another.

the other side. However, the intricacy of each one depends on the particularity of network environment at hand, the mobility model of the nodes, the performance objectives to attain, and other criteria.

This chapter will survey and classify various research works that have considered routing schemes for DTNs. Actually, there are different approaches to categorize these schemes [60, 108]. Hereafter, we propose a classification that is based on the way that these schemes make use of the knowledge of potential contact opportunities of the nodes in order to perform routing. Specifically, depending on whether these contact opportunities are scheduled, controlled, predicted or opportunistic, these approaches can be grouped into one of the four following families.

1.3 Scheduled-contact based routing

This section surveys the routing approaches that attempt to improve the performance of a sparse network when its dynamics are known in advance such as for instance Low-earth Orbiting satellites (LEO) based networks. In a given network scenario, the most important metrics of interest are the following : contact times between nodes (their starting times and durations), queue lengths at the nodes, the network traffic load. The complete knowledge of these three metrics by the routing protocol allows it to select optimal routes between the nodes. Despite that the implementation of the complete knowledge in a distributed environment is a very hard task, its evaluation is important as it constitutes the best case scenario compared with other cases where only a partial knowledge is available to the routing protocol. On the other side, the approaches that use no knowledge constitute the worst case scenario.

Jain *et al.* in [55] use the delay of a link as a cost function, and define the cost of a route to be the sum of its link costs. The authors propose four different techniques that use different degrees of knowledge. The first proposal is the Minimum Expected Delay (MED) where only the expectation of the link delay (excluding queueing delay) is known by the routing protocols. The second is the Earliest Delivery (ED) where the instantaneous link delay is available. The third is the Earliest Delivery with Local Queueing (EDLQ) where in addition to the use of the instantaneous delay, the delay at the local queue node is known. The last is the Earliest Delivery with All Queues (EDAQ) where in addition to the link delays, all the delays of the nodes queues are known. All these approaches were evaluated using simulations and compared to the zero knowledge and the complete knowledge cases. Their conclusion is that in networks with plentiful communication opportunities, the need for smart algorithms that require more knowledge is minimal. In situation where resources are limited, smarter algorithms (EDLQ and EDAQ) may provide significant benefits.

1.4 Controlled-contact based routing approaches

In this section, we discuss some routing approaches in DTNs which control the mobility of some dedicated additional mobile nodes in order to improve the network performance by increasing the contact opportunities between participating nodes. The additional mobile nodes can either have fixed predetermined paths conceived in a way to permit them to meet a large number of nodes, or their paths can be adjusted dynamically to meet traffic flows between the nodes. Their main task is to relay packets between the participating nodes by providing a store-carry-forward service. Indeed, by controlling the mobility of the additional nodes, a DTN network administrator would be able to limit the delivery delay and to provide bounds on some other performance metrics of the network. In the literature, several research works have discussed the integration of some special mobile nodes and the design of travel paths of these nodes to meet certain optimization criteria.

Jain *et al.* in [56] have introduced and modeled an architecture of a sparse network constituted by fixed sensors and extended by mobile nodes called *MULEs* (Mobile Ubiquitous LAN Extensions). MULEs move in the network area according to a random mobility model. Their task is to collect data from sensors, buffer it and drop it off later to a set of fixed base stations representing data sinks. The main objective of the architecture is to enhance power saving by allowing sensor nodes to exploit the random mobility of MULEs by transmitting their data to these mobile nodes over short range radio links when they pass nearby. To characterize data success ratio and queueing delay at the sensor buffer, the authors introduce a simple stochastic model based on renewal theory and bulk queueing theory. Through simulations, they have also investigated other performance metrics when the system parameters, the number of access points and the number of MULEs scale. Their basic observations confirm that an increase in the MULE density will improve system performance and leverage resource consumption.

Another controlled-contact routing work that is based on a proactive approach has been introduced in [110],[111] by Zhao *et al.*. The approach is termed proactive in the sense that the trajectories of the special mobile nodes, termed as message ferries (MF), are already determined and fixed. Under the assumption of mobility of network nodes, the authors consider two schemes of messages ferries, depending on whether nodes or ferries initiate the proactive movement. In the Node-Initiated Message Ferrying (NIMF) scheme, ferries move around the area according to known routes, collect messages from the nodes and deliver the messages later to their corresponding destinations. Aware of the ferries routes, the mobile nodes can adapt their trajectories to meet the ferries in order to transmit and receive messages. In the Ferry-Initiated Message Ferrying (FIMF), the ferries will move upon service requests to meet the nodes. Specifically, when a node has packets to send or to receive, it generates a service request and transmits it to a chosen ferry using a long range radio. When the ferry receives the request, it adapts its trajectory to meet with the node for packet exchanging using short range radio.

In their former work [110], the focus was on the design of ferry routes to meet certain constraints on throughput requirement and delivery delay in networks with stationary nodes using a single ferry. By formulating the problem as two optimization sub-problems, the authors developed algorithms to design the ferry route. In a recent work [112], the same authors considered the case of multiple ferries with the possibility of interaction between the ferries. The addition of multiple ferries has the advantage of improving the system performance and robustness to ferry failure at the cost of increasing the complexity of the problem. Based on several assumptions regarding whether the ferries follow the same or different routes and whether they interact with each other, the authors investigated four different route design algorithms that attempt to meet the traffic demand and minimize the delivery delay. Simulation results show that when the traffic load is low, the impact of increasing the number of ferries on the delivery delay is minor. However, for high traffic load scenarios, the impact is significant.

In [24], the authors propose an algorithm called MV routing which, on one side, exploits the movement patterns of participating nodes, that is the *meeting* and *visit* behaviors, and on the other side, attempts to control the motions of some additional external nodes. Their aim is to improve network efficiency in terms of bandwidth and latency of message delivery. The algorithm is seen as being constituted by two separate mechanisms. Building on their previous work in [33], the first mechanism is a slightly modified variant of the Drop-Least-Encountered technique that is used as a routing strategy instead of a buffer management technique as it has been used in [33]. The second mechanism of the algorithm consists in adapting dynamically the movement paths of some additional nodes to meet the traffic demands in the network while optimizing some performance criterion. Travel path adjustment is carried out through multi-objective control algorithms with the objective of optimizing simultaneously several metrics related to bandwidth and delay. Simulation results demonstrate that exploiting node mobility patterns in conjunction with multi-objective control for autonomous nodes has the most significant performance improvements.

1.5 Predicted-contact based routing approaches

Predicted routing techniques attempt to take advantage of certain knowledge concerning the mobility patterns or some repeating behavioral patterns. Based on an estimation of that knowledge, a node will decide on whether to forward the packet or to keep it and wait for a better chance. Basically, each node is assigned a set of metrics representing its likelihood to deliver packets to a given destination node. When a node holding a packet meets another node with a better metric to the destination, it passes the packet to it, hence increasing the packet likelihood of being delivered to its corresponding destination. According to the nature of knowledge, we propose to reclassify the algorithms falling under this category as based on mobility-pattern or based on history.

1.5.1 Mobility-pattern based approaches

Approaches falling under this section attempt to take advantage of common behaviors of node mobility patterns in the network in order to derive decisions on packet forwarding. In fact, by letting the nodes learn the mobility pattern characteristics in the network, efficient packet forwarding decisions can be taken. Two main issues are related to these approaches. The first issue concerns the definition and the characterization of the node mobility pattern where several ways can exist to characterize and acquire such a pattern. For instance, the appearance of stable node clusters in the network, or the acquisition of statistical information related to meeting times or to the visit frequencies of nodes to a given set of locations are examples of mobility patterns that can be exploited by the nodes. The second issue is related to the way through which a node can learn and acquire its own pattern as well as those of other nodes. In particular, the presence of some external signals to the nodes such as GPS coordinates or some fixed beacons help greatly the nodes to acquire easily the mobility patterns in the network. Alternatively, nodes can also learn their own mobility patterns without any external signal by relying only on previous observations and measurements, or by exchanging pattern information with other nodes. Several routing works in DTNs that use mobility patterns to derive forwarding decisions have appeared in the literature.

In [92], the authors develop a routing algorithm that exploits the presence of concentration points (CPs) of high node density in the network to optimize forwarding decisions. The appearance of CPs is seen as the result of a general mobility model where nodes will have a high concentration inside these CPs with random movements over time between these islands of connectivity. The basic idea of their algorithm is to make use of the neighbor set evolution of each node without using any external signals. Specifically, nodes that belong to a given concentration point will collaborate between them to assign a label to their CP. Nodes will learn the labels of other nodes when they move in the network between the different CPs. Using the knowledge of the CP graph, and the positions of the source and destination nodes in the graph, the message is forwarded from the source to its destination through a sequence of CPs using the shortest path between the respective CPs. Even though the algorithm performs well, the need to manage and update the labels introduces some complexity in the algorithm mechanism.

In another work [70], the authors introduce a virtual-location routing scheme which makes use of the frequency of visit of nodes to a discrete set of locations in the network area in order to decide on packet forwarding. Specifically, they define a virtual Euclidean space, termed as *MobySpace*, where the dimension degree and the type of the coordinate space depend on the mobility pattern of the nodes. For instance, for a network with L possible node locations, the MobySpace is an n-dimensional space where $n = |L|$. Each node is represented in that space by a virtual coordinate termed as *MobyPoint*. A source node X with a message to send at time t will forward its message to a node Y among the set of its neighbors $W_X(t)$ for which the Euclidean distance to the destination is the smallest. Observe that the MobyPoint of a node is not related to its physical GPS coordinate. The acquisition of the visit frequencies of the nodes to the location set is obtained by computing the respective fraction of time of being in a given location.

1.5.2 History based approaches

History based approaches are developed mainly for heterogeneous mobility movements. They rely on the observation that the future node movements can be effectively predicted based on repeating behavioral patterns. For instance, if a node had visited a location at some point in time, it would probably visit that location in another future time. Actually, if at any point in time a node can move randomly over the network area, an estimate based on previous contacts is of no help to decide on packet forwarding. However, if the mobility process has some locality, then last encounter times with other nodes can be associated with some weights that can be ranked based on their likelihood to deliver the messages to the corresponding destinations. The following works illustrate the working mechanisms of some of these approaches.

One of the pioneer work that considered history-based routing in sparse mobile networks is the work of Davis *et al.* in [33]. The objective of their work is to study the impact of different buffer management techniques on an extended variant of the epidemic protocol [100] on nodes with limited buffer size. Even though their work is not related to routing, the way by which the packets are sorted upon a contact influences implicitly the performance of the routing protocol. More precisely, when two nodes meet, they will first transfer the packets destined to each other, then they will exchange the lists of their remaining stored packets. The combined list of remaining packets is next sorted according to the used buffer management strategy, and each node will request the packets it does not have among the top K sorted packets. The authors have explored four different buffer management techniques, among them the *Drop-Least-Encountered* (DLE) technique which makes use of previous contacts with other nodes to decide on packet ranking. Basically, nodes using the DLE technique keep a vector indexed by addresses of other nodes where each entry estimates the likelihood of meeting the corresponding node. At each time step, a given node A updates its likelihood meeting values for every other node C with respect to the co-located node B according to the temporal difference rule (see [98]). If node A meets B, it is likely that A meets B again in the future, and hence A is a good candidate for passing the packets to B. Thus, node A should increase its likelihood for node B. If B has a high encounter for node C, then A should increases its likelihood of meeting C by a factor proportional to the likelihood of meeting between B and C. Last, if at a given time step, node A did not meet any other node, the different likelihood values decrease in a constant rate.

In [61], the authors propose a wireless peer-to-peer networking architecture, called ZebraNet system, which is designed to support wildlife tracking for biology research. The network is basically a mobile sensor network, where animals equipped with tracking collars act as mobile nodes which cooperate between them in a peer-to-peer fashion to deliver collected data back to researchers. Researcher base stations, mounted on cars, are moving around sporadically to collect logged data. The design goal is to use the least energy, storage and other resources necessary to maintain a reliable system with a very high data delivery success rate. To attain these objectives, they propose the use of a history-based protocol to handle packet transfer between neighbor peer nodes. More precisely, each node will be assigned a hierarchical level based on its past successes of transferring data to the base station. The higher the level of the node, the higher the probability that this node is within range of base station or within range of some other nodes near the base station. Therefore, it has a high likelihood of relaying the data back to the base station either directly or indirectly through minimal number of other nodes. The mechanism works as follows : each time a node scans for peer neighbors, it requests the hierarchy level of all of its neighbors. Collected data is then sent to the neighbor with the highest hierarchy level. Whenever a node comes within range of the base station, its hierarchy level is increased while it is decreased over time at a given rate when it is out-of-range.

Jones *et al.* in [59] propose a variant of the MED approach of [55] called Minimum Estimated Expected Delay (MEED). Alternatively to the MED approach where the expected delay of a link is computed using the future contact schedule, MEED uses an estimation of the observed contact history. The estimator implements a sliding history window with an adjustable size. To minimize the overhead induced by the frequent updates of the estimated link delay, the authors propose to filter update samples having small difference with the actual information in the network. Through simulations, MEED has shown to overcome the performance of MED as it is more responsive to network changes, and its performance approaches that of the epidemic protocol. However, the algorithm lacks the presence of an adjustment mechanism of its window size.

The authors in [71] propose PROPHET (Probabilistic Routing Protocol using History of Encounters and Transitivity), a single copy history-based routing algorithm for DTNs. Similarly to [33], each node in PROPHET will attempt to estimate a delivery predictability vector containing an entry for each other node. For a given node X, the entry $P(X, Y) \in [0, 1]$ will represent the probability of node X to deliver a message to a given node, for instance node Y in this case. The entries of the predictability vectors will be used to decide on packet forwarding. Specifically, when two nodes meet, a message is forwarded to the other node if the delivery predictability for the destination of the message is higher at the other node. In addition to the predictability vector, a summary vector of stored packets will be also exchanged upon contact. The information in the summary vector is used to decide on which messages to request from the other node. The entry update process occurs upon each contact and works as follows. Nodes that are often within mutual ranges have a high delivery predictability for each other, and hence they will increase their corresponding delivery predictability entries. Alternatively, nodes that rarely meet are less likely to be good forwarders of messages to each other, and hence they will reduce their corresponding delivery predictability entries.

1.6 Opportunistic-contact based routing approaches

Opportunistic based approaches are generally characterized by random contacts between participating nodes followed by potential pair-wise exchanges of data. Given that meeting, and consequently, data exchanges are subject to the characteristics of the mobility model which are in general unpredicted, these approaches rely on multi-copy schemes to speed up data dissemination within the network. In the following, we subdivide these approaches into epidemic-based approaches and coding based approaches.

1.6.1 Epidemic based approaches

Epidemic based approaches imitate the spread of contagious disease in a biological environment. Similarly to the way an infected individual passes on a virus to those who come into contact, each node in an epidemic-based system will spread copies of packets it has received to other susceptible nodes. The number of copies that an infected node[2] is allowed to make, termed as the fan-out of the dissemination, and the maximum number of hops that a packet is allowed to travel between the source and the destination nodes, represented by a hop count field in the packet, define the epidemic variant of the algorithm. These two parameters can be tuned to trade delay for resource consumption. Clearly, by allowing the packet to spread throughout the mobile nodes, the delay until one of the copies reaches the destination can be significantly reduced. However, this comes at the cost of introducing a large overhead in terms of bandwidth, buffer space and energy consumption. Several variants of epidemic-based approaches have been proposed and their performance in terms of delay and resource consumption have been evaluated.

One of the pioneer work in this domain is the epidemic routing protocol of Vahdat and Becker [100]. The protocol is basically a flooding mechanism accommodated for mobile wireless networks. It relies on pair-wise exchanges of messages between nodes as they get in contact with each other to eventually deliver the messages to their destinations. Each node manages a buffer containing messages that have been generated at the current node as well as messages that has been generated by other nodes and relayed to this node. An index of the stored messages called a summary vector is kept by each node. When two nodes meet, they will exchange their summary vectors. After this exchange, each node can determine then if the other node has some messages that was previously unseen by it. In this case, it will request the missing messages from the other node. To limit the resource utilization of the protocol, the authors propose to use a hop count field at each message that specifies the total number of epidemic exchanges that a particular message may be subject to. They showed that by appropriately choosing the maximum hop count, delivery rates can still be kept high while limiting resource utilization.

In [42], Grossglauser and Tse introduce a one copy two-hop relay protocol. Basically, at any time, they will be one copy of the packet in the network, however, the copy can make at most two hops between the source node and the destination node. Their packet dissemination algorithm can be seen as an epidemic-like protocol with a fan-out of one and a hop count of two. The key goal of their work is to show that the capacity of a mobile network can scale with the number of nodes by exploiting the mobility of these nodes through a two-hop relay protocol.

Building on [100] and [42], several research works have appeared subsequently which proposed analytical models to evaluate the performance of these protocols. In [41] Groenevelt *et al.* introduce a multicopy two-hop relay protocol (MTR), a variant of the two-hop relay protocol. In MTR, the source forwards a copy of the packet to any other relay node that it encounters. Relay nodes are only allowed to forward the packets they carry to their destinations. By modeling the successive meeting times between any pair of mobile nodes

[2] We use henceforth the term infected node to designate a node that carries a copy of the packet.

by Poisson processes, the authors characterize the distribution of the delivery delay and that of the total number of copies generated until the packet delivery. This work was extended in [6] by Al Hanbali *et al.* under the assumption of limited lifetime of the packets, and in [5] under the assumption of general distribution of inter-meeting times. Zhang *et al.* in [106] extend the work in [41] by evaluating several variations of the epidemic protocol and some infection-recovery schemes.

A biological acquisition system termed as the shared wireless infostation model (SWIM) has been introduced in [94] as a way of routing collected measurement traces between a set of sensors attached to whales and a set of fixed infostations acting as collecting nodes. Infostations act as base stations which connect the users to the network. Mobile nodes represented by the tagged whales move randomly within the area and connect to the infostations when they are within range to offload their data. When two tagged whales meet, an epidemic exchange mechanism takes place in order to accelerate the delivery of the packets at the cost of increasing the storage space at the nodes. Through simulations, the authors showed that sharing the data among the whales as well as increasing the number of SWIM stations reduce significantly the end-to-end delay. The positions of infostations as well as the mobility of whales greatly affect the system performance.

Spyropoulos *et al.* introduce a new routing algorithm for sparse networks in [96], termed Spray and Wait algorithm. The algorithm disseminates a number of copies of the packet to other nodes in the network, and then waits until one of these copies meets the destination. It consists of two phases. In the first phase, the source node will generate a total of L copies of the message it holds, then spreads these copies to other nodes for delivery to the destination node. The spreading process works as follows. When an active node holding $n > 1$ copies meets another node, it hands off to it $F(n)$ copies and keeps for itself the remaining $n - F(n)$ copies and so forth until a copy of the message reaches the destination. F is the function that defines the spreading process. For instance, for binary spray and wait, $F(n) = \frac{n}{2}$. In the second phase, the wait phase, if the destination is not found among the L copy-carrying nodes, then these latter nodes will perform direct transmissions to the destination node. Using simulations, the authors show that this technique can achieve a trade-off between efficient packet delivery and low overhead if the parameters are carefully designed.

1.6.2 Coding based approaches

The approaches in Section 1.6.1 are primarily based on packet flooding in order to improve the efficiency of packet delivery. Unfortunately, these improvements come at the expense of introducing large overhead in the network due to redundant packet transmissions. The approaches presented in this section alleviates the effect of flooding through the use of smarter redundant algorithms that are based on coding theory. In the following, we consider two main coding algorithms that appeared in the literature and which have shown their suitability to the opportunistic contact networks, namely the erasure coding and the network coding.

In the erasure coding scheme, upon receiving a packet of size m, the source produces n

data blocks of size $l < m$. The coding algorithm composes these blocks in a such way to allow the destination to retrieve the original message on receiving any subset of these blocks [76]. More precisely, the transmission of the packet is completed when the destination receives the kth block, regardless of the identity of the $k \approx m/l < n$ blocks it has received. The blocks are forwarded to the destination through the relay nodes according to store-carry-and-forward approach. The performance analysis of this approach in opportunistic contact network has shown to improve significantly the worst case delay with fixed amount of overhead [5, 101]. Further, in [54] it has been shown that erasure coding improve the probability of packet delivery in DTNs with transmissions failures.

In the network coding scheme, instead of simply forwarding the packets, nodes may transmit packets with linear combinations of previously received ones. For example, consider the three nodes case where nodes A and C want to exchange packets via the intermediate node B. A (resp. C) sends a packet a (resp. c) to B, which in turn broadcasts $a\ xor\ c$ packet to A and C. Both A and C can recover the packet of interest, while the number of transmissions is reduced. In [103], different aspects of network coding with limited storage resources have been discussed and different techniques have been proposed. The main result is that network coding benefits more from node mobility and performs well in scenarios of high packet drop rate where simple flooding approaches fail.

1.7 General comparison and discussion

Table 1.1 compares the various proposals that have been addressed throughout the chapter by summarizing the main distinguishable features of each one. Our comparison is based on four features. The first feature defines the degree of knowledge that the nodes have about their future contact opportunities. Future contact opportunities are identified as being scheduled, controlled, predicted or opportunistic. The second feature lists the key relevant performance metrics that each proposal attempts to optimize. For instance, these metrics range from increasing the packet delivery ratio to reducing the end-to-end delivery delay, energy consumption and/or buffer occupancy of the nodes. The third and fourth features list the characteristics of the mobility patterns of the network nodes and the dedicated special nodes, whenever employed. Precisely, nodes of the network could be stationary where in this case they are the special nodes that move around according to some predetermined or dynamic paths to assist in packet routing to the fixed nodes. Alternatively, node mobility pattern could be either random, where there is no means to predict the potential future contacts of a node, or heterogeneous with some location dependency or some correlated meeting among the nodes. Observe that the properties we have listed are not exhaustive and other properties can be included in addition. For instance, the complexity of the proposal in terms of implementation or computation, or the requirement to exchange some control information can also be considered. However, we restricted the comparison to the previous four features, which we think are the most relevant according to the classification that we made before.

Tab. 1.1 – Summary of the routing approaches in DTNs and their main properties.

Proposal	**Contact opportunities**	**Metric to optimize**	**Mobility pattern of network nodes**	**Mobility pattern of special nodes**
Jain *et al.* protocol [55]	Scheduled	Delay	Random	–
MULE protocol [56]	Controlled	Power usage, Buffer overhead	Stationary	Random
MF protocol [110]	Controlled	Delivery rate, Power usage	Stationary	Predetermined paths
Extended MF protocol [111]	Controlled	Delivery rate, Power usage	Random, Stationary	Predetermined, Dynamic paths
MV protocol [24]	Controlled	Delivery rate, Delay	Meeting and Visit dependant	Metric dependant paths
Island hopping protocol [92]	Predicted	Delay, Transmission overhead	Heterogeneous mobility	–
MobySpace protocol [70]	Predicted	Delivery rate, Power usage	Location dependant	–
DLE protocol [33]	Predicted	Buffer usage	Heterogeneous mobility	–
ZebraNet [61]	Predicted	Delivery rate, Power, Storage	Heterogeneous mobility	–
MEED protocol [59]	Predicted	Delay, Transmission overhead	Random	–
PROPHET [71]	Predicted	Delivery rate, Power usage	Heterogeneous mobility	–
Epidemic protocol [100]	Opportunistic	Delivery ratio, Delay	Random	–
Two-hop protocol [42]	Opportunistic	Network capacity	Random	–
MTR protocol [41]	Opportunistic	Delay, Transmission overhead	Random	–
SWIM protocol [94]	Opportunistic	Delivery rate, Delay	Random, Stationary	–
Spray and Wait [96]	Opportunistic	Delivery rate, Power usage	Random	–
Erasure coding [101]	Opportunistic	Delivery rate	Random	–
Network coding [103]	Opportunistic	Delivery rate	Random	–

1.8 Concluding remarks

Delay and/or disruption tolerant networks (DTNs) are a new class of emerging mobile wireless networks that are characterized by frequent disconnected paths among the nodes.

Under these conditions, message routing is a major challenge, and new concept of routing, called store-carry-and-forward, had appeared. To alleviate the problem of frequent partitions within the network, the DTNs routing protocols exploit nodes mobility where an intermediate node stores the packet, carries it until meeting another relay node or the destination node, and forwards it to the other relay node it meets.

In this chapter, we have provided an overview of various routing approaches for delay tolerant networks that have appeared recently in the literature and which follow the store-carry-and-forward model. We have also provided a classification of these approaches which is based on the way that they serve of the knowledge of potential contact opportunities of the nodes in order to perform routing. Based on that knowledge, we have classified them into one of the four categories : scheduled-contact based, controlled-contact based, predicted-contact based and opportunistic-contact based approaches. Generally, these approaches attempt to tradeoff routing performance, for instance packet delivery ratio, delivery latency and/or network capacity, versus node resource consumptions such as power and/or buffer usage.

Chapitre 2

Inter-meeting Time Distribution between Mobile and Fixed Nodes

Contents

In this chapter, we characterize inter-meeting time distribution between a mobile node and a throwbox. Through simulations, we show that mobile-throwbox inter-meeting time distribution can be modelled by an exponential distribution. We derive also approximation expressions of mobile-throwbox meeting rates under two common random mobility models, namely the random waypoint and the random direction models. Additionally, we derive approximation expressions of the mean contact time between a mobile and a throwbox for the last two mentioned mobility models. The various expressions are validated through simulations.

Note : Part of the results contained in this chapter have been published in [52].

2.1 Introduction

A Mobile ad hoc Network (MANET) is a self-configuring network consisting of mobile nodes that are communicating through wireless links and where the network does not assume any centralized mechanism to coordinate the communications among nodes. Nodes are free to move and each participating node can initiate a peer-to-peer data exchange with any other node through one-hop or multi-hop paths. Recently, with the proliferation of MANET applications, a new class of MANETs which is characterized by sparse node density started to gain attention from the networking community. With respect to that, the network turns to be partitioned for relatively large time periods and new approaches are appeared to handle communication among the nodes. Mainly, the latter approaches, termed as mobility-assisted approaches, exploit the mobility of nodes where participating ones store-carry and forward messages around the network as part of the routing protocol.

Specifically, mobility-assisted routing approaches consist of a series of independent forwarding decisions that take place when two nodes meet each other. Since message transfer occurs only at these instants, the time elapsed between two meetings, called inter-meeting time, plays a major role as it constitutes the basic component that influences the delivery rate and delay performances. Therefore, analyzing the statistical properties of inter-meeting times is essential in order to be able to evaluate the performance of any mobility-assisted routing scheme. Intuitively speaking, these properties are expected to depend on the specific mobility model in hand, and on nodes and network parameters.

In this chapter, we are interested in studying *the distribution and the intensity of inter-meeting time* between a mobile node moving according to some random mobility model and N throwboxes. A throwbox is a stationary and low cost wireless device that acts as a fixed relay which is mainly used to increase contact opportunities among mobile nodes, and consequently, the system performance.

The chapter is organized as follows. In Section 2.2, we give an overview of random mobility models and we describe the functional behavior and the statistical properties of two commonly used ones, namely the random waypoint and the random direction models. Next to that in Section 2.3, we show through simulations that the inter-meeting time between a mobile, moving according to random waypoint or random direction models, and any throwbox of a set of N throwboxes can be modelled by an exponential distribution. In Section 2.4, we derive analytical approximations of the meeting rate between a mobile and a throwbox corresponding to each mobility model, and for two different throwbox spatial distribution. Further, we derive accurate approximations of the meeting contact times between a mobile and a throwbox before concluding in Section 2.5.

2.2 Random mobility models

In the study of mobile wireless networks, the mobility model of the nodes takes a major consideration. Generally, a particular mobility model is conceived in order to mimic the movements of some mobile devices in a given real scenario or according to some collected mobility trace. These mobility models can be divided into entity models and group models [28]. Primarily, for entity mobility models, the movement of each device, noted as entity, is independent of the movements of other moving devices in the network. On the opposite, under group mobility model, mobile devices belonging to a particular mobility group can have the same mobility pattern. Hereafter in the thesis, we have focused only on mobile nodes moving according to random entity mobility model. Essentially, these models attempt to represent *random mobility movements*. They are mainly characterized by a memoryless mobility pattern where there is no information concerning node past location and speed values, and where node future movement is completely unpredictable. Moreover, since they are entity models, the node mobility process is identical and independently distributed from all others.

There are several random mobility models that have been proposed over the last ten years for mobile wireless networks. Here, we describe two commonly studied models, namely the *Random Waypoint (RWP)* model and the *Random Direction (RD)* model [10],[28]. In the following subsections, we describe each model with its associated statistical properties.

2.2.1 Random waypoint model

The random waypoint mobility model has been introduced for the first time by Johnson *et al.* in [58]. The model is defined as follows. At time zero, i.e. when the simulation starts, each node is assigned an initial position, let us called W_0 within the area which is sampled according to a uniform distribution. The position of the node is called *waypoint*. Starting from that position, each node chooses next waypoint, W_i, according to the uniform distribution to move to during its next trip. It also chooses a speed for that trip, V_i according to a uniform distribution where $V_i \in [V_{min}, V_{max}]$. Once the speed for the next trip is chosen, the node travels at the constant speed V_i till reaching the next waypoint. At the next waypoint, the node might pauses for some random time, called pause time, which is drawn according to some general distribution. Once the pause time is finished, the process starts again and the node chooses new waypoint and speed for its next trip.

Due to its common use in the simulations of mobile network, the steady-state behavior of RWP, noted as stationary behavior, have been widely studied over the last few years. As a matter of fact, the knowledge of the steady-state spatial and speed distributions of node are essential for all analyses for which the instantaneous location and speed of the mobile node are important.

Based on a set of observations in [13],[105],[79],[69], it has been reported that the stationary distribution of the location of a node, sampled at random time instants, is more concentrated near the center of the area in which the nodes move. In other terms, nodes moving according to RWP, which are traveling between uniformly chosen points, spend more time near the center than near the edges. Yoon *et al.* [105] have noticed also that the stationary distribution of the speed, sampled at random time instants, differs from their uniform distribution at the waypoints. Specifically, they showed that it is more likely to find a mobile with a low speed in the case when the speed sampling is done according to a uniform distribution at the waypoints. In addition, if the minimum speed V_{min} is taken to be 0, the mean node speed approaches 0. These observations are at the origin of a series of extensive research works that attempted to provide answers to these issues.

In [13],[12],[11], Bettstetter *et al.* have derived the stationary distribution of location for one dimensional segment and its approximation for a two-dimensional square area. They have also derived the transition length and time of a mobile node between two waypoints. In another work [51],[50], the authors have analyzed the stationary spatial distribution of nodes moving according to the RWP in a square area and a disk. They provided an explicit expression for the spatial density which is in the form of a one-dimensional integral that needs to be computed numerically.

Navid and Camp [79] have derived explicit expressions for the stationary distributions of speed, location, and pause time for a node moving in a rectangular area under the random waypoint mobility model. By claiming that the initial distributions of these metrics are not representative of their stationary distributions when the simulation progresses, they introduced a procedure in [80] which samples them from their corresponding stationary distributions. In fact, when they are sampled at the beginning of simulations from the stationary distributions, their convergence to steady-state distributions is immediate.

Using Palm calculus, Le Boudec *et al.* [69],[68] have also provided closed-from expressions for location and speed stationary distributions for a set of mobility models. Moreover, they provided necessary and sufficient conditions for a stationary regime to exist in the case of RWP model, and they extended the stationary results to the cases of non-convex and mutli-sites areas. They presented also an algorithm, called *perfect simulation*, which starts the simulation in steady-state. Observe that all the simulations that we have carried out throughout the thesis for RWP and RD mobility models are based on the perfect simulation code that is described in [87].

2.2.2 Random direction model

The random direction model has been first introduced for modeling users movement within cellular systems [48],[43]. Mainly because of its simplicity, it has been introduced recently as a mobility model for mobile wireless networks [10],[28],[8]. The model is defined as following. As in RWP model, at time zero each node is assigned an initial position uniformly distributed within the area. Starting from that position, each node chooses a direction θ_i uniformly distributed in $[0, 2\pi]$ for its next trip. It also chooses a speed for that trip, V_i according to a uniform distribution for which $V_i \in [V_{min}, V_{max}]$. Moreover, it selects a duration for its trip, T_i, according to some distribution, commonly an exponential distribution. Next, the node travels in the direction θ_i with the constant speed V_i for a duration T_i. Once travel time duration T_i expires, the node might pauses for a pause time which is drawn according to some general distribution. Next to that, it restarts the process again with new and independent values for the direction, speed and travel time. Observe that under this definition, the mobile node might hit the boundaries of the area. Hence, two variants of the RD model have appeared. Under the first variant, called wrap around [45], when the mobile hits a boundary, it reappears instantaneously on the other opposite boundary where it continues its trip. As a result, the location of the mobile is only changed while the direction and the speed vector remain unchanged. In the second variant, called with reflection [10],[8], when the node hits a boundary, it bounces off as in billiards. Consequently, the location and speed remain unchanged while the new direction is orthogonal to the previous one before the hit, i.e. angle of reflection equals angle of incidence. Note that in the literature, the two variants of RD model can be found under other different terminologies.

Unlike the case of random waypoint, the stationary distributions of speed and direction for the RD model, sampled at random time instants, have the same distributions of those at transition instants [78],[69]. Additionally, the stationary distribution of location, sampled also at random time instants, has a uniform distribution regardless of the variant being with reflection or with wrap around.

2.3 Characterizing the distribution of inter-meeting times between a mobile and N throwboxes

In this section, we investigate the statistical nature of the consecutive meeting times between a mobile node and a throwbox, and between a mobile node and any throwbox of a set of N throwboxes located in a square area $\mathcal{A}$ of size $L \times L$.

2.3.1 Related work

On the opposite to extensive research works that have analysed the steady-state statistical properties of mobility models, there are a few works that have addressed the statistical properties of inter-meeting times between two mobile nodes or between a mobile and a fixed nodes.

By using Markov chain to model the data dissemination process between two whales or to model the offloading process between a whale carrying a message and a fixed infostation, the authors in [94] have implicitly assumed exponential distributions for both inter-meeting times. However, the authors did not provide analytical expressions to compute these two meeting rates, and they have simply estimated their values using simulations.

In [41], it was shown through extensive simulations that one may accurately model the successive meeting times by Poisson processes when both nodes move according to the same *random mobility pattern*. A necessary condition for that assumption to hold is that the communication radius r should be *small* with respect to the side length L of the area. Using simulations, the authors have validated that assumption for three mobility patterns : the random waypoint (RWP) without pause, the random direction (RD) model with reflection or with wrap around, and the random walker. They have also derived approximation formulas for the intensity of the Poisson processes for both the RWP and the RD mobility models. The corresponding expressions are given in [40, Chapter 4].

Using real traces of six distinct experimental data set, Chaintreau *et al.* have observed in [30] that the inter-meeting times between two wireless devices carried by humans have a heavy tailed distribution. These traces report inter-meeting times between same pairs of wireless devices that are carried by humans moving in three campus universities, company offices and in a conference. In clear difference to the results reported in [41], the discrepancy between the two reported distributions in [41] and [30] is not surprising in our view. It is simply due to the fact that movement of humans in these environments are not i.i.d. , and the considered networks there are relatively dense. In particular, the carried wireless devices do not move randomly around the area. For instance, they spend a large amount of time moving or settling in some particular locations, and from time to time, they roam among these locations. Moreover, movement patterns are not identical. Their frequencies of visit to the set of locations are generally different, and some individuals might be more mobile than others.

In another work [97], Spyropoulos *et al.* have studied the expected meeting times between two mobile nodes, and between a mobile and a static node when mobile nodes are moving according to the RWP and RD models. By considering that the number of travel trips separating two meetings between a pair of nodes is geometrically distributed, the authors have provided closed form and approximated expressions for the expected meeting times. To represent real node movement, they have also introduced a community mobility model which is based on a two-states Markov model and presented an expression of the expected inter-meeting time between a mobile and a static node under this model.

In a recent work, and by analyzing real traces of inter-meeting times between mobile devices, the authors in [62] have shown that the corresponding CCDF [1] features a dichotomy property. Below a certain time value termed as characteristic time which is in the order of half a day, the CCDF of inter-meeting times follows a power-law distribution as has been reported in [30]. On the other hand, beyond the characteristic time, the distribution decays exponentially. They have also reported that in many cases, the mean inter-contact time is in the same order of the characteristic time, and hence the exponential decay property is of relevance.

2.3.2 Inter-meeting times characterization

Two nodes may only communicate when they are within each others transmission range. When at least one of two nodes is mobile, this event occurs in a random way at certain points in time, called *meeting times*. The time that elapses between two consecutive meeting times is called the *inter-meeting time*. We define *contact time* as the time that elapses between a meeting time, say t, of the mobile and the throwbox and the first time after time t when they are no longer within transmission range of one another. We highlight here that our definition of inter-meeting time period encompasses the relative contact time interval resulting from the meeting.

Similarly to [41], to characterize the inter-meeting times distribution we have first proceeded through simulations. The simulation scenario is as follows. There are N throwboxes and one mobile node. Throwboxes are placed in the area $\mathcal{A}$ independently of each other according to the uniform distribution. The mobile node moves according to either the RWP

[1] CCDF stands for complementary cumulative distribution function.

[18] without pause (with constant speed $V = 7m/sec$) or to the RD model with reflection at the boundaries [10],[8] (with exponentially distributed travel times with mean $200sec$, and constant speed $V = 7m/sec$). Side length of the area is set equal to $L = 4Km$.

Let τ_N be the stationary inter-meeting time between the mobile node and any one of the N throwbox. This time is defined as the time that elapses between two consecutive instants when the mobile node meets any one of the N throwboxes. Figure 2.1 plots, on a linear-log scale, the complementary cumulative distribution function (CCDF) of τ_N for $N \in \{1, 5, 10\}$, in the case where the throwboxes are placed in the area $\mathcal{A}$ independently of each other according to the uniform distribution, for RWP model. For each $N \in \{1, 5, 10\}$, at least $100,000$ meeting times have been simulated. The transmission range is constant for all nodes and set to $200m$ in 2.1-(a) and to $100m$ in 2.1-(b). Jointly to each curve, we have also plotted the CCDF of an exponential distribution with rate μ_N for $N \in \{1, 5, 10\}$, where μ is given by Equation (2.6). Similarly, figures 2.2-(a) and 2.2-(b) plot the same results

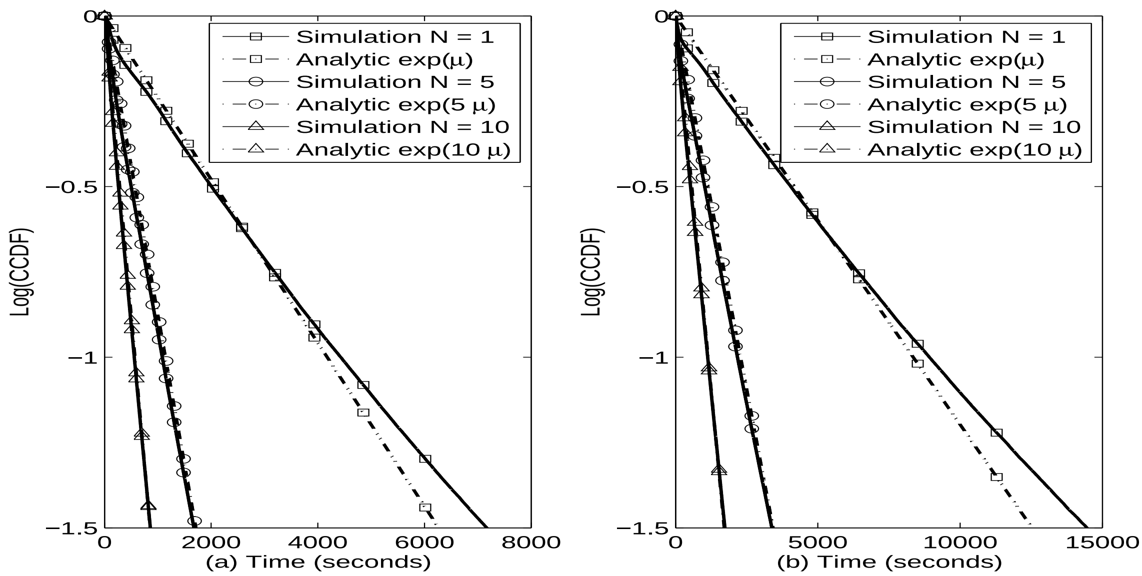

FIG. 2.1 – Tail distribution of inter-meeting times between N throwboxes and a mobile node moving according to RWP model. **(a)** : Tx range = 200m, **(b)** : Tx range =100m.

under the RD model with reflection (ERF). Here, the rate μ for the analytical exponential distribution is given by Equation (2.8).

For each $N \in \{1, 5, 10\}$, we observe that τ_N is well-approximated by an exponential random variable (r.v.), with parameter μ_N. By estimating the slope μ_N associated with each mapping $t \to \log P(\tau_N > t)$ for $N = 1, 5, 10$, we have also observed that μ depends linearly on N as $\mu_N \approx \mu_1 N$. This implies that τ_N is well-approximated by the minimum of N independent exponential r.v.s with common parameter $\mu := \mu_1$.

It is also worth to point out that through various simulations that we have conducted, we have also validated that property under different values of the simulation parameters and for different throwboxes spatial distributions (for instance stationary distribution of the

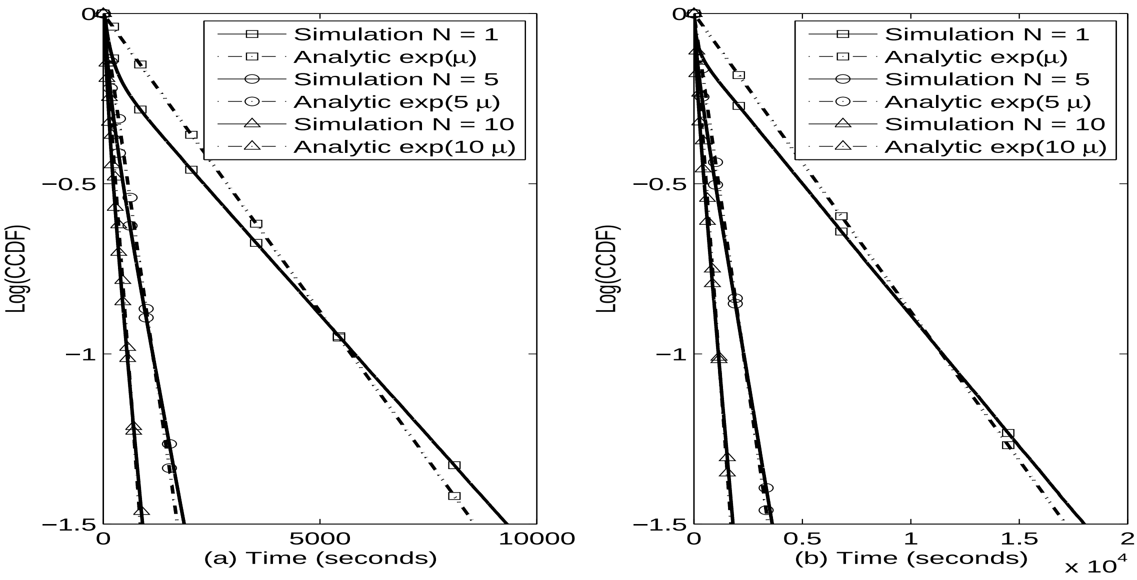

FIG. 2.2 – Tail distribution of inter-meeting times between N throwboxes and a mobile node moving according to RD with reflection model. **(a)** : Tx range = 200m, **(b)** : Tx range =100m.

RWP [68, 69]). However, and in analogy to the observation outlined in [41], the property of exponential distribution between a mobile node and a throwbox is shown to hold as long as the transmission range r is small compared to L. In fact, this assumption is in accordance with the underlying characteristics of disruption tolerant networks.

We have also conducted a statistical analysis using the classical autocorrelation function (ACF) method to test the mutual independence of inter-meeting times between a mobile and a throwbox. Recall that for a discrete time series $x = (x_1, \ldots, x_n)$ of length n, an estimate of the ACF of x (denoted by $\hat{R}$) is

$$\hat{R}(k) = \frac{1}{(n-k)\sigma^2} \sum_{i=1}^{n-k} (x_{i+k} - \bar{x})(x_i - \bar{x}), \quad k = 1, \ldots, n-1$$

where $\bar{x} = (1/n)\sum_{i=1}^{n} x_i$ and $\sigma^2 = (1/(n-1))\sum_{i=1}^{n}(x_i - \bar{x})^2$ denote the empirical mean and the empirical variance, respectively, of x.

Figure 2.3 displays the autocorrelation function of the inter-meeting times over a range of 10 lags (i.e. $k = 1, 2, \ldots, 10$) when the mobile moves according to the random waypoint model (left) or to the random direction model (right). For either mobility model, we observe that at each lag the autocorrelation is very close to zero, thereby justifying the assumption that inter-meeting times are mutually independent.

We conclude from the above that the consecutive meeting times between a mobile node and N throwboxes can be represented by a Poisson process with rate $N\mu$.

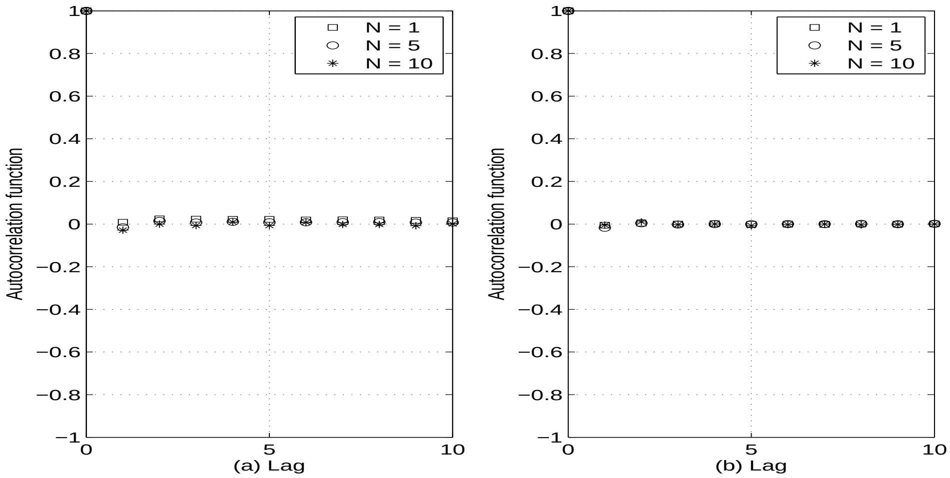

FIG. 2.3 – Autocorrelation function of inter-meeting times over 10 lags between N throwboxes and a mobile node moving according to the RWP model **(a)** and the RD model with reflection **(b)**.

2.4 Characterizing the intensity of inter-meeting times between a mobile and N throwboxes

By showing that the successive meeting times between a mobile node and N throwboxes are well-represented by a Poisson process with rate $N\mu$, we derive in this section approximation formulas of the parameter μ under the RWP and the RD models. Recall that μ is defined as the intensity at which the mobile node meets a single throwbox (i.e. $N = 1$). Even though the following presented formulas hold for the RWP and the RD models, the underlying analysis hold for any other mobility model given that its stationary spatial distribution is known.

Let π_c be the stationary probability that the mobile lies within the communication range of the throwbox. If we denote by $f(\cdot,\cdot)$ the density of the stationary location of the mobile node, i.e. the stationary spatial distribution of the mobility model, then

$$\pi_c = \int_{\mathcal{K}} f(x,y)dxdy,$$

where $\mathcal{K}$ is the area covered by the throwbox. Conditioned on the location (x, y) of the throwbox, if $r << L$ then we may approximate π_c by

$$\pi_c \approx \pi r^2 f(x,y) \tag{2.1}$$

Recall that the stationary contact time C is defined as the time that elapses between a meeting time, say t, of the mobile and the throwbox and the first time after time t when they are no longer within transmission range of one another. By an argument from renewal theory where a renewal cycle corresponds an inter-meeting time interval, we have then

$$\pi_c = \mathbf{E}[C]/\mathbf{E}[\tau] = \mu\,\mathbf{E}[C] \tag{2.2}$$

Plugging the expression of Equation (2.2) into Equation (2.1), we conclude that

$$\mu \approx \frac{\pi r^2 f(x,y)}{\mathbf{E}[C]} \tag{2.3}$$

conditioned on the location (x, y) of the throwbox. Unconditioning on throwbox location, we have that

$$\mu \approx \frac{\pi r^2}{\mathbf{E}[C]} \int_{\mathcal{A}} f(x,y)g(x,y)dxdy, \tag{2.4}$$

with $g(\cdot,\cdot)$ is the density of the throwbox location. Hence, to continue with the derivation of μ, it remains to evaluate $\mathbf{E}[C]$.

2.4.1 Mean contact time between a mobile and a throwbox

To derive the mean contact time $\mathbf{E}$[C], we assume that the mobile node does not change its direction when it crosses $\mathcal{K}$, the coverage area of the throwbox. Therefore, if σ denotes the length of the trip of the mobile when it crosses the area $\mathcal{K}$ and V_c denotes its corresponding speed during the contact, we have that $C = \sigma/V_c$.

When the mobile crosses the coverage area $\mathcal{K}$ of the throwbox, the intersection of length σ can be seen as a random line intercepting a convex area, so that $\mathbf{E}[\sigma] = \pi F/P = \pi r/2$, where F and P are respectively the area and the perimeter of $\mathcal{K}$ [91] (note that in this calculation we have implicitly assumed that $\mathcal{K} \cap \mathcal{A} = \mathcal{K}$, which is justified if $r << L$).

If we further assume that the speed V_c of the mobile at a meeting time is independent of σ, an assumption which holds for both the RWP and the RD models, we conclude from the above that $\mathbf{E}[C] = \mathbf{E}[\sigma]\,\mathbf{E}[V_c^{-1}] = \pi r \mathbf{E}[V_c^{-1}]/2$, so that

$$\mu \approx \frac{2r}{\mathbf{E}[V_c^{-1}]} \int_{\mathcal{A}} f(x,y)g(x,y)dxdy, \tag{2.5}$$

Therefore, it remains to find $\mathbf{E}[V_c^{-1}]$ for both the RWP and the RD models. As we will show in the next subsections, the value of $\mathbf{E}[\frac{1}{V_c}]$ shall depend on the mobility model used. For a mobile node moving according to the random waypoint model, the speed distribution of the mobile node sampled at the contact time instant is shown to be approximated by a uniform distribution and thus, $\mathbf{E}[\frac{1}{V_c}]$ is equal to $\frac{\log V_{max} - \log V_{min}}{V_{max} - V_{min}}$, while it can be approximated by a linear distribution for the case of random direction model, and $\mathbf{E}[\frac{1}{V_c}]$ will be equal to $\frac{2}{V_{max}+V_{min}}$. Table 2.1 lists the measured values of the mean contact time obtained through different simulations with the corresponding numerical values obtained through Equation (2.5) and the corresponding relative errors. The results are reported for the random waypoint model, random direction with reflection and with wrap around models. L is set equal to $2Km$, and the throwbox is located at coordinate $(1400, 1400)$.

2.4.1.1 Speed distribution for the random waypoint mobility model

Contrary to the concave distribution of speed that has been reported in [69] for an arbitrary speed sampling of a mobile node moving according to the random waypoint model,

Tab. 2.1 – Simulated and analytical mean contact time comparison for a two dimensional network.

RWP	$\mathbf{E}[C]_{measured}$	$E[C]_{computed}$	Relative Error %
r = 100, V = const = 5	31.696	31.4159	0.8863
r = 100, V = const = 10	15.837	15.708	0.817
r = 50, V = const = 5	15.763	15.708	0.3495
$r = 100, V_{max} = 12, V_{min} = 2$	28.0799	28.1449	0.2314
RD with Reflection	$\mathbf{E}[C]_{measured}$	$\mathbf{E}[C]_{computed}$	Relative Error %
r = 100, V = const = 5	31.4219	31.4159	0.0190
r = 100, V = const = 10	15.6835	15.708	0.156
r = 50, V = const = 5	15.7232	15.708	0.0971
$r = 100, V_{max} = 12, V_{min} = 2$	22.4628	22.4399	0.1017
RD with Wrap around	$\mathbf{E}[C]_{measured}$	$\mathbf{E}[C]_{computed}$	Relative Error %
r = 100, V = const = 5	31.4759	31.4159	0.1906
r = 100, V = const = 10	15.7197	15.708	0.0744
r = 50, V = const = 5	15.693	15.708	0.0953
$r = 100, V_{max} = 12, V_{min} = 2$	22.4636	22.4399	0.1055

mobile nodes moving according to the random waypoint model are shown to have a uniform distribution of speed when speed is sampled at contact time. Intuitively reasoning, given that the speed is uniformly distributed between V_{min} and V_{max} at a given waypoint, and given that the current path to the next waypoint crosses the coverage zone of throwbox, an observer located at the throwbox is more likely to observe during contact time the same distribution of speed as the one at the starting waypoint of the current path. In other terms, the realization of a given speed is independent of the event of crossing the throwbox, and hence, for an arbitrary observer located on mobile node travel path, speed distribution is the same as the one at the starting waypoint. Moreover, this observation is independent of the current position of the throwbox within the area $\mathcal{A}$.

Figure 2.4(a) shows a histogram of instantaneous speeds sampled at contact times in the case of a mobile node moving according to random waypoint model. For this scenario, mobile node speed is uniformly distributed between $V_{min} = 2\ m/s$ and $V_{max} = 12\ m/s$ at waypoints, and the transmission range is set equal to $r = 100m$. Area side length is set equal to $L = 2Km$ and the throwbox is uniformly distributed within that area. Clearly, the speed, sampled at contact time, presents a uniform distribution between V_{min} and V_{max} in accordance with our previous statement.

2.4.1.2 Speed distribution for the random direction mobility model

Unlike the case of random waypoint model, an observer located at throwbox and sampling the speed of a mobile node moving according to the random direction model when it passes within coverage, is more likely to observe high speed samples than low speed ones. Actually, by conditioning on given trajectory direction and travel time of the current path trip, the event of crossing the throwbox is clearly dependent on the value of the speed.

There is more opportunities to cross the throwbox for higher speeds than for lower ones. Therefore, given that the distribution of speed is uniform at waypoints, a linearly increasing density function of speed sampled at contact time instants is expected to be obtained. This assumption is confirmed through simulation results where figure 2.4(b) displays the histogram of speeds sampled at contact times of a mobile node moving according to random direction model.

Figure 2.4(b) shows a histogram of instantaneous speeds sampled at contact times in the case of a mobile node moving according to random direction model. Simulation parameters are the same as for the simulation of Figure 2.4-(a). In addition here, travel times are exponentially distributed with mean $200sec$. As the figure illustrates, the density probability of the speed, $f_{V_c}(x)$, sampled at meeting times is approximately a linear function (i.e. $f_{V_c}(x) = ax$, with a a normalizing constant). This yields $\mathbf{E}[V_c^{-1}] \approx 2/(V_{\max} + V_{\min})$ for the RD model with reflection.

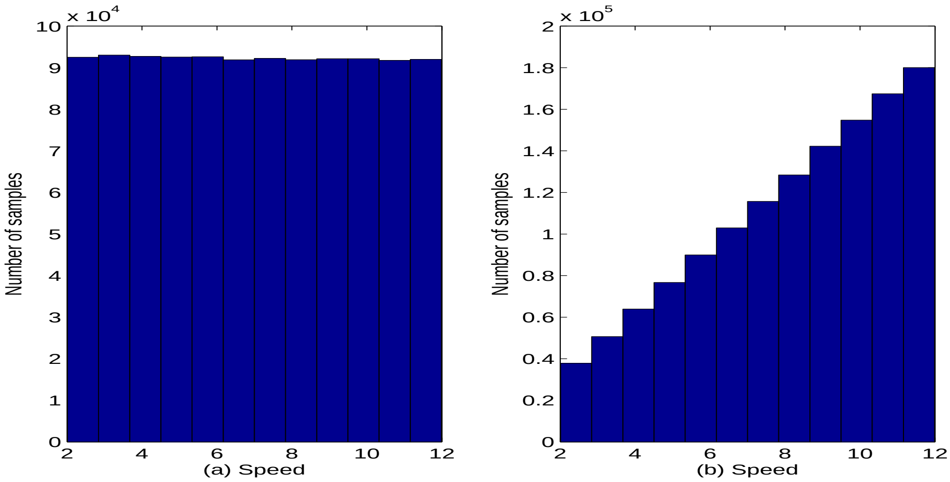

FIG. 2.4 – Histogram of speed sampled at meeting times of mobile and throwbox for the RWP model **(a)**, and the RD model with reflection **(b)**.

2.4.2 Approximation formula of μ

Recall the expression of meeting rate between a mobile and a throwbox μ which we have derived at the beginning of this section. A general approximation of μ is given by Equation (2.5) as follows :

$$\mu \approx \frac{2r}{\mathbf{E}[V_c^{-1}]} \int_{\mathcal{A}} f(x,y)g(x,y)dxdy,$$

As the equation illustrates, the meeting rate μ shall depend on the mobility model through the stationary density function of the location $f(.,.)$ and through the mean of the inverse of speed sampled at contact times $\mathbf{E}[\frac{1}{V_c}]$. Moreover, μ will depend on the spatial distribution of the throwbox in the area $\mathcal{A}$ through the spatial density function $g(.,.)$.

2.4.2.1 Case of random waypoint model

In the case where the nodes move according to the random waypoint model, closed-form expressions of the stationary density of location $f(.,.)$ are given in [69] and [50]. In addition, when the speed at waypoints V_i is sampled uniformly between V_{min} and V_{max}, the speed V_c is approximated by a uniform distribution, and $\mathbf{E}[\frac{1}{V_c}]$ is equal to $\frac{\log V_{max} - \log V_{min}}{V_{max} - V_{min}}$. On the other hand, when the speed at waypoints is constant, i.e. $V_i = V_{const}$, $\mathbf{E}[\frac{1}{V_c}]$ turns simply into $\frac{1}{V_{const}}$.

When the throwbox is distributed randomly in the area, the density function $g(.,.)$ converts to $\frac{1}{L^2}$, and therefore Equation (2.5) reads as

$$\mu \approx \frac{2r}{L^2 \mathbf{E}[\frac{1}{V_c}]} \tag{2.6}$$

However, a higher value of meeting rate can be obtained when the throwbox is distributed near the center of the area, for instance, according to RWP stationary density of location, i.e. $g(.,.) = f(.,.)$, and μ reads hence as

$$\mu \approx \frac{2r1.3683}{L^2 \mathbf{E}[\frac{1}{V_c}]} \tag{2.7}$$

Figure 2.1 plots the mapping $t \to -\mu_N t$ for $N \in \{1, 5, 10\}$ when the intensity μ is given by Equation (2.6). As the curves show, there is a good matching between the measured and the approximated intensity μ.

2.4.2.2 Case of random direction model

In the case where the nodes move according to the random direction model, the stationary density of location $f(.,.)$ is uniformly distributed ([78],[69]) and hence, $f(.,.)$ is equal to $\frac{1}{L^2}$. Therefore, Equation (2.5) reads as

$$\mu \approx \frac{2r}{L^2 \mathbf{E}[\frac{1}{V_c}]} \tag{2.8}$$

and consequently, the meeting rate between a throwbox and a mobile moving according to the RD model turns to be independent of throwbox location density function $g(.,.)$. The speed V_c sampled at contact times is show to be approximated by a linear distribution when the speed at waypoints V_i is sampled uniformly between V_{min} and V_{max} and thus $\mathbf{E}[\frac{1}{V_c}] = \frac{2}{V_{max}+V_{min}}$. When the speed V_i is constant, i.e.$V_i = V_{const}$, $\mathbf{E}[\frac{1}{V_c}] = \frac{1}{V_{const}}$

Figure 2.2 plots the mapping $t \to -\mu_N t$ for $N \in \{1, 5, 10\}$ when the intensity μ is given by Equation (2.8). Similarly to the case of random waypoint model, there is a good matching between the measured and the approximated intensity μ for the RD model.

2.5 Concluding remarks

In this chapter, we have considered a scenario where a mobile node comes at random instants within contacts of N throwboxes, which are essentially fixed relays, to exchange messages. With respect to this scenario, we have analyzed through simulations the distribution of inter-meeting time between the mobile and the N throwboxes and we showed that the latter distribution can be well approximated by an exponential one. This assumption is shown to hold as long as the network is sparse which results either from low nodes density or from small transmission ranges with respect to the area side lengths. We have also derived accurate approximate formulas for the intensity of the corresponding meeting rates in the case where the mobile moves according to the random waypoint or random direction models. The derived formulas take also into account the spatial distribution of the throwboxes within the network area. In addition, we have derived an approximate formulas for the mean contact time between a mobile node and a throwbox. The various derived expressions have been validated through simulations. Note that the results derived in this chapter will be used throughout the next two chapters to analyze the delay and energy consumption for a set of routing protocols running over sparse mobile ad hoc networks.

Chapitre 3

Performance Evaluation of Multicopy Two-Hop and Epidemic Routing Protocols in Throwbox-Embedded DTNs

Contents

This chapter addresses the delay analysis and resource consumption in throwbox-embedded DTNs. We consider two common routing protocols, namely the Multicopy Two-hop Routing protocol and the Epidemic Routing protocol, in the cases where the throwboxes are fully disconnected or mesh connected. To this end, we use a Markovian model where the successive meeting times between any pair of mobile nodes (resp. any mobile node and any throwbox) are represented by a Poisson process with intensity λ (resp. μ). We derive closed-form expressions for the distribution of the delivery delay of a packet and for the distribution of the total number of copies of a packet that are generated, the latter metric reflecting the overhead induced by the routing protocol. These results are then compared to simulation results. Through a mean-field approach we also provide asymptotic results when the number of nodes (mobile nodes and/or throwboxes) is large.

Note : Part of the results contained in this chapter have appeared in [52].

3.1 Introduction

The present chapter addresses the performance of two commonly used relaying protocols in a DTN[1] extended with throwboxes. A throwbox is a stationary wireless device with enhanced storage and power capabilities which acts as a fixed relay between mobile nodes passing by it at different times [109]. The considered networking scenario is as follows. Mobile nodes move according to some random mobility model, and throwboxes are spread out in the network area according to some general distribution. The throwboxes may either be fully disconnected or mesh connected. The network operates in intermittent connectivity conditions, for instance due to low node density, high node mobility, etc, and messages routing is handled according to store-carry-and-forward paradigm at mobile nodes (see Section1.2) and to store-and-forward paradigm at throwboxes.

With respect to this scenario, we examine the performance of the Multicopy Two-hop Relay protocol (MTR)[41] and the Epidemic Routing (ER) protocol [100]. Both protocols are defined in Section 1.6.1. Recall that in MTR when the source meets either a mobile node different from the destination node or a throwbox in the present considered scenario, it sends a copy of the packet to it ; a mobile node or a throwbox that carries a copy of the packet are only allowed to forward it to the destination when they meet. ER behaves the same as MTR except that a relay node (a mobile node or a throwbox) that carries a copy of the packet is allowed to generate another copy and to forward it to any other node that does not already carry a copy. In both MTR and ER, when the source meets the destination, it forwards the packet to it.

Performance efficiency of both protocols is investigated in terms of the delivery delay of a single packet and the number of copies generated at both mobile nodes and throwboxes until packet delivery. The latter metric reflects the power consumption that is induced by each routing protocol. To this end, we model contact opportunities between nodes (mobile nodes and throwboxes) by independent Poisson processes, and we develop a Markovian model for each routing protocol. Using the latter models, we derive closed-form expressions for the distribution of the delivery delay of a packet and for the distribution of the total number of copies that are generated. Through ordinary differential equations (ODEs), we examine also the asymptotic behavior of the MTR as the number of mobile nodes and/or the number of throwboxes become large.

Hereafter, we mention some related works that have considered message delay and energy consumption in sparse MANETs. Small and Haas in [94] provides a model, based on ODEs, to evaluate the performance of sparse MANETs embedded in an infostation network architecture. They consider the case where the ER protocol is used to relay data from the mobile nodes to the infostations. An infostation can be seen as a wireless access port to the Internet or to some private networks [39].

In [41], Groenevelt *et al.* introduced the so-called Multicopy Two-hop Relay protocol (MTR) which is a variant of the two-hop relay protocol that was proposed in [42]. By modeling the successive meeting times between any pair of mobile nodes by Poisson processes, Groenevelt *et al.* characterize the distributions of the packet delivery delay and of

[1] In the following, we use alternatively the term DTN to refer to a sparse MANET.

the number of copies generated before a packet is delivered to its destination. This work is extended in [6] by Al Hanbali *et al.* by imposing lifetime constraints. Building on [41], Zhang *et al.* in [106] extend the work in [94] by studying variations of the ER protocol, including probabilistic routing (see [64]) and recovery infection schemes.

In [109] Zhao *et al.* investigate the capacity enhancement of using throwboxes in Disruption Tolerant Networks (DTNs). Using a linear programming approach, they find the maximum total rate (λ) such that the network can sustain a rate of λb_{ij} between nodes i and j for all i and j, where $[b_{ij}]$ is the traffic matrix. In a follow-up of this work, Banerjee *et al.* in [7] report experimental results collected on the UmassDieselNet testbed (a 40 bus DTN) showing that the use of a single throwbox improves the packet delivery by 37% and reduces the message delivery latency by more than 10%.

The rest of the chapter is organized as follows. In Section 3.2, we introduce Markov models to characterize the packet delivery delay and the resource consumption, both for the MTR protocol 3.2.1 and for the ER protocol 3.2.2. The asymptotic behavior of the MTR protocol is addressed in Section 3.3, as the number of mobile nodes and/or the number of throwboxes become large. Our results are validated through simulations in Section 3.4 and concluding remarks are given in Section 3.5.

3.2 Markovian analysis

In this section we introduce and analyze simple stochastic models which will allow us to quantify, under the MTR and the ER protocols, the impact on the performance (packet delivery delay, resource consumption, etc.) of adding throwboxes in sparse MANETs and of using them as additional relay nodes.

Throughout the focus is on the transmission of a *single packet* between a source node and a destination node, both nodes being mobile nodes. Observe that the single packet model has been considered previously by various authors [94],[95],[41],[81] due to its simple mathematical tractability, and to its ability to be extend it to support the case of multiple packets. We assume that the transmission time of the packet or one of its copies between two nodes is instantaneous and successful. This setting typically corresponds to a situation where the network is sparse. The network is constituted by $N + 1$ identical mobile nodes, one source, one destination and $N - 1$ potential relay nodes, and M identical throwboxes, all located in a two-dimensional area, say a square $\mathcal{A}$. Mobile nodes are assumed to move independently of each other, according to the same mobility model.

Recall that according to our assumption in Section2.3, a meeting occurs either when two mobile nodes enter within each other transmission range or when a mobile node crosses the coverage area of a throwbox[2]. Motivated by the results in [41] for mobile nodes and by our observations in Chapter 2 for a mobile node and several throwboxes, we assume in the model that the inter-meeting times between any pair of mobile nodes form a renewal sequence of r.v.s with a common exponential distribution with parameter λ, and that the

[2]We assume implicitly that nodes use periodic beacon messages to discover the presence of other nodes in their neighborhood.

inter-meeting times between any pair composed of a mobile node and of a throwbox form a renewal sequence of r.v.s with a common exponential distribution with parameter μ. All these renewal sequences are assumed to be mutually independent. It is noteworthy to mention that this assumption has been validated when the nodes move either according to the random waypoint (RWP) [18] or random direction (RD) models [10],[28]. Subsequently, we assume that mobile nodes are moving according to one of these two mobility models.

Let T_{MTR} (resp. T_{ER}) be the first time where the packet or one of its copies, whichever event occurs first, reaches the destination under the MTR (resp. ER) protocol. Let G_{MTR} (resp. G_{ER}) be the number of copies generated until time T_{MTR} (resp. T_{ER}) under MTR (resp. ER) (we assume that the source is ready to transmit the packet at time $t = 0$ and that no copy has been generated yet).

For each protocol, we investigate the case where throwboxes are fully disconnected or mesh connected. The later case models situations when throwboxes are linked to each others (e.g. via wired lines or wireless links). In this case, a throwbox that receives a copy of the packet will forward it to all other throwboxes instantaneously.

The following notation will be used throughput this section. Let $R(t) \in \{1, 2, \ldots, N\}$ be the number of mobile nodes including the source that carry a copy of the packet. Similarly, let $B(t)$ be the number of throwboxes that carry a copy of the packet. Observe that when throwboxes are disconnected, $B(t) \in \{0, 1, 2, \ldots, M\}$, while when they are connected $B(t) \in \{0, M\}$.

We define the stochastic process $I(t) = (R(t), B(t))$ for $0 \leq t < T_{MTR}$ (resp. $0 \leq t < T_{ER}$) and $I(t) = a$ for $t \geq T_{MTR}$ (resp. $t \geq T_{ER}$). $I(t) = a$ indicates that the packet has been delivered to the destination at time t. Throughout this section we will assume $I(0) = (1, 0)$. Our modeling assumptions imply that, for both the MTR and the ER protocols, the process $\mathbf{I} = \{I(t),\, t \geq 0\}$ is a continuous-time Markov chain with one absorbing state $\{a\}$.

3.2.1 MTR protocol

In this section, we consider the multicopy two-hop relay protocol. Under this protocol, we study the network performance in terms of packet latency and total packet transmissions for the two cases of fully disconnected and mesh connected throwboxes. Note that the following analysis hold with appropriate modifications for a more general case of different meeting Poisson rates between the nodes. However, for the sake of simplicity, we will consider the case of identical exponential parameters.

3.2.1.1 Disconnected throwboxes

In this case, the state space of the absorbing Markov chain is $\xi = \{1, \cdots, N\} \times \{0, \cdots, M\} \cup a$. Its transition rate diagram is shown in Figure 3.1, from which we readily find

$$\begin{aligned}
q((i,m),(i+1,m)) &= (N-i)\lambda, \quad i=1,\ldots,N-1, \quad m=0,\cdots,M \\
q((i,m),(i,m+1)) &= (M-m)\mu, \quad i=1,\ldots,N, \quad m=0,\cdots,M-1 \\
q((i,m),a) &= i\lambda+m\mu, \quad i=1,\ldots,N, \quad m=0,\cdots,M \\
q((i,m),(i,m)) &= -(N\lambda+M\mu), \quad i=1,\ldots,N, \quad m=0,\cdots,M \\
q(i,j) &= 0, \text{ otherwise.}
\end{aligned}$$

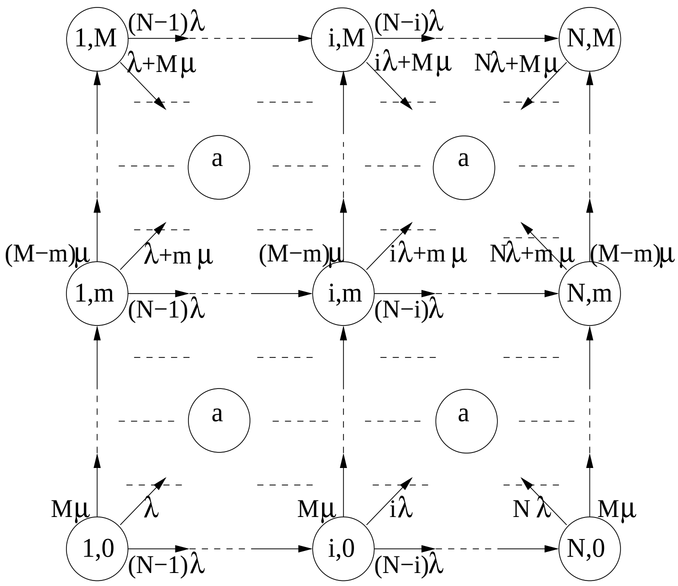

FIG. 3.1 – Transition rate diagram of the Markov chain in the case of MTR protocol and disconnected throwboxes.

Let us first determine the probability distribution of the delivery delay under this scenario, which will denote by T^d_{MTR}[3]. Following a similar approach to the one used in [5], let D_{sd} be the time for the source to encounter the destination, D^r_n be the time for the mobile relay node n to deliver a copy of the packet to the destination, and D^t_m be the time for a throwbox m to deliver a copy of the packet to the destination. Thanks to our modeling assumptions we see that D_{sd} is distributed according to an exponential r.v. with rate λ, D^r_n is distributed as the sum of two independent exponential r.v.s each with the same parameter λ, and D^t_m is distributed as the sum of two independent exponential r.v.s with the same parameter μ. Clearly,

$$T^d_{MTR} \stackrel{st}{=} \min(D_{sd}, D^r_1, \ldots, D^r_{N-1}, D^t_1, \ldots, D^t_M).$$

These r.v.s being all mutually independent, we have that

$$P(T^d_{MTR} > t) = P(D_{sd} > t) \prod_{n=1}^{N-1} P(D^r_n > t) \prod_{m=1}^{M} P(D^t_m > t),$$

[3]The superscript "d" stands for "disconnected" throwboxes.

so that

$$\begin{aligned} P(T^d_{MTR} > t) &= e^{-\lambda t} \prod_{n=1}^{N-1} e^{-\lambda t}(\lambda t + 1) \prod_{m=1}^{M} e^{-\mu t}(\mu t + 1) \\ &= e^{-(\lambda N + \mu M)t}(\lambda t + 1)^{(N-1)}(\mu t + 1)^M \end{aligned} \tag{3.1}$$

From (3.1) we can compute the expected delivery delay under MTR, given by

$$\mathbf{E}[T^d_{MTR}] = \frac{1}{\alpha} \sum_{n=0}^{N-1} \sum_{m=0}^{M} \binom{N-1}{n} \binom{M}{m} (n+m)! \left(\frac{\lambda}{\alpha}\right)^n \left(\frac{\mu}{\alpha}\right)^m \tag{3.2}$$

where $\alpha := N\lambda + M\mu$. Asymptotic results for $\mathbf{E}[T^d_{MTR}]$ when N and M are large will be obtained in Section 3.3.

Let us now turn our attention to the probability distribution of G^d_{MTR}, the number of copies generated by the MTR protocol until the delivery of the packet to the destination when throwboxes are disconnected. Define $P_a(n,m)$ as the probability that the last state visited by the Markov chain $\mathbf{I(t)}$ before absorption in state $\{a\}$ is (n,m) given that $I(0) = (1,0)$. The probability of the event $\{G^d_{MTR} = k\}$ is equal to the sum of the probabilities $P_a(n,m)$ for which $n+m=k$ with $1 \le n \le N$ and $0 \le m \le M$, that is,

$$P(G^d_{MTR} = k) = \sum_{n=max(1,k-M)}^{min(k,N)} P_a(n, k-n), \quad k = 1, 2, \ldots, N+M. \tag{3.3}$$

It remains to find $P_a(n,m)$. The Markov chain $\mathbf{I(t)}$ stays for an exponential amount of time with rate α in state $(i,j) \in \mathcal{E}$ after which it jumps into state $(i+1,j)$ with probability $\lambda(N-i)/\alpha$, it jumps in state $(i,j+1)$ with probability $\mu(M-j)/\alpha$ and it is absorbed with probability $(\lambda i + \mu j)/\alpha$. By writing down the probabilities of the different paths joining state $(1,0)$ to state (n,m) (there are $\binom{n+m-1}{n-1}$ such paths) we see that they are all as likely to occur, so that (Hint : take the path $(1,0) \to (N-1,0) \to (N-1,1) \to (N,M)$)

$$\begin{aligned} P_a(n,m) &= \binom{n+m-1}{n-1} \frac{n\lambda + m\mu}{\alpha} \prod_{i=1}^{n-1} \frac{(N-i)\lambda}{\alpha} \prod_{j=0}^{m-1} \frac{(M-j)\mu}{\alpha} \\ &= \frac{\lambda^{n-1}\mu^m(n\lambda + m\mu)}{\alpha^{n+m}} (n+m-1)! \binom{N-1}{n-1} \binom{M}{m}. \end{aligned} \tag{3.4}$$

The expected number of copies generated until the delivery of the packet, $\mathbf{E}[G^d_{MTR}]$, may be obtained from (3.3)-(3.4).

Observe that there exists an alternative way of computing $\mathbf{E}[G^d_{MTR}]$, which is based on the observation that T^d_{MTR} is the sum of G^d_{MTR} independent, identical, exponential r.v.s with parameter α (each of them corresponds to the sojourn time of the Markov chain $\mathbf{I(t)}$ in a state). Hence, by Wald's formula we obtain that $\mathbf{E}[T^d_{MTR}] = \mathbf{E}[G^d_{MTR}]/\alpha$, that is $\mathbf{E}[G^d_{MTR}] = \alpha \mathbf{E}[T^d_{MTR}]$, where $\mathbf{E}[T^d_{MTR}]$ is given in (3.2).

3.2.1.2 Connected throwboxes

In this scenario, we assume that throwboxes are mesh connected. Hence, whenever a throwbox receives the first copy of a packet from the source, it will flood it to all the other throwboxes instantaneously, and subsequently, the other throwboxes will abstain from receiving any new copies of the packet. Thus, the state space of the absorbing Markov chain is $\xi = \{1, \cdots, N\} \times \{0, M\} \cup a$. Its transition rate diagram is shown in Figure 3.2, from which we readily find

$$\begin{aligned}
q((i,0),(i+1,0)) &= (N-i)\lambda, \quad i = 1,\ldots,N-1, \\
q((i,0),(i,M)) &= M\mu, \quad i = 1,\ldots,N, \\
q((i,0),a) &= i\lambda, \quad i = 1,\ldots,N, \\
q((i,0),(i,0)) &= -(N\lambda + M\mu), \quad i = 1,\ldots,N, \\
q((i,M),(i+1,M)) &= (N-i)\lambda, \quad i = 1,\ldots,N-1, \\
q((i,M),a) &= i\lambda + M\mu, \quad i = 1,\ldots,N, \\
q((i,M),(i,M)) &= -(N\lambda + M\mu), \quad i = 1,\ldots,N, \\
q(i,j) &= 0, \text{ otherwise.}
\end{aligned}$$

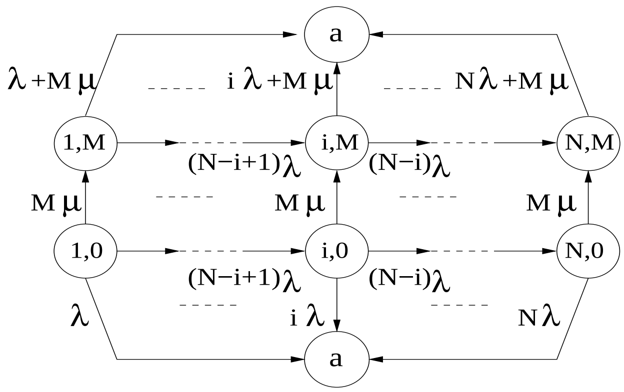

FIG. 3.2 – Transition rate diagram of the Markov chain in the case of MTR protocol and connected throwboxes.

Referring to the transition rate diagram of Figure 3.2, it is easily to recognize that the connected throwboxes scenario is a special case of the more general disconnected throwboxes scenario. In particular, the connected throwboxes setting corresponds to a network made of only one throwbox where its meeting rate with a mobile is hence equal to $M\mu$. Therefore, the analysis of this case is easily obtained from the previous one of disconnected throwboxes case by letting $M = 1$ and then by replacing μ by $M\mu$ in the various formulas. Consequently, the distribution of the delivery delay reads as

$$T^c_{MTR} \stackrel{st}{=} \min(D_{sd}, D^r_1, \ldots, D^r_{N-1}, D^t_1)$$

These r.v.s being all mutually independent, we have then

$$P(T^c_{MTR} > t) = P(D_{sd} > t) \prod_{n=1}^{N-1} P(D^r_n > t)\, P(D^t_1 > t),$$

so that

$$\begin{aligned} P(T^c_{MTR} > t) &= e^{-\lambda t} \prod_{n=1}^{N-1} e^{-\lambda t}(\lambda t + 1)\, e^{-M\mu t}(M\mu t + 1) \\ &= e^{-(\lambda N + \mu M)t}(\lambda t + 1)^{(N-1)}(M\mu t + 1) \end{aligned} \tag{3.5}$$

and hence, the expected delivery delay is given by

$$\mathbf{E}[T^c_{MTR}] = \frac{1}{\alpha} \sum_{n=0}^{N-1} \sum_{m=0}^{1} \binom{N-1}{n} \binom{1}{m} (n+m)! \left(\frac{\lambda}{\alpha}\right)^n \left(\frac{M\mu}{\alpha}\right)^m \tag{3.6}$$

where $\alpha := N\lambda + M\mu$. Similarly, $P_a(n, m)$ is equal to

$$\begin{aligned} P_a[(n,0)] &= \frac{n\lambda}{N\lambda + M\mu} \prod_{l=1}^{n-1} \frac{(N-l)\lambda}{N\lambda + M\mu} \\ &= n\frac{(N-1)!}{(N-n)!}\left(\frac{\lambda}{\alpha}\right)^n \end{aligned} \tag{3.7}$$

for $m = 0$ and

$$\begin{aligned} P_a[(n,M)] &= n\frac{M\mu}{\alpha}\frac{n\lambda + M\mu}{\alpha} \prod_{l=1}^{n-1} \frac{(N-l)\lambda}{\alpha} \\ &= nM\frac{(N-1)!}{(N-n)!}\frac{\lambda^{n-1}\mu(n\lambda + M\mu)}{\alpha^{n+1}} \end{aligned} \tag{3.8}$$

for $m = M$, and hence the distribution of the total packet transmission at the delivery time reads as

$$P(G^c_{MTR} = k) = \sum_{n=max(1,k-1)}^{min(k,N)} P_a(n, M(k-n)), \quad k = 1, 2, \ldots, N+1. \tag{3.9}$$

Similarly , the mean delivery delay $E[T^c_{MTR}]$ is given by

$$E[T^c_{MTR}] = \frac{E[G^c_{MTR}]}{\alpha} = \frac{E[G^c_{MTR}]}{N\lambda + M\mu} \tag{3.10}$$

3.2.2 Epidemic routing protocol

In this section, we consider nodes that relay packets according to the epidemic routing protocol. Similarly to the previous section, in the following we will evaluate the two settings when throwboxes are fully disconnected and mesh connected. The objective still the evaluation of the delivery delay and the total number of packet transmissions at delivery time. However, due to the complex structure of the Markov chain $\mathbf{I(t)}$ under the epidemic protocol, a direct approach like the one used in Section 3.2.1 will not be possible to determine the probability distributions of T_{ER} and G_{ER}. Instead, we will rely on the theory of absorbing Markov chains to identify these calculations.

3.2.2.1 Disconnected throwboxes

We will first consider the case where throwboxes are fully disconnected. Let $\mathbf{Q}^d_{ER}$ denote the infinitesimal generator of the Markov chain $\mathbf{I(t)}$ under this scenario. The generator $\mathbf{Q^d_{ER}}$ can be written as

$$\mathbf{Q^d_{ER}} = \left(\begin{array}{c|c} \mathbf{M} & \mathbf{R} \\ \hline \mathbf{0} & 0 \end{array}\right),$$

where $\mathbf{M} = [q(\mathbf{u},\mathbf{v})]_{\mathbf{u}\in\mathcal{E},\mathbf{v}\in\mathcal{E}}$, $\mathbf{R} = (q((1,0),a),\ldots,q((N,M),a))^T$, and $\mathbf{0}$ is the row vector of dimension $N\times(M+1)$ whose all components are equal to 0. The entries of the matrix $\mathbf{M}$ are given by

$$\begin{aligned}
q((n,m),(n+1,m)) &= (N-n)\beta(n,m), \quad n=1,\ldots,N-1, \quad m=0,\cdots,M \\
q((n,m),(n,m+1)) &= n(M-m)\mu, \quad n=1,\ldots,N, \quad m=0,\cdots,M-1 \\
q((n,m),(n,m)) &= -((N-n+1)\beta(n,m)+n(M-m)\mu), \quad n=1,\ldots,N, m=0,\cdots,M,
\end{aligned}$$

while the entries of the vector $\mathbf{R}$ are

$$q((n,m),a) = \beta(n,m), \quad n=1,\ldots,N, \quad m=0,\cdots,M$$

with $\beta(m,n) = n\lambda + m\mu$.

From the theory of absorbing Markov chains, we find that the probability distribution of the delivery delay T_{ed} is given by (see lemma 2.2.2 [82])

$$P(T^d_{ER} < x) = 1 - \mathbf{b}\, e^{\mathbf{M}x}\, \mathbf{e} \tag{3.11}$$

where $\mathbf{b}$ is a row vector of dimension $N\times(M+1)$ with the first entry equals to one and the remaining ones equal to zero, and $\mathbf{e}$ is a column vector of dimension $N\times(M+1)$ with all entries equal to one.

In particular, the k-th order moment of T^d_{ER} is equal to

$$\mathbf{E}[{T^d_{ER}}^k] = (-1)^k k! \sum_{j=1}^{N\times(M+1)} m^*_k(1,j), \tag{3.12}$$

where $m^*_k(i,j)$ is the (i,j)-entry of the matrix $(\mathbf{M}^{-1})^k$. Define $m^*(i,j) = m^*_1(i,j)$ for all entries (i,j).

The inverse of the matrix $\mathbf{M}$, which enters the calculation of the moments of T^d_{ER}, can be obtained in closed-form, as shown below. It is easily seen that $\mathbf{M}$ is an upper block bi-diagonal matrix with the $M+1$ square matrices $\mathbf{A}_m := [q((i,m),(j,m))]_{\{1\le i,j\le N\}}$, $m = 0,1,\ldots,M$, on the diagonal, and with the M square matrices $\mathbf{B}_m := [q((i,m),(j,m+1))]_{\{1\le i,j\le N\}}$, $m = 0,1,\ldots,M-1$, on the upper diagonal. The matrix $\mathbf{M}$ can be written as

$$\mathbf{M} = \begin{pmatrix} \mathbf{A}_0 & \mathbf{B}_0 & & & & \\ & \ddots & \ddots & & & \\ & & \mathbf{A}_m & \mathbf{B}_m & & \\ & & & \ddots & \ddots & \\ & & & & \mathbf{A}_{M-1} & \mathbf{B}_{M-1} \\ & & & & & \mathbf{A}_M \end{pmatrix}$$

The matrices $\mathbf{A}_m$, $m = 0, 1, \ldots, M$, are upper bi-diagonal matrices with strictly negative diagonal entries, so that they are all invertible. The matrices $\mathbf{B}_m$, $m = 0, 1, \ldots, M$, are all diagonal matrices. It can be checked that

$$\mathbf{M}^{-1} = \begin{pmatrix} \mathbf{A}_0^{-1} & \mathbf{U}_{0,1} & & \cdots & & \mathbf{U}_{0,M} \\ & \ddots & \ddots & & & \\ & & \mathbf{A}_m^{-1} & \mathbf{U}_{m,m+1} & \cdots & \mathbf{U}_{m,M} \\ & & & \ddots & \ddots & \\ & & & & & \mathbf{A}_M^{-1} \end{pmatrix}$$

where $\mathbf{U}_{m,l} = (-1)^{l-m}\left(\prod_{j=m}^{l-1} \mathbf{A}_j^{-1}\mathbf{B}_j\right)\mathbf{A}_l^{-1}$ for $0 \le m \le M-1$ and $m+1 \le l \le M$.

It remains to find $\mathbf{A}_m^{-1}$ in closed-form in order to derive $\mathbf{M}^{-1}$. Let $a_m^*(i,j)$ denote the (i,j)-entry of $\mathbf{A}_m^{-1}$. Since $\mathbf{A}_m$, $m = 0, 1, \ldots, M$, is an upper-bidiagonal square matrix, we find that

$$a_m^*(i,j) = \begin{cases} \dfrac{1}{q((i,m),(i,m))}, & i = j \\ \dfrac{(-1)^{j-i}}{q((j,m),(j,m))} \displaystyle\prod_{k=i}^{j-1} \dfrac{q((k,m),(k+1,m))}{q((k,m),(k,m))}, & j \ge i+1 \\ 0, & \text{otherwise.} \end{cases} \tag{3.13}$$

We now derive the distribution of G_{ER}^d, the number of copies generated by the ER protocol until the delivery of the packet to the destination. Similar to the MTR protocol, each transition in the Markov chain $\mathbf{I(t)}$ corresponds to a packet transmission. Starting from the state $(1,0)$, the largest path length is equal to $N+M$, therefore $P(G_{ed}) > N+M) = 0$. For $1 \le k < N+M$, $P(G_{ed} = k)$ is the probability that the last state visited before the absorption is one of the states (n,m) for which $n+m = k$. Therefore, it remains to find the probability $P_a(n,m)$ that the absorption occurs in state (n,m). By splitting the absorption state a into $N \times (M+1)$ independent absorption states $\{a_1, \ldots, a_{N\times(M+1)}\}$, one can show that (see [37, Lemma 1.c] for a similar approach)

$$P_a(n,m) = -m^*(1, n+mN)q((n,m),a), \tag{3.14}$$

for $n \in \{1, \cdots, N\}$ and $m \in \{0, \cdots, M\}$. Hence,

$$P(G_{ER}^d = k) = \sum_{n=max(1,k-M)}^{min(k,N)} P_a(n, k-n) = -\sum_{n=max(1,k-M)}^{min(k,N)} m^*(1, n+(k-n)N)q((n,k-n),a). \tag{3.15}$$

3.2.2.2 Connected throwboxes

In direct analogy with the MTR protocol, the performance of the ER protocol in the case where all throwboxes are connected can be obtained from the results in Section 3.2.2.1 by substituting M by 1 and μ by μM. Consequently, the different expressions for the distribution and the expectation of the delivery delay and the total packet transmission can be obtained by simple substitutions to those reported in the case of disconnected throwboxes.

3.3 Asymptotic analysis

Here, we conduct an asymptotic analysis of the expected delivery delay and of the expected number of copies generated under the MTR protocol as the number of mobile nodes and/or the number of throwboxes are large. Note that under the ER protocol, the analysis was hard to carry on and we were not able to derive its asymptotic behavior analytically ; even though, we describe the corresponding analysis briefly at the end of this section.

To proceed with the asymptotic analysis, we will use a mean-field approach [66]. In the context of MANETs, this approach was recently used in [94] and [106] to study the ER protocol. In these papers, the authors model the packet relaying in a MANET by ordinary differential equations (ODEs) by making an analogy with the infection of disease in a biological environment.

We still consider the stochastic model introduced at the beginning of Section 3.2 : $N+1$ mobile nodes, M throwboxes, and the successive meeting times between a mobile node and a throwbox (resp. between two mobile nodes) form a Poisson process with intensity μ (resp. λ), all these Poisson processes being mutually independent.

3.3.1 MTR protocol

We first address the case where the throwboxes are fully disconnected. Let $Y(t)$ (resp. $X(t)$) denote the expected number of throwboxes (resp. expected number of mobile nodes including the source) that hold a copy of the packet at time t or, equivalently, that have been infected at time t to reinforce the analogy with the spreading of an epidemic. When MTR is enforced, the rate at which mobile nodes are infected at time t is equal to λ, the intensity at which a mobile node meets the source, multiplied by $N - X(t)$, the number of mobile nodes yet to be infected at time t. Similarly, the infection rate of throwboxes at time t is equal to μ, the intensity at which the source meets a throwbox, multiplied by $M - Y(t)$, the number of throwboxes yet to be infected at time t.

This gives rise to the following set of ODEs :

$$\dot{X}(t) = \lambda(N - X(t)) \tag{3.16}$$

$$\dot{Y}(t) = \mu(M - Y(t)) \tag{3.17}$$

for $t > 0$. Solving these ODEs with the initial conditions $X(0) = 1$ and $Y(0) = 0$ (these conditions reflect the assumption that at time $t = 0$ only the source carries the packet), yields

$$X(t) = N - (N-1)e^{-\lambda t}, \quad Y(t) = M - Me^{-\mu t}, \quad t \geq 0.$$

Let $F(t)$ denote the cumulative distribution function of the delivery delay T^d_{MTR} at instant t, i.e. $F(t) = P(T^d_{MTR} < t)$, with $F(0) = 0$. In a small time interval of length h, the probability that a node, whether it is fixed or mobile, infects the destination is equal to $(\lambda X(t) + \mu Y(t))h$. Hence, similarly to [94], $F(t)$ writes as

$$\dot{F}(t) = (\lambda X(t) + \mu Y(t))(1 - F(t)), \quad t > 0. \tag{3.18}$$

Solving for (3.18) with the initial condition $F(0) = 0$ gives

$$F(t) = 1 - e^{-\int_0^t (\lambda X(u) + \mu Y(u))du}, \quad t \geq 0 \tag{3.19}$$

From (3.19) we find that the expected delivery delay is given by

$$\mathbf{E}[T^d_{MTR}] = \int_0^\infty e^{-\int_0^t (\lambda X(u) + \mu Y(u))du} dt.$$

Observe that when either N or M is large enough, then $e^{-\int_0^t (\lambda X(u) + \mu Y(u))du}$ is extremely small except near $t = 0$, where the main contribution to the above integral comes from. Therefore, by expanding the function $-\int_0^t (\lambda X(u) + \mu Y(u))du$ in Taylor series in the vicinity of $t = 0$ at the order three, the average delivery delay can be approximated by

$$\begin{aligned}
\mathbf{E}[T^d_{MTR}] &\approx \int_0^\infty e^{\frac{-(\lambda^2(N-1)+\mu^2 M)t^2}{2}} e^{\frac{(\lambda^3(N-1)+\mu^3 M)t^3}{6}} dt \\
&\approx \int_0^\infty e^{\frac{-(\lambda^2(N-1)+\mu^2 M)t^2}{2}} (1 + \frac{(\lambda^3(N-1)+\mu^3 M)t^3}{6}) dt \\
&\approx \sqrt{\frac{\pi}{2(\lambda^2(N-1)+\mu^2 M)}} + \frac{\lambda^3(N-1)+\mu^3 M}{3(\lambda^2(N-1)+\mu^2 M)^2}
\end{aligned} \tag{3.20}$$

when either N or M is large.

If $M = o(N)$ (i.e. $\lim_{N\to\infty} \frac{M}{N} \to 0$) then, as expected, we retrieve the asymptotic expression for the expected delivery delay reported in [41] and [106], that is,

$$\mathbf{E}[T^d_{MTR}] \approx \frac{1}{\lambda}\left(\sqrt{\frac{\pi}{2N}} + \frac{1}{3\lambda N}\right) = O\left(\frac{1}{\sqrt{N}}\right). \tag{3.21}$$

In this case, the introduction of throwboxes only marginally improves the performance of the MTR protocol. On the other hand, if $N = o(M)$ (i.e. $\lim_{M\to\infty} \frac{N}{M} \to 0$) then

$$\mathbf{E}[T^d_{MTR}] \approx \frac{1}{\mu}\left(\sqrt{\frac{\pi}{2M}} + \frac{1}{3M}\right) = O\left(\frac{1}{\sqrt{M}}\right). \tag{3.22}$$

Also of interest is the case where N and M grow to infinity in the same proportion, that is typically $N = \theta M$ with $\theta > 0$. From (3.20) we see that in this case

$$\mathbf{E}[T^d_{MTR}] \approx \frac{\lambda^3\theta + \mu^3}{3\,(\lambda^2\theta + \mu^2)^2}\,\frac{\theta}{N}.$$

Let us briefly consider the expected number of copies generated until the delivery of the packet to the destination. Clearly, it is given by $\int_0^\infty (X(t) + Y(t))dF(t)$, which yields (same formula as the one established at the end of Section 3.2.1)

$$\mathbf{E}[G^d_{MTR}] = (\lambda N + \mu M)\mathbf{E}[T^d_{MTR}]. \tag{3.23}$$

From (3.23) and the asymptotic formulas derived earlier for $\mathbf{E}[T^d_{MTR}]$ as N and/or M are large, we can derive asymptotics for $\mathbf{E}[G^d_{MTR}]$ as N and/or M are large.

When the throwboxes are fully connected, the same analysis yields

$$\dot{X}(t) = \lambda(N - X(t)) \quad (3.24)$$

$$\dot{Y}(t) = \mu M(1 - Y(t)) \quad (3.25)$$

$$\dot{F}(t) = (\lambda X(t) + M\mu Y(t))(1 - F(t)) \quad (3.26)$$

and hence, $\mathbf{E}[T^c_{MTR}]$ reads as

$$\mathbf{E}[T^c_{MTR}] \approx \sqrt{\frac{\pi}{2(\lambda^2(N-1) + \mu^2 M^2)}} + \frac{\lambda^3(N-1) + \mu^3 M^3}{3(\lambda^2(N-1) + \mu^2 M^2)^2}. \quad (3.27)$$

In this case, the "cut-off value" occurs when M is of the order of $\sqrt{N}$ as opposed to M being of the order of N as in the case where the throwboxes are all disconnected.

3.3.2 Epidemic routing protocol

The same analysis can be done for the ER protocol. When throwboxes are disconnected, the following ODEs hold

$$\dot{X}(t) = \lambda X(t)(N - X(t)) + \mu Y(t)(N - X(t)) \quad (3.28)$$

$$\dot{Y}(t) = \mu X(t)(M - Y(t)) \quad (3.29)$$

$$\dot{F}(t) = (\lambda X(t) + \mu Y(t))(1 - F(t)) \quad (3.30)$$

As opposed to the system of differential equations for the case of MTR, we could not be able to obtain explicit expressions for the different concerned means at the exception of the mean number of infected mobile nodes $\mathbf{E}[X(t)]$. Specifically, through an integration by part of equation (3.28) then through a substitution in equation (3.30), $\mathbf{E}[X(t)]$ reads as

$$\mathbf{E}[X(t)] = \frac{N+1}{2} \quad (3.31)$$

which is equal to the mean number of infected mobile nodes in an equivalent network without throwboxes (see [41] and [106]).

3.4 Simulation results

In this section, we report numerical resutls to validate the analytical models and results for the various schemes that have been considered throughout the previous sections (MTR/ER protocols, fully connected/disconnected throwboxes). Hereafter in the next section, we first address the validation of our model and investigate the impact on the performance of the presence of uniformly distributed throwboxes. Then, for both the MTR and the ER protocols, we investigate the effect on their performance induced by two different throwbox deployment scenarios.

Throughput the subsequent simulations, we have considered the following common settings. The mobility of the nodes has been simulated by using the code available at [87].

This code avoids the various pitfalls associated with the simulations of the RWP/RD mobility models. For the sake of clarity, we only present results obtained for the random waypoint mobility model. However, we point out that the results obtained under the random direction mobility model are as accurate as those obtained for the RWP model. Mobile node speed v is held constant and is equal to $v = 10m/sec$, and nodes do not pause at the end of a movement.

Each node is assigned a constant and fixed transmission range $r = 100m$. The transmission of a packet between two nodes is assumed to occur instantaneously when the nodes are located at a distance less than or equal to r of one another.

The intensity μ of the meeting times between a mobile node and a throwbox is taken according to Equation 2.6 when throwboxes are uniformly distributed within the area and to Equation 2.7 when they are distributed according to the stationary distribution of the RWP model. The intensity λ of the meeting times between two mobile nodes is taken from [41], and it is given by $\lambda = 10.94\frac{rv}{\pi L^2}$ where L is the area side length.

3.4.1 Validation

To validate our models, we have considered small sized networks, where throwboxes are *uniformly distributed* in a square area of size $2 \times 2\ Km^2$.

Figures 3.3-3.4 compare the expected delivery delay obtained through simulations (solid lines) to those obtained by the Markov models (dashdot lines), under both the MTR and the ER protocols, when throwboxes are either fully connected or fully disconnected. Three sets of networks have been considered in which the first one is composed of an equal number of mobile nodes and throwboxes, the second one is composed only by throwboxes with the mobile source and the mobile destination, and the third one is composed only by mobile nodes (no throwbox). For the different considered cases, we observe that our model is able to accurately predict the expected delivery delay with a small relative error when the number of nodes becomes large (this is a consequence of the fact that as the number of nodes increases then the inter-meeting times are no longer mutually independent). We also observe that when relaying is performed according to the MTR protocol and throwboxes are connected, the expected delivery delays in networks where $N = M$ and where there are no relay nodes are very close whenever $\sqrt{N} << M$. This observation is in agreement with the asymptotic results obtained in Section 3.3.

Results for the expected number of copies generated until the packet delivery are reported in Figures 3.5-3.6. Here too there is a good fit between the analytical models and the simulations, across all scenarios that have been considered.

For either relay protocols, we can also assess the impact of having throwboxes on the expected delivery delay by distinguishing the cases where $M > 0$ from the case where $M = 0$ in Figures 3.3-3.4.

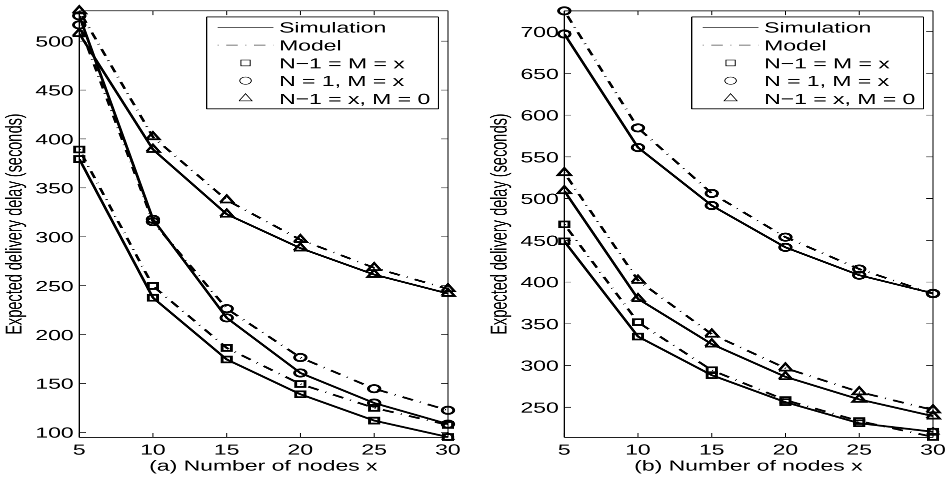

FIG. 3.3 – $\mathbf{E}[T_{MTR}]$: **(a)** Throwboxes are connected, **(b)** Throwboxes are disconnected.

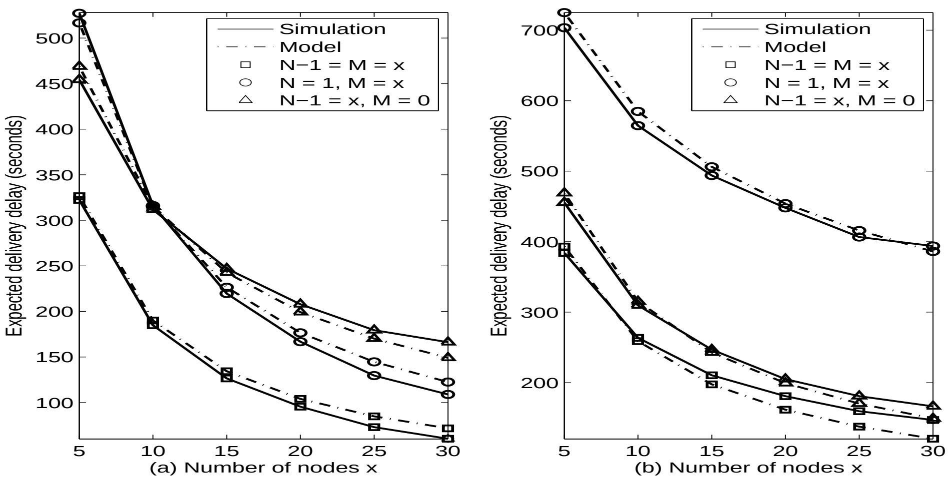

FIG. 3.4 – $\mathbf{E}[T_{ER}]$: **(a)** Throwboxes are connected, **(b)** Throwboxes are disconnected.

3.4.2 Impact of different throwbox deployments

The aim of this section is to compare, for the performance of both MTR and the ER protocols, the impact of two different spatial deployments of throwboxes. We consider a large network of size $10 \times 10\ Km^2$ with an equal and varying number of mobile relay nodes and throwboxes. Figure 3.7 plots the expected delivery delay under both the MTR and the ER protocols for two different distributions of throwboxes : the uniform distribution and the stationary distribution of the RWP model [68]. Recall that in the latter distribution, the throwboxes will be mainly distributed near the center of the area. Since the nodes are moving according to the RWP model, it is not too surprising to observe that the best performance is obtained when the throwboxes are deployed according to the stationary distribution of this mobility model. Note also the excellent agreement between the analytical results and

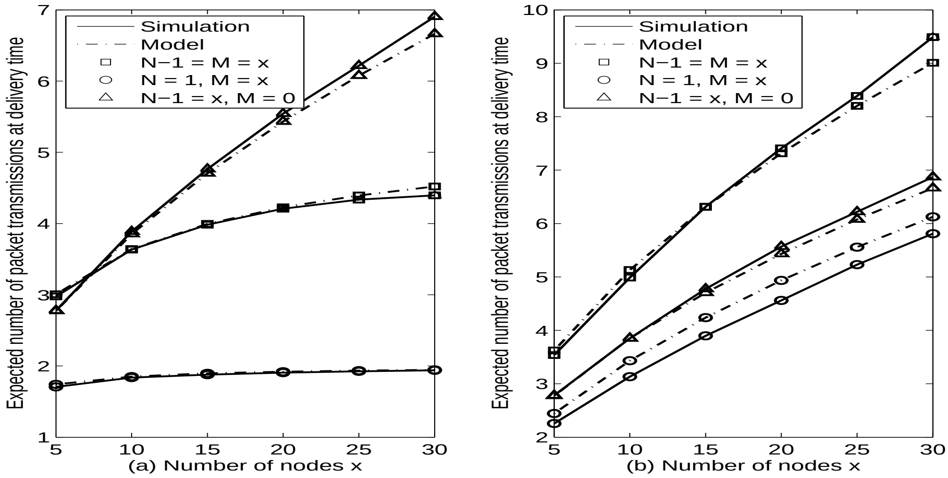

FIG. 3.5 – $\mathbf{E}[G_{MTR}]$: **(a)** Throwboxes are connected, **(b)** Throwboxes are disconnected.

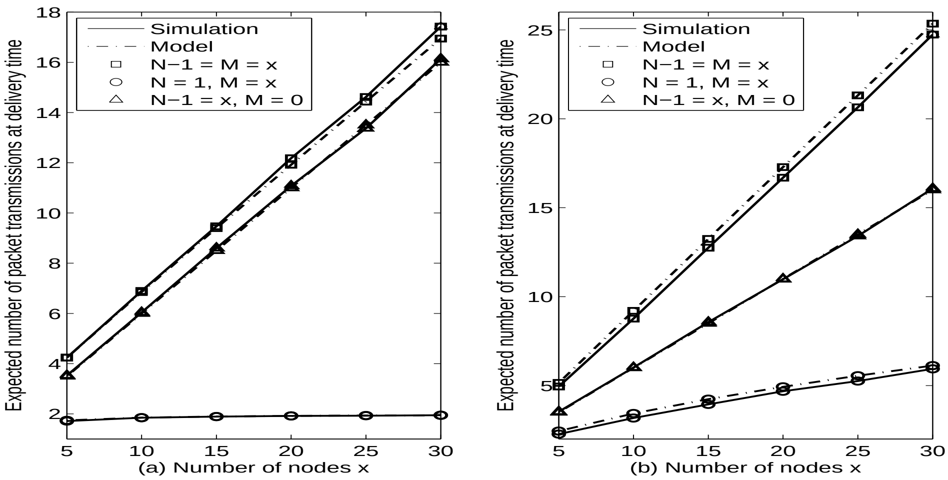

FIG. 3.6 – $\mathbf{E}[G_{ER}]$: **(a)** Throwboxes are connected, **(b)** Throwboxes are disconnected.

the simulation results.

3.5 Conclusion

In this chapter, we have developed Markovian models to analyze the impact on the performance of sparse MANETs (delivery delay, resource consumption) of adding throwboxes. These models are based on the assumptions that (i) the successive meeting times between any pair of mobile nodes (resp. between a mobile node and a throwbox) form a Poisson process, and (ii) that these Poisson processes are mutually independent. We have considered two different store-carry-and-forward protocols, the multicopy two-hop routing protocol and the epidemic routing protocol. Both the situations where the throwboxes are

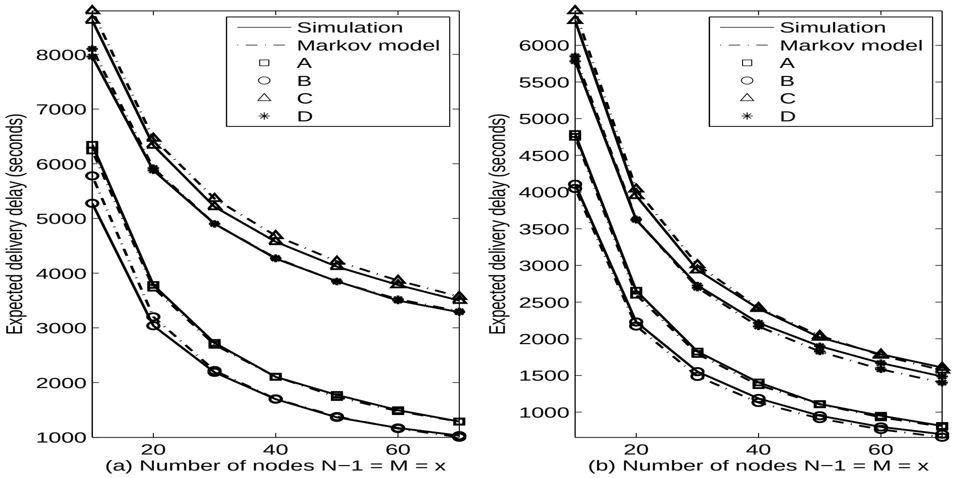

FIG. 3.7 – $\mathbf{E}[T_{MTR}]$**(a)** and $\mathbf{E}[T_{ER}]$ **(b)** : **(A)** Throwboxes are connected and uniformly distributed, **(B)** Throwboxes are connected and stationary distributed, **(C)** Throwboxes are disconnected and uniformly distributed, **(D)** Throwboxes are disconnected and stationary distributed.

fully disconnected and mesh connected have been studied. The derived results have been validated through simulations when nodes are moving either according to the random waypoint model or to the random direction model. In all cases, we have found excellent matches between the results predicted by the analytical models and the simulations.

This chapter aimed to develop analytical models to study the delay-energy performance of two commonly used relay protocols for sparse MANETs in presence of throwboxes. In the next chapter, we develop a general analytical framework which allows to study the delay-energy performance and the underlying trade-offs of a wide set of relay protocols in a similar network setting.

Chapitre 4

Performance Evaluation Framework for Throwbox Embedded DTNs

Contents

The objective of the present chapter is to give guidelines on the design and the evaluation of routing techniques for opportunistic delay tolerant networks augmented with throwboxes. We consider a network model comprising both mobile relays and throwboxes, and we introduce a framework which allows to evaluate analytically the performance of a large set of routing techniques in such networks. We then propose five routing strategies for such network settings where each strategy trades off delivery delay efficiency versus energy consumption. Using the analytical framework, we evaluate the two latter performance metrics for each one of the proposed strategies. The obtained results highlight the various design solutions that are left to network designers, and show how the proposed framework can be applied.

4.1 Introduction

In this chapter, we consider the issue of routing in opportunistic disruption tolerant networks that are equiped with throwboxes. By opportunistic, we mean that mobile nodes have zero knowledge regarding their past and future contact opportunities with other mobile nodes or with throwboxes.

With respect to this networking scenario, our goal is to design and evaluate the efficiency of multi-copy routing schemes that exploit the presence of throwboxes in the network to speed up packet dissemination, and consequently to enhance network performance. The underlying design objectives aim to enhance the latency of messages delivery while minimizing the number of transmissions that need to be made by the mobile relays since throwboxes are assumed to have enhanced power and storage capabilities. Concerned with this task, we propose a design methodology which is based on three consecutive steps :

1. First, we discuss the challenges related to the design of opportunistic disruption tolerant networks, illustrating the various aspects related to the operation of opportunistic disruption tolerant networks and highlighting the performance metrics that the corresponding applications require.
2. Next, we introduce an analytical framework which allows us to evaluate the performance of a wide set of routing strategies and system configurations without complex and time-consuming simulations. The latter framework allows to easily model and evaluate a wide range of opportunistic communication scenarios, accounting for the network configuration (e.g., number of mobile relays and throwboxes) and for a wide range of relaying schemes.
3. Then, we consider an opportunistic network comprising both throwboxes and mobile relays, and propose and evaluate through the introduced analytical framework five routing strategies.

In the literature, there are few works that have considered the issue of routing in throwbox-embedded DTNs. To the best of our knowledge, the pioneer work in this domain is the work of [113], where the authors have focused on improving the network throughput using throwboxes. They have considered three different deployment scenarios for throwboxes, and for each scenario they evaluated three different routing scenarios, namely an epidemic, single path and multi-path routing schemes. Using `ns-2` simulations, they have provided several insights to guide the design and deployment of throwboxes within such networks. In [52] where the results are reported in the previous chapter, we have focused on studying the effectiveness of adding throwboxes on the performance of epidemic and multi-copy two hop relay protocols, but we did not consider the issue of designing routing approaches for these networks. The obtained results have shown that, under the epidemic protocol, the performance of the network increases greatly in the presence of throwboxes at the cost of increasing the number of copy transmissions in the network. On the other hand, minor improvement is observed under the two-hop relaying protocol when throwboxes are disconnected.

The remainder of the chapter is organized as follows. Section 4.2 describes the challenges related to the design of opportunistic DTNs, and outlines the various performance requirements that need to be taken into account. Section 4.3 introduces our analytical framework for evaluating the performance of opportunistic DTNs, and details the necessary steps to be performed for applying it. Section 4.4 applies such framework to the evaluation of routing trade-offs in throwbox-embedded opportunistic DTNs. Section 4.5 discusses the network design space that is possible to explore with the proposed framework. Finally, Section 4.6 provides the concluding remarks.

4.2 Designing an opportunistic disruption tolerant network

This section highlights the key factors that impact the design and the main performance requirements for developping an opportunistic network that is operating in a disconnected regime and using opportunistic forwarding for achieving network communications [23, 1].

4.2.1 Key design factors

The first step when designing an opportunistic DTN is to account for all the specific requirements arising from the scenario at hand, and to understand the space that is left to network designers. In the following, we provide a taxonomy of the many aspects contributing to the operation of an opportunistic communication system :

- **network infrastructure** : the basic infrastructure of an opportunistic DTN is composed by (i) *relaying nodes*, which are mobile nodes generating, relaying and receiving messages. Their number determines the "density" of the network, which is then reflected in the contact opportunities and therefore in the speed of the message diffusion process. (ii) *throwboxes*, which are dedicated nodes, with greater resources in terms of storage, battery and processing capabilities. Their utilization can greatly increase the performance of the system, especially in the case where relaying nodes are constrained by extremely limited resources. In this case, throwboxes are placed in strategic positions of the served area, and act as fixed relays of the network. Where needed, throwboxes can be connected, thus providing a backbone for the diffusion of data.
- **forwarding strategy** : one of the major problems in disconnected mobile networks is to envision efficient mechanisms for disseminating information and letting messages eventually reach their intended destination [47] without consumming much resources. A specific forwarding strategy regulates the way messages are relayed from one node to the other, and represents one of the key design choices of any opportunistic communication system.
- **mobility pattern of nodes** : at the basis of the message diffusion pattern there are opportunistic communications among nodes, in which a message is relayed when nodes get within mutual communication range. The diffusion of messages is therefore driven by the underlying mobility pattern. Mobility can be constrained in a spatial region, as in the case of a city-wide deployment, can be heterogeneous, as in

the case of pocket switched networks where nodes mobility is associated with the social behavior of users, or can be periodic, as in the case of opportunistic communications systems constituted by buses [107]. Each mobility pattern leads to a different performance of the system, and needs to be deeply investigated during the system design phase.
- **communication technology** : the specific communication technology determines (i) when two nodes are within mutual communication range (ii) the amount of data that is possible to exchange at any contact opportunity. Clearly, this highly impacts the networking performance of these systems, and needs to carefully evaluated.

4.2.2 Application requirements

Most of the applications related to the research in opportunistic DTNs arise from considerations relative to scenarios where devices may be disconnected due to some physical constraints. As an example, we can consider the use of buses and other public transportation means for carrying and disseminating information [20]. Such systems may be used, e.g., to diffuse information about traffic situation, parking availability, local advertisement and similar tasks. Alternatively, people, with their personal devices, can be carriers of the network [30]. This gives rise to many social networking applications, where information is diffused on the basis of the physical movement of people.
Independently from the specific scenario, the metrics that mostly affect the performance of such systems, and that needs to be taken into consideration when evaluating the requirements of a related application can be summarized as follows :
- **delivery delay** : this is the average time that is needed for delivering a message from a source to a destination. This metric is extremely important as it reflects the dynamic behavior of the system, and in particular how efficient a certain routing solution is in letting messages reach their intended destination. Clearly, this metric is affected by many aspects such as, e.g., network infrastructure, mobility pattern of nodes, etc, and is therefore one of the most difficult metrics to evaluate at design phase.
- **utilized resources** : forwarding operations rely on the ability of a node to keep (even for a rather long time) a message in its internal memory. This is justified by the fact that a node may be doomed to remain isolated for a long time, but should still be able to forward the messages it received. Clearly, the larger the number of copies of a message in the system, (i) the faster it reaches its destination (ii) the more it is robust with respect to the nodes mobility and node/link failures. On the other hand, in order to have more copies of the same message traveling in the network at the same time, a larger amount of network resources has to be exploited. However, resources are typically scarce, especially on mobile nodes, and it is therefore an application requirement that needs to be specified when considering the application constraints.
- **packet delivery ratio** : due to the intrinsic randomness of the many factors influencing the performance of opportunistic networks, only a fraction of the sent packets can reach their intended destinations. For instance, due to the limited re-

sources available on nodes it can occur that messages are discarded and eventually not reaching the intended destination. For some application scenarios, where reliability is not a critical, this can be tolerated. The packet delivery ratio provides a measure of how likely is this phenomenon to occur.

4.3 Modeling framework

In this section, we define the network model we are considering, the performance metrics we are interested in and a Markovian-based framework that evaluates analytically these metrics for a set of routing strategies running over our network model.

4.3.1 Model assumptions

We consider the same network model that we have considered in Chapter 3. The following is a summary of the model :

- The network is composed of $N+1$ identical mobile nodes and M identical throwboxes, all located in a two-dimensional square area $\mathcal{A}$. The network is sparse and operates under intermittent connectivity conditions.
- Among the $N+1$ nodes, there is one source and one destination nodes. The source has a packet to send to the destination and the remaining $N-1$ mobile nodes as well as the M throwboxes will assist in copies spreading according to a given relay protocol.
- A transmission occurs when two nodes come within each other transmission range. Transmissions are assumed to be instantaneous and successful.
- Unless otherwise mentioned, throwboxes are assumed to be disconnected from each other, and they are deployed in the area $\mathcal{A}$ according to some random distribution, for instance a uniform distribution.
- Mobile nodes are assumed to move independently of each other according to some random mobility model for which the observations made in [41] and in Section 2.3 of independent and exponentially distributed inter-meeting times between a pair of nodes hold.
- In particular, inter-meeting times between any pair of mobile nodes are independent identically distributed (i.i.d.) r.v.s with a common exponential distribution with rate λ.
- Inter-meeting times between a pair of a mobile node and a throwbox are i.i.d. r.v.s with a common exponential distribution with rate μ.

With respect to that model, we are interested in evaluating the performance of several proposed routing strategies in terms of delivery delay and energy consumption at the mobile nodes. Henceforth, we use the subscript s to refer to a given routing strategy s.

Correspondingly, we define the r.v. T_s as the delivery delay of the packet defined as the elapsed time from the packet generation time at $t=0$ till one of its copies is delivered to the destination. We denote by $\mathbf{E}[T_s]$ its expected value. We define the r.v. G_s as the number of *valuable* copy transmissions made during time T_s. Those *valuable* copy transmissions comprise those made only by the mobile nodes, source included. G_s includes also the potential packet transmission to the destination that would be made at delivery instant by the source or one of the mobile nodes. Throwboxes are assumed to have sufficient power and storage capabilities than mobile nodes, and therefore, the number of their copy transmissions during T_s is of no interest to our study of energy consumption. Let $\mathbf{E}[G_s]$ denote the expected value of G_s. Observe the difference with the mean number of transmissions that we have considered in the previous chapter where it accounts in addition for those transmissions made by the throwboxes.

4.3.2 Analytical framework

In this section, we introduce an analytical framework which allows us to evaluate the performance of various routing strategies that would be applied on our network model without complex and time consuming simulations. Considering a routing strategy s, the framework solves analytically the corresponding $\mathbf{E}[T_s]$ and $\mathbf{E}[G_s]$ by mapping them to some variables and then by solving these variables iteratively. Note that the framework can be applied to other general network models as long as the inter-meeting times between nodes are i.i.d. exponentially distributed.

The framework is based on an absorbing three dimensional continuous time Markov chain noted $\mathbf{A_s(t)} = (I_s(t), J_s(t), K_s(t))$ which is defined as follows. For $0 \leq t < T_s$, the first component $I_s(t) \in \{1, 2, \ldots, N\}$ represents the number of mobile relays, source included, which have been infected[1] directly by the source during time t. The second component $J_s(t) \in \{0, 1, \ldots, N - I_s(t)\}$ indicates the number of mobile relays which have been infected by throwboxes during time t. And last, the third component $K_s(t) \in \{0, 1, \ldots, M\}$ represents the number of infected throwboxes during time t. For $t \geq T_s$, the packet had been delivered to the destination and $\mathbf{A_s(t)} = \{a\}$ where $\{a\}$ denotes the absorbing state.

Let $\mathcal{E}_s$ and $\mathbf{Q_s}$ denote respectively the state space and the infinitesimal generator of $\mathbf{A_s(t)}$ under routing strategy s. Figure 4.1 plots the various possible outgoing transitions from a given state (i, j, k). Specifically, following a sojourn time in state (i, j, k) which is exponentially distributed with rate $-q((i, j, k), (i, j, k))$ noted hereafter as $S(i, j, k)$, the chain jumps either into state $(i + 1, j, k)$ with rate $q((i, j, k), (i + 1, j, k))$ noted as $\alpha(i, j, k)$, or into state $(i, j + 1, k)$ with rate $q((i, j, k), (i, j + 1, k))$ noted as $\beta(i, j, k)$, or into state $(i, j, k + 1)$ with rate $q((i, j, k), (i, j, k + 1))$ noted as $\gamma(i, j, k)$. In addition, it is absorbed with rate $q((i, j, k), a)$ noted as $\theta(i, j, k)$.

Once the state space $\mathcal{E}_s$ and the transition rates of the generator $\mathbf{Q_s}$ corresponding to the considered routing strategy is well defined, we can proceed with the framework to compute $\mathbf{E}[T_s]$ and $\mathbf{E}[G_s]$. To compute these metrics, we follow a recursive approach similar to the one developed in [63, Chapter 7] for two-dimensional Markov chain with absorbing states where here we extend it to the three-dimensional Markov chain $\mathbf{A_s(t)}$. The approach proceeds as in the following.

Let F_s denote a three-dimensional matrix representing a given performance metric.

[1]We recall that by infected we mean that the node that is carrying a copy of the packet, termed as an infected node, has transmitted a copy of it to a susceptible node.

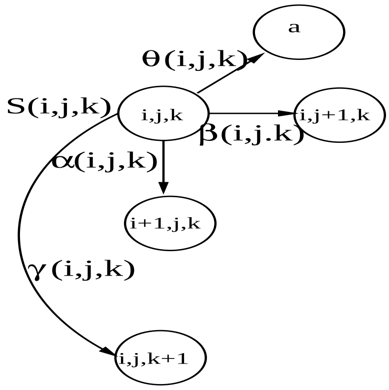

FIG. 4.1 – Transition rates out of state (i,j,k).

The size of F_s, noted as Dim_s, is defined as

$$Dim_s = max[I_s(t)] \times max[J_s(t)] \times max[K_s(t)] \tag{4.1}$$

where $max[V_s(t)]$ denotes the maximum value that can take the r.v. $V_s(t)$ under strategy s. The entries of the matrix F_s are defined as follows. Starting from state (i, j, k), let $F_s(i, j, k)$ denote the mean value of the considered performance metric till absorption under routing strategy s. In accordance with this definition, and by considering the possible transitions going out from state (i, j, k), $F_s(i, j, k)$ satisfies the following general recursive equation :

$$\begin{aligned} F_s(i,j,k) &= u_s(i,j,k) + \tfrac{\alpha(i,j,k)}{S(i,j,k)} F_s(i+1,j,k) \\ &+ \tfrac{\beta(i,j,k)}{S(i,j,k)} F_s(i,j+1,k) + \tfrac{\gamma(i,j,k)}{S(i,j,k)} F_s(i,j,k+1) \end{aligned} \tag{4.2}$$

where $u_s(i, j, k)$ denotes the mean value of the considered performance metric at state (i, j, k). What we are interested in is to compute $F_s(i, j, k)$ starting from the initial state which would result hence in the mean value of the corresponding metric till the absorption time. Therefore, to compute $F_s(1, 0, 0)$, the approach starts first by computing $F_s(i, j, k)$ at the last state of $\mathbf{A_s(t)}$. The value of the metric at that state is generally known and it is equal to its mean value $u_s(i, j, k)$ at that state. Consequently, Equation (4.2) can be iterated back state by state till reaching the initial state $(1, 0, 0)$. **Algorithm 1** shows an implementation of the preceding approach to compute $F_s(1, 0, 0)$. Note that *last_value* in the algorithm could be either $N - 1$ or 0 depending on the state space of the considered strategy.

Now, we apply **Algorithm 1** in order to derive $\mathbf{E}[T_s]$ and $\mathbf{E}[G_s]$. To compute $\mathbf{E}[T_s]$, we define the matrix $\mathbf{T_s}$ with entry $T_s(i, j, k)$ representing the mean sojourn time of $\mathbf{A_s(t)}$ till the absorption starting from state (i, j, k). The size of $\mathbf{T_s}$ is given according to Equation (4.1). The mean sojourn time in state (i, j, k) is equal to the sum of the total rates going out from that state which is equal to $\frac{1}{S(i,j,k)}$, and hence Equation (4.2) translates into

$$\begin{aligned} T_s(i,j,k) &= \tfrac{1}{S(i,j,k)} + \tfrac{\alpha(i,j,k)}{S(i,j,k)} T_s(i+1,j,k) \\ &+ \tfrac{\beta(i,j,k)}{S(i,j,k)} T_s(i,j+1,k) + \tfrac{\gamma(i,j,k)}{S(i,j,k)} T_s(i,j,k+1) \end{aligned} \tag{4.3}$$

Algorithm 1 : Computes the mean value of the metric F starting from state $(1, 0, 0)$ till the absorption time

Require: N, M, λ and μ
Ensure: $F_s(1, 0, 0)$
for $k = M$ to 0 **do**
  for $j = last\_value$ to 0 **do**
    for $i = (N - j)$ to 1 **do**
      $\alpha(i, j, k) = q((i, j, k), (i + 1, j, k))$
      $\beta(i, j, k) = q((i, j, k), (i, j + 1, k))$
      $\gamma(i, j, k) = q((i, j, k), (i, j, k + 1))$
      $\theta(i, j, k) = q((i, j, k), a)$
      $S(i, j, k) = \alpha(i, j, k) + \beta(i, j, k) + \gamma(i, j, k) + \theta(i, j, k)$
      if $(k = M$ and $(i + j) = N)$ **then**
        $F_s(i, j, k) = u_s(i, j, M)$
      else
        $F_s(i, j, k) = u_s(i, j, k) + \frac{\alpha(i,j,k)}{S(i,j,k)} F_s(i + 1, j, k) + \frac{\beta(i,j,k)}{S(i,j,k)} F_s(i, j + 1, k) + \frac{\gamma(i,j,k)}{S(i,j,k)} F_s(i, j, k + 1)$
      end if
    end for
  end for
end for
return $F_s(1, 0, 0)$ {expected value of the metric F during T_s}

with the value of $T_s(i, j, k)$ at the last state is equal to $T(last\_state) = \frac{1}{S(last\_state)} = \frac{1}{\theta(last\_state)}$. With respect to this definition, the mean delivery delay $\mathbf{E}[T_s]$ is nothing than the mean sojourn time starting from state $(1, 0, 0)$ till the absorption which is equal to $T_s(1, 0, 0)$. The latter variable can be easily computed through **Algorithm 1**. Thus, we have

$$\mathbf{E}[T_s] = T_s(1, 0, 0) \tag{4.4}$$

Recall that the mean number of *valuable* copy transmissions $\mathbf{E}[G_s]$ is defined as the mean number of those transmissions that are made only by the source and by the mobile nodes during T_s. In respect to that, $\mathbf{E}[G_s]$ is consequently equal to the sum of the following four components :

- The mean number of mobile relays which are infected by the source and by other mobile relays. In the following, we address routing strategies where the only ways to infect a mobile relay is either directly by the source or through an infected throwbox. Thus, we need to consider as valuable transmissions only those made by the source to mobile relays.
- The mean number of infected throwboxes. Basically, the only way to infect a throwbox is through the source or a mobile node.
- The probability that the source is the node who delivers the message to the destination.
- The probability that a mobile relay is the node who delivers the message to the

destination (in the case when routing strategy s allows that interaction).

Therefore, to evaluate $\mathbf{E}[G_s]$, we introduce four matrices where each matrix corresponds to one of the four previous metrics. Similarly to matrix $\mathbf{T_s}$, the size of each one of these matrices is given by Equation (4.1). Hence, we define matrix $\mathbf{RbyS_s}$ with entry $RbyS_s(i,j,k)$ representing the mean number of mobile relays that are infected by direct transmissions from the source starting from state (i,j,k) till the absorption time. Note that $RbyS_s(i,j,k)$ does not include those relays who are infected by the throwboxes. Using Equation (4.2), $RbyS_s(i,j,k)$ reads as

$$\begin{aligned} RbyS_s(i,j,k) &= i\frac{\theta(i,j,k)}{S(i,j,k)} + \frac{\alpha(i,j,k)}{S(i,j,k)}RbyS_s(i+1,j,k) \\ &+\frac{\beta(i,j,k)}{S(i,j,k)}RbyS_s(i,j+1,k) + \frac{\gamma(i,j,k)}{S(i,j,k)}RbyS_s(i,j,k+1) \end{aligned} \tag{4.5}$$

We define matrix $\mathbf{B_s}$ with entry $B_s(i,j,k)$ representing the mean number of infected throwboxes starting from state (i,j,k) till the absorption time. Using Equation (4.2), $B_s(i,j,k)$ reads as

$$\begin{aligned} B_s(i,j,k) &= k\frac{\theta(i,j,k)}{S(i,j,k)} + \frac{\alpha(i,j,k)}{S(i,j,k)}B_s(i+1,j,k) \\ &+\frac{\beta(i,j,k)}{S(i,j,k)}B_s(i,j+1,k) + \frac{\gamma(i,j,k)}{S(i,j,k)}B_s(i,j,k+1) \end{aligned} \tag{4.6}$$

Also, we define matrices $\mathbf{PrS_s}$ and $\mathbf{PrR_s}$ with entries $PrS_s(i,j,k)$ and $PrR_s(i,j,k)$ as respectively the probability of delivery by the source and by a mobile relay given that the chain starts from state (i,j,k). Using Equation (4.2), $PrS_s(i,j,k)$ reads as

$$\begin{aligned} PrS_s(i,j,k) &= \lambda\frac{1}{S(i,j,k)} + \frac{\alpha(i,j,k)}{S(i,j,k)}PrS_s(i+1,j,k) \\ &+\frac{\beta(i,j,k)}{S(i,j,k)}PrS_s(i,j+1,k) + \frac{\gamma(i,j,k)}{S(i,j,k)}PrS_s(i,j,k+1) \end{aligned} \tag{4.7}$$

while $PrR_s(i,j,k)$ reads as

$$\begin{aligned} PrR_s(i,j,k) &= \lambda\frac{(i+j-1)}{S(i,j,k)} + \frac{\alpha(i,j,k)}{S(i,j,k)}PrR_s(i+1,j,k) \\ &+\frac{\beta(i,j,k)}{S(i,j,k)}PrR_s(i,j+1,k) + \frac{\gamma(i,j,k)}{S(i,j,k)}PrR_s(i,j,k+1) \end{aligned} \tag{4.8}$$

The mean number of *valuable* copy transmissions $\mathbf{E}[G_s]$ is the sum of $RbyS_s(i,j,k)$, $B_s(i,j,k)$, $PrS_s(i,j,k)$ and $PrR_s(i,j,k)$ when the chain $\mathbf{A_s(t)}$ starts from state $(1,0,0)$. By applying **Algorithm 1** on each one of these variables, we get [2]

$$\begin{aligned} \mathbf{E}[G_s] &= (RbyS_s(1,0,0)-1) + B_s(1,0,0) + \\ &\quad PrS_s(1,0,0) + PrR_s(1,0,0) \end{aligned} \tag{4.9}$$

It is noteworthy to mention that the five metrics (e.g. T_s, $RbyS_s$, etc.) that we have just defined are in accordance with the two considered performance metrics, i.e. $\mathbf{E}[T_s]$ and $\mathbf{E}[G_s]$, and the routing strategies that we will propose in the next section. However, the framework is general enough to compute the numerical values of other metrics that we did not consider in this work.

[2] The -1 in $\mathbf{E}[G_s]$ stands for the source packet.

4.4 Evaluating routing trade-offs in throwbox-embedded opportunistic networks

We now apply the methodology introduced in Section 4.2, and show how the proposed modeling framework can be used to investigate the performance of different routing schemes which include throwboxes within their infection mechanism. Under these objectives, we propose five routing strategies. Specifically, each one of the considered techniques represents a combination of possible relaying interactions between (i) the source node (ii) the mobile relays (iii) the throwboxes (iv) and the destination node.

Commonly for each strategy, the source node passes a copy of its packet upon a contact with a susceptible throwbox or the destination. In addition, whenever the destination or a susceptible mobile relay meets an infected throwbox, the throwbox will transmit a copy to the mobile node. On the other side, these strategies differ only by the restrictions that we impose on an infected mobile relay for its ability of infecting a susceptible throwbox or the destination. The five strategies with their possible interactions are presented in Table 4.1.

Tab. 4.1 – Possible interactions among various nodes composing opportunistic network : Source, Destination, Throwbox and Relays.

Strategy I	**Strategy II**	**Strategy III**	**Strategy VI**	**Strategy V**
Interactions between source, relay and throwbox (thr.) nodes				
Source → Thr. *Thr. → Relay*	*Source → Thr.* *Thr. → Rel.* *Source → Rel.*	*Source → Thr.* *Thr. → Relay* *Relay → Thr.*	*Source → Thr.* *Thr. → Relay* *Relay → Thr.*	*Source → Thr.* *Thr. → Relay* *Relay → Thr.* *Source → Relay*
Interactions with the *destination* node.				
Source. → Dest. Thr. → Dest. Relay → Dest.	Source → Dest. Thr. → Dest. Relay → Dest.	Source → Dest. Thr. → Dest.	Source → Dest. Thr. → Dest. Relay → Dest.	Source → Dest. Thr. → Dest. Relay → Dest.

In strategy I, the source can make copies to throwboxes or to the destination. However, the only way to infect a mobile relay is via an infected throwbox. Once infected, a mobile relay can then give a copy to the destination.

Strategy II, is a slight extension to strategy I where the source can also generate copies to the mobile nodes. Obviously, in this strategy, there will higher number of transmissions made by the source node than in strategy I at the advantages of decreasing the delivery delay.

In strategy III, mobile nodes act only as relays between the throwboxes. As in the previous two strategies, the source can pass a copy to a throwbox or to the destination. An infected throwbox can hence relay a copy to a mobile node that subsequently will operate in an epidemic scheme upon meeting with other susceptible throwboxes. However, the only way to infect the destination is via the source directly or through an infected throwbox. This strategy can be seen as a routing technique in a non-cooperative network, where a node with a packet to deliver makes copies only to throwboxes. Afterward, infected throwboxes can exploit the movement of the mobile relays inside the network in order to spread the data to other throwboxes to enhance delivery performance.

Strategies IV and V are extensions of the first three basic mechanisms. Specifically, strategy IV is a compound of strategies I and III, while strategy V is a compound of strategies II and III.

4.4.1 Analyzing the scenario

The first step of our design methodology consists in analyzing the specific application scenario under consideration.

We consider a scenario composed by $N+1$ identical mobile nodes that move according to a random mobility model in a relatively large square area $\mathcal{A}$ of size $10 \times 10\ km^2$. We further assume that M identical throwboxes are uniformly distributed within the area $\mathcal{A}$. The transmission range of all nodes is fixed and set equal to $r = 100$ m. In the following, we present results only for the RWP model even though similar results are also obtained for the random direction model. Nodes mobility has been simulated using the code of the random trip model in [87]. Nodes move at a constant speed $v = 10$ m/s and do not pause.

In order to apply **Algorithm 1** of the framework, we need to estimate (i) the expected meeting rate λ between two mobile nodes and (ii) the expected meeting rate μ between a mobile and a throwbox. For the mobility pattern under consideration, namely the random waypoint with constant speed and no pausing, we used the approximation formulas of $\lambda = \frac{8rv\omega}{\pi L^2}$ and $\mu = \frac{2rv}{L^2}$, as given in [41, Lemma 1] and by Equation 2.6, respectively. These formulas have proved to represent a good approximation for sufficiently sparse networks and as long as $r << L$. In a real setting, such values need to be approximated.
Table 4.2 provides a concise summary of the constraints deriving from the scenario under consideration.

TAB. 4.2 – Summary of the scenario under consideration.

Mobility model	RWP
Playground	$10 \times 10\ km^2$
Nodes speed	$v = 10$ m/s
Nodes comm. range	$r = 100$ m
Mobile/mobile meeting rate	$\lambda = \frac{8rv\omega}{\pi L^2}$
Mobile/thr. meeting rate	$\mu = \frac{2rv}{L^2}$
Forwarding Strategy	*Epidemic, Strategy I,II, III,IV and V*

4.4.2 Defining the associated $\mathbf{A_s(t)}$

To evaluate the performance of a chosen routing strategy s, the second step consists in defining the associated state space $\mathcal{E}_s$ and the generator matrix $\mathbf{Q_s}$. Recall that $\mathcal{E}_s$ and $\mathbf{Q_s}$ determine respectively the sequence of transient states with the corresponding transition rates for the considered routing strategy s.

Hereafter, we will present the detailed results for strategies II, while we refer to the appendix for the details of the remaining schemes. In particular, under strategy **II**, the state space $\mathcal{E}_{II}$ is defined as $\mathcal{E}_{II} = \{(i,j,k), i \in (1,2,\ldots,N), j \in (0,1,\ldots,N-i), k \in (0,\ldots,M)\}$, and the non-zero entries of the generator $\mathbf{Q_{II}}$ are given by

$$\begin{aligned}
q((i,j,k),(i+1,j,k)) &= \alpha(i,j,k) = \lambda(N-i-j), \\
& i = 1,\cdots,N-1, \quad j = 0,\ldots,N-i, \quad k = 0,\cdots,M \\
q((i,j,k),(i,j+1,k)) &= \beta(i,j,k) = \mu k(N-i-j), \\
& i = 1,\ldots,N-1, \quad j = 0,\ldots,N-i, \quad k = 1,\cdots,M \\
q((i,j,k),(i,j,k+1)) &= \gamma(i,j,k) = \mu(M-k), \\
& i = 1,\ldots,N, \quad j = 0,\ldots,N-i, \quad k = 0,\cdots,M-1 \\
q((i,j,k),a) &= \theta(i,j,k) = (i+j)\lambda + k\mu, \quad \forall\ (i,j,k) \in \mathcal{E}_{II} \\
q((i,j,k),(i,j,k)) &= -S(i,j,k), \quad \forall\ (i,j,k) \in \mathcal{E}_{II}\ .
\end{aligned}$$

The state space $\mathcal{E}_{II}$ and the generator matrix $\mathbf{Q_{II}}$, together with the mobile nodes meeting rate λ and mobile node-throwbox meeting rate μ, provide all the necessary information to iterate back the recursive equations until reaching the initial state $(1,0,0)$ of the chain, which provides the mean value of the various performance figures considered.

4.4.3 Validating the framework

We validated **Algorithm 1** by running numerical simulations of the scenario summarized in Table 4.2, and by comparing the results with those evaluated through the analytical framework. Tables 4.3-4.4 report the accuracy of **Algorithm 1** by comparing the values of various metrics computed by the algorithm for two network sizes, in particular for $N = M = 10$ and $N = M = 100$ nodes. For the sake of brevity, we present the results for strategies **I** and **II**, although similar conclusions apply for the other forwarding strategies.

As illustrated in both tables, the analytical results obtained through the algorithm match well with those reported by the simulations even when the number of nodes becomes large. The relative errors are mainly due to the approximations made on the meeting rates λ and μ.

Consequently in the following, we rely on the analytical results reported by **Algorithm 1** to evaluate numerically the performance of the various proposed strategies and discuss the routing trade-offs for each one of them in throwbox-embedded DTN networks.

TAB. 4.3 – Strategy **I** : Analytical (**Algorithm 1**) versus simulation results.

N	Metric	Alg. 1	Simul.	Err. %
10	T(1,0,0)	1.12 10^3	1.08 10^3	3.4
	RbyS(1,0,0)	0.0	0.0	0.0
	B(1,0,0)	1.98	1.78	11.32
	PrS(1,0,0)	0.39	0.38	2.37
	PrR(1,0,0)	0.35	0.35	3.2
100	T(1,0,0)	2.94 10^3	2.80 10^3	4.9
	RbyS(1,0,0)	0.0	0.0	0.0
	B(1,0,0)	5.70	5.29	7.71
	PrS(1,0,0)	0.10	0.09	3.08
	PrR(1,0,0)	0.71	0.72	1.98

TAB. 4.4 – Strategy **II** : Analytical (**Algorithm 1**) versus simulation results .

N	Metric	Alg. 1	Simul.	Err. %
10	T(1,0,0)	8.15 10^3	7.95 10^3	2.54
	RbyS(1,0,0)	3.21	3.27	1.75
	B(1,0,0)	1.48	1.37	7.83
	PrS(1,0,0)	0.27	0.28	4.21
	PrR(1,0,0)	0.55	0.56	1.15
100	T(1,0,0)	2.30 10^3	2.6 10^3	2.75
	RbyS(1,0,0)	8.35	8.46	1.29
	B(1,0,0)	4.34	4.61	5.81
	PrS(1,0,0)	0.07	0.08	6.48
	PrR(1,0,0)	0.79	0.78	3.0

4.4.4 Performance evaluation

Using the analytical framework, we investigate here the delay-energy performance of each strategy and compare its efficiency with the epidemic protocol and with the other proposed strategies.

Figure 4.2 plots the relative mean delivery delay $\mathbf{E}[T]$ (on the left) and the relative mean number of valuable transmissions $\mathbf{E}[G]$ (on the right) obtained under each strategy. The relative values are taken with respect to those corresponding to the epidemic protocol. Each point on the x-axis corresponds to a network made of an equal number of mobile nodes N and of throwboxes M taking values in the range $\{1, 2, \ldots, 100\}$. As it is possible to observe, given a specific setting, there is a significant difference among the performance of the various strategies. This highlights the space that is left to network designers when choosing the most suitable solution, given the application scenario being developed.

Figure 4.3 plots the mean number of valuable transmissions $\mathbf{E}[G]$ versus the mean delivery delay $\mathbf{E}[T]$ obtained under the epidemic protocol and the five strategies. Each (x,y) coordinate corresponds to a network made of an equal number of mobile nodes N and of throwboxes M taking values in the range $\{1, 2, \ldots, 100\}$. Commonly for the various

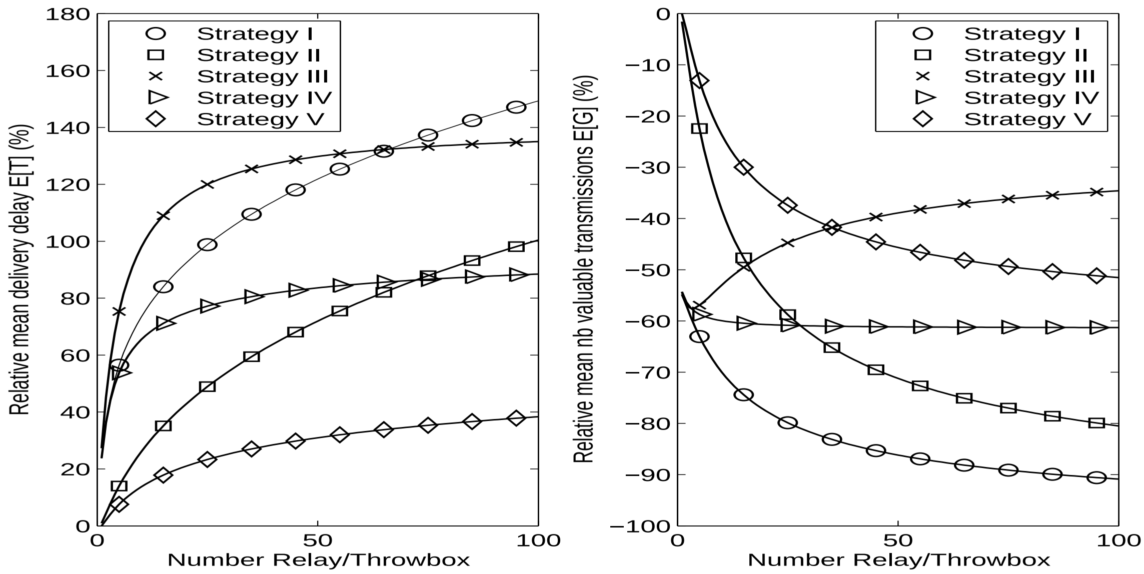

FIG. 4.2 – Relative mean delivery delay $\mathbf{E}[T]$ (left) and relative mean number of transmissions $\mathbf{E}[G]$ (right) obtained under each strategy with respect to epidemic protocol.

strategies, when node density is large ($N = M = 100$), the mean delivery delay $\mathbf{E}[T]$ reaches its lowest value while the mean number of valuable transmissions $\mathbf{E}[G]$ attains its highest value. When node density decreases, the mean delivery delay increases while the mean number of valuable transmissions decreases correspondingly. In comparison with our proposed strategies, the epidemic protocol achieves the lowest delivery delay at the expense of highest transmission overhead. The second best delivery delay performance is achieved by strategy **V**. The other strategies represent a trade-off between achieving low delivery delay and low transmission overhead.

4.5 Discussion and forwarding mechanisms design guidelines

The aim of this section is to provide some insights on the design and the evaluation of routing schemes for opportunistic disruption tolerant networks working in the presence of throwboxes. Through the various strategies we presented in Section 4.4 and which are simply diverse variants of permitted pairwise exchanges between the different nodes and throwboxes, the designer of applications oriented for opportunistic networks can have a relatively large spectrum of choices to trade-off delivery delay efficiency versus energy consumption.

Considering the previous setting, the epidemic protocol is shown to achieve the lowest mean delivery delay (see Figure 4.2). However, such metric is not always the most stringent one and needs to be accommodated with the overhead required for running the specific relaying scheme. For some application scenarios, it would be mode advantageous for instance to have 20% higher delivery delivery delay at the benefit of achieving 40% less transmission overhead (case of strategy **V** around $N = M = 50$). When comparing the performance of the

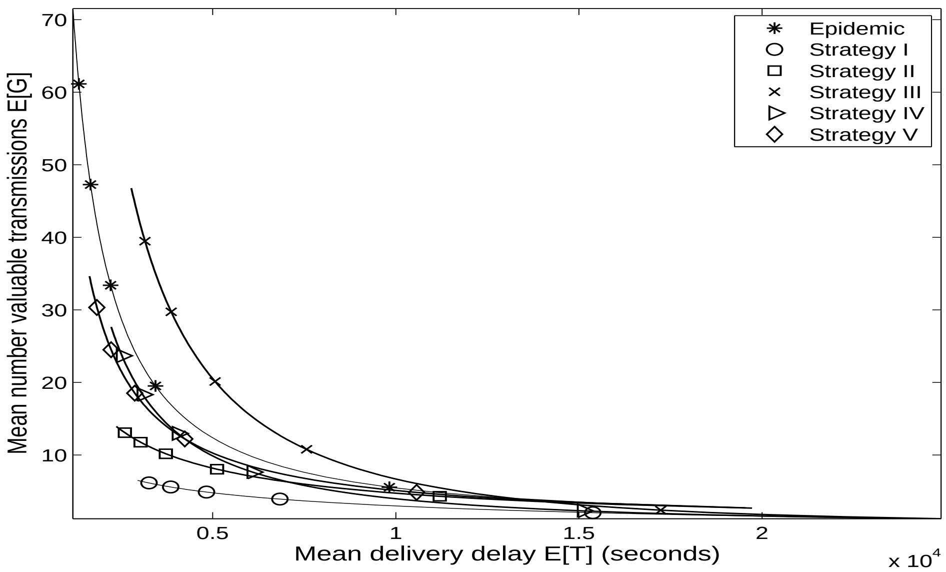

FIG. 4.3 – Mean number of valuable transmissions $\mathbf{E}[G]$ versus mean delivery delay $\mathbf{E}[T]$ obtained under the epidemic and the five routing strategies for decreasing node density.

five strategies to each others (Figures 4.2-4.3), the numerical results have shown that routing strategies which permit more copies exchange among the mobile nodes, for instance between the source and the mobile relays (e.g. strategies **II** and **V**) or between the mobile relays and the destination (e.g. strategies **I**, **II**, **IV** and **V**) perform better in terms of delivery delay. Observe that strategy **IV** which is simply strategy **III** extended with the possibility of an infected mobile relay to transmit a copy to the destination achieves roughly 50% less delivery delay. Similarly, strategy **V** which extends strategy **IV** with the possibility of the source to transmit copies to mobile relays achieves almost 50% less delivery delay than strategy **IV**. Similar observation can also be made regarding strategies **I** and **II**, where the latter strategy is simply extended with the possibility of the source to transmit a copy to a mobile relay.

The framework that we developed is shown to be a helpful tool to evaluate accurately the performance of these strategies. It is noteworthy to mention that the framework can be easily extended to study other strategies and performance metrics as long as the inter-meeting times among the nodes can be approximated by exponential distributions.

The framework is also useful to solve problems related to the choice of the optimal strategy to use in a given opportunistic scenario and/or the dimensioning of the number of throwboxes. For instance, let an opportunistic network that is constituted by N mobile nodes where packet routing is handled according to the classical epidemic protocol. Under such protocol, let $\mathbf{E}[T_{ep}]$ and $\mathbf{E}[G_{ep}]$ denote respectively the mean delivery delay and the mean number of copy transmissions made by the mobile nodes till the delivery of the packet. An optimization problem that a designer would encounter is to choose among the set S of possible strategies a routing strategy s with the following objective :

- Minimize the number of throwboxes $m \in \{1, \ldots, M\}$ to be added to the network, where M denotes their maximum number.

and subject to the following constraints :

– Achieve $\mathbf{E}[T_s]$ subject to $\mathbf{E}[T_s]$ - $\mathbf{E}[T_{ep}] \leq \alpha_T$ where $\alpha_T \geq 0$ is some threshold on the delivery delay.
– Achieve $\mathbf{E}[G_s]$ subject to $\mathbf{E}[G_s] \leq \alpha_G \leq \mathbf{E}[G_{ep}]$ where $\alpha_G \geq 0$ is some threshold on the energy consumption.

Algorithm 2 illustrates a heuristic algorithm implementing a simple procedure to solve the previous problem. For each considered strategy s and number of throwboxes m, the procedure solicits **Algorithm 1** to obtain the corresponding values of $\mathbf{E}[T_s^m]$ and $\mathbf{E}[G_s^m]$. At the end, it returns back the optimal strategy to use and the minimum number of throwboxes to deploy with respect to the considered constraints.

Algorithm 2 : A heuristic algorithm to solve for the minimum number of throwboxes and the optimal routing strategy to use.

Require: α_T, α_G and M
Ensure: Minimal m and optimal strategy $v \in \mathcal{E}_s = \{I, \ldots, S\} \backslash (\mathbf{E}[T_v^m] - \mathbf{E}[T_{ep}] \leq \alpha_T)$ and $(\mathbf{E}[G_v^m] \leq \mathbf{E}[G_s^m]) \ \forall s \in \mathcal{E}_s$
Define $min_m = M$
Define $optimal_s$
for $m = M$ to 1 **do**
  for $s = I$ to S **do**
    Compute $\mathbf{E}[T_s^m]$ and $\mathbf{E}[G_s^m]$ using **Algorithm 1**
    if $(\mathbf{E}[T_s^m] - \mathbf{E}[T_{ep}] \leq \alpha_T)$ and $(\mathbf{E}[G_s^m] \leq \alpha_G)$ **then**
      $min_m = m$
      $optimal_s = s$
    end if
  end for
end for
return min_m and $optimal_s$

Additionally, the proposed framework allows to better explore the performance of a specific relaying scheme. As an example, **Algorithm 1** can be used to create two-dimensional contour plots, highlighting the performance of a relaying mechanism for a variable number of mobile relays and throwboxes.

For instance, Figure 4.4 displays a filled two-dimensional contour plot of the mean delivery delay $\mathbf{E}[T_I]$ (on the left) and of the mean number of valuable transmissions $\mathbf{E}[G_I]$ (on the right) obtained under strategy **I**. The x-axis displays the number of mobile relays varying in the range $\{1, 2, \ldots, 100\}$ while the y-axis displays the number of throwboxes varying in the same range. Each (x, y) coordinate maps into a colored point representing a particular value of $\mathbf{E}[T_I]$ and $\mathbf{E}[G_I]$. Each contour line represents a curve of equal values, and two adjacent contour lines bound an interval of diverse values where those belonging to the same interval have the same color. The colorbars shown next to each figure display the range of values with their respective colors. The contour plot figure can be used as a design map to determine the number of throwboxes to be added to a given network in order to achieve given delivery delay and number of transmissions falling in certain bounds.

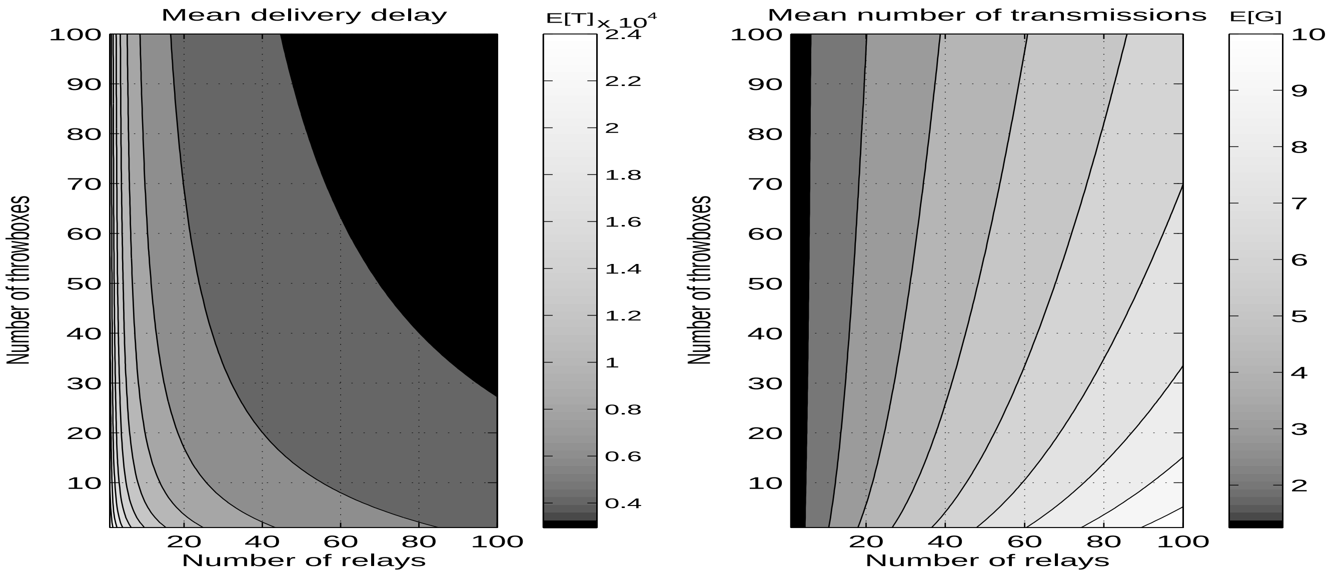

FIG. 4.4 – Contour plot of mean delivery delay $\mathbf{E}[T_I]$ and mean number of transmissions $\mathbf{E}[G_I]$ obtained under strategy **I**.

4.6 Conclusion

In this chapter, we presented a design methodology to engineer and evaluate routing techniques for opportunistic disruption tolerant networks. The proposed methodology is based on an analytical framework, which allows network designers to easily evaluate the performance of a wide range of technological solutions. In particular, it accounts for the peculiarities of the specific scenario (e.g., number of nodes/throwboxes, nodes mobility), different routing strategies and provides various performance figures such as delivery delay, or number of copies generated in the network by mobile nodes.

The proposed framework has been applied and validated in the case of throwbox-embedded opportunistic networks for several routing schemes. The results we obtain give insights on the various trade-offs that are left to network designers when deciding the most suitable solution for a given application scenario. The framework is general enough and can be easily extended to evaluate analytically other routing strategies and performance metrics as long as the inter-meeting times among the nodes can be approximated by exponential distributions.

4.A State space and transition rates of $\mathbf{Q_s}$

We recall the notations we used in **Algorithm 1** to denote the non-zero entries of the generator $\mathbf{Q_s}$.

$$\begin{aligned}
q((i,j,k),(i+1,j,k)) &= \alpha(i,j,k)\\
q((i,j,k),(i,j+1,k)) &= \beta(i,j,k)\\
q((i,j,k),(i,j,k+1)) &= \gamma(i,j,k)\\
q((i,j,k),a) &= \theta(i,j,k)\\
q((i,j,k),(i,j,k)) &= -S(i,j,k) = -(\alpha(i,j,k) +\\
&\quad \beta(i,j,k) + \gamma(i,j,k) + \theta(i,j,k)).
\end{aligned}$$

Strategy I : Under strategy **I**, $\mathcal{E}_I = \{(i,j,k), i \in (1), j \in (0,1,\ldots,N-1), k \in (1,\ldots,M) \cup (1,0,0)\}$. The non-zero entries of the generator $\mathbf{Q_I}$ are given by

$$\begin{aligned}
\alpha(i,j,k) &= 0, \quad \forall\ (i,j,k) \in \mathcal{E}_I\\
\beta(i,j,k) &= \mu k(N-i-j), \quad i=1, \quad j=0,1,\ldots,N-1, \quad k=1,\cdots,M\\
\gamma(i,j,k) &= \mu(M-k), \quad i=1, \quad j=0,1,\ldots,N-1, \quad k=0,\cdots,M-1\\
\theta(i,j,k) &= (i+j)\lambda + k\mu, \quad \forall\ (i,j,k) \in \mathcal{E}_I\\
-S(i,j,k), &\qquad \forall\ (i,j,k) \in \mathcal{E}_I
\end{aligned}$$

Strategy III : Similarly to strategy **I**, strategy **III** has the same state space $\mathcal{E}_{III} = \{(i,j,k), i \in (1), j \in (0,1,\ldots,N-1), k \in (1,\ldots,M) \cup (1,0,0)\}$. The non-zero entries of the generator $\mathbf{Q_{III}}$ are given by

$$\begin{aligned}
\alpha(i,j,k) &= 0, \quad \forall\ (i,j,k) \in \mathcal{E}_{III}\\
\beta(i,j,k) &= \mu k(N-i-j), \quad i=1, \quad j=0,1,\ldots,N-1, \quad k=1,\cdots,M\\
\gamma(i,j,k) &= \mu j(M-k), \quad i=1, \quad j=0,1,\ldots,N-1, \quad k=0,\cdots,M-1\\
\theta(i,j,k) &= i\lambda + k\mu, \quad \forall\ (i,j,k) \in \mathcal{E}_{III}\\
-S(i,j,k), &\qquad \forall\ (i,j,k) \in \mathcal{E}_{III}\ .
\end{aligned}$$

Strategy VI : Similarly to strategies **I** and **III**, strategy **IV** has the same state space $\mathcal{E}_{IV} = \{(i,j,k), i \in (1), j \in (0,1,\ldots,N-1), k \in (1,\ldots,M) \cup (1,0,0)\}$. The non-zero entries of the generator $\mathbf{Q_{IV}}$ are given by

$$\begin{aligned}
\alpha(i,j,k) &= 0, \quad \forall\ (i,j,k) \in \mathcal{E}_{IV}\\
\beta(i,j,k) &= \mu k(N-i-j), \quad i=1, \quad j=0,1,\ldots,N-1, \quad k=1,\cdots,M\\
\gamma(i,j,k) &= \mu j(M-k), \quad i=1, \quad j=0,1,\ldots,N-1, \quad k=0,\cdots,M-1\\
\theta(i,j,k) &= (i+j)\lambda + k\mu, \quad \forall\ (i,j,k) \in \mathcal{E}_{IV}\\
-S(i,j,k), &\qquad \forall\ (i,j,k) \in \mathcal{E}_{IV}\ .
\end{aligned}$$

Strategy V : Under strategy $\mathbf{V}$, $\mathcal{E}_V = \{(i,j,k), i \in (1,2,\ldots,N), j \in (0,1,\ldots,N-i), k \in (0,\ldots,M)\}$. The non-zero entries of the generator $\mathbf{Q_V}$ are given by

$$
\begin{aligned}
\alpha(i,j,k) &= \lambda(N-i-j), \quad i=1,\cdots,N-1, \quad j=0,\ldots,N-i, \quad k=0,\cdots,M \\
\beta(i,j,k) &= \mu k(N-i-j), \quad i=1,\ldots,N-1, \quad j=0,\ldots,N-i, \quad k=1,\cdots,M \\
\gamma(i,j,k) &= \mu(i+j)(M-k), \quad i=1,\ldots,N, \quad j=0,\ldots,N-i, \quad k=0,\cdots,M-1 \\
\theta(i,j,k) &= (i+j)\lambda + k\mu, \quad \forall\ (i,j,k) \in \mathcal{E}_V \\
-S(i,j,k), & \qquad \forall\ (i,j,k) \in \mathcal{E}_V\ .
\end{aligned}
$$

Chapitre 5

On the Diameter of Epidemic Routing Protocol

Contents

In this chapter, we examine the average length of forwarding path in opportunistic DTNs where packet relaying is carried out according to the epidemic protocol. We consider a network made of N nodes moving according to some random mobility model where meetings between nodes follow a Poisson process. With respect to that model, we derive a closed-form expression for the average number of hops that a copy of the initial packet makes before being delivered. The derived expression shows that the network diameter depends only on the number of nodes in the network and it is in the order of $\log(N)$. Using numerical results, we also investigate the impact of limiting the number of hops on the performance of epidemic dissemination protocols, in particular when the maximum number of hops approaches the network diameter.

5.1 Introduction and motivation

In this chapter, we analyze the mean length of the delivery path between a source and a destination in an opportunistic disruption tolerant network where the epidemic routing protocol is employed, and where nodes move independently of each other according to some random mobility model. Specifically, we define the *diameter of the epidemic protocol* as the mean number of hops that a copy of the source packet travels over time between the source and the destination when packet dissemination is performed according to the epidemic protocol (see also [31]).

The main motivation behind the present work is to characterize analytically the diameter of the epidemic protocol, and consequently, to investigate the impact of limiting the number of hops on the performance of an epidemic dissemination protocol. In particular, we are interested to look whether the diameter of the network corresponds to some optimal length for the forwarding path, and if some performance improvement can be achieved when the maximum number of travel hops is set equal to the diameter of the corresponding network.

In the following, using the theory of recursive trees, we provide a closed-form expression for the diameter of the epidemic protocol when nodes are moving according to a random mobility model for which the assumption of exponentially distributed inter-meeting times holds. The derived expression shows that the diameter varies as $\Theta(\log(N))$ where N is the total number of nodes within the network. Using the analytical framework that we developed in Section 4.3.2, we compare analytically the performance of different epidemic routing protocols that limit the maximum number of travel hops the packet can traverse to some given values.

In the literature, there are few works that have investigated the statistical characteristics of path delivery in MANETs or in opportunistic DTNs. Using simulation traces, the authors in [90] have addressed the statistical properties of path lifetime duration between pairs of nodes in multi-hop MANETs under different mobility models. For moderate and high nodes velocities, they have shown that the path lifetime duration follows approximately an exponential distribution for paths of two or more hops. They have also provided an empirical formula of the distribution parameter which depends on the relative speed of the mobility model, the node transmission range and the number of hops in the path. The exponential distribution of path duration in multi-hop MANETs have been also established in [46] using Palm theory under the assumption of large number of hops.

In a recent work [31], Chaintreau *et al.* have analyzed the characteristics of forwarding paths in opportunistic networks where wireless devices are carried by the participants of a conference. By assuming that the maximum delay and the number of hops of a forwarding path depends logarithmically on the number of relay nodes, they have derived conditions for the existence of such paths as well as their expected number between a given source-destination pair. Their results have been validated empirically on multiple human contact traces for sparse and dense networks composed of 40 up to 100 nodes and for different contact rates. They have observed that the network diameter generally varies from 3 to 6 hops in accordance with logarithmic increase assumption.

The rest of this chapter is organized as follows. In Section 5.2, we derive a closed-form expression for the diameter of the epidemic protocol. Section 5.3 validates through simulation the accuracy of the derived expression. Section 5.4 investigates through analytical results the impact of limiting the number of hops and Section 5.5 provides some concluding remarks.

5.2 Deriving the diameter expression

Hereafter, we are interested in computing the mean length of forwarding path for the epidemic protocol in opportunistic disruption tolerant networks. We consider a network that is made of $N+1$ mobile nodes, with one source node (S), one destination (D) and $N-1$ relay nodes. The nodes move according to some random mobility model, and inter-meeting times between two nodes are assumed to follow an exponential distribution. The source node has a packet to send to the destination and the remaining $N-1$ mobile nodes operate as relays to assist in copies spreading process according to the epidemic protocol.

To proceed with the analysis, we make an analogy between the temporal evolution of the infection process of the nodes under the epidemic protocol and the temporal evolution of a recursive tree of order $N+1$. A recursive tree of order n, RT_n, is defined as any tree in a connected graph with no cycles ([77, 34]). In addition, the vertices in a recursive tree are labeled in an increasing order of their joining instants to the tree. Specifically, a recursive tree RT_n of order n evolves from the recursive tree RT_{n-1} of order $n-1$ by choosing at random a vertex from RT_{n-1} and then by joining the vertex labeled n to it, where all $n-1$ vertices of RT_{n-1} are equally likely to be the potential parent. Conventionally, the vertex labeled 1 is defined as the root of the tree and the children are ordered from left to right within one level. Therefore, starting from the root, there is a unique path with increasing vertex labels that joins it to any random vertex of the tree. Observe that for the tree RT_n, there are $(n-1)!$ manners to label a given set of $(n-1)$ vertices, and consequently, $(n-1)!$ recursive trees.

The statistical properties of distances between labeled vertices in recursive trees have been extensively studied. Here the term *distance* denotes the number of edges separating two vertices. In [77], the author provided closed form expressions for the mean and the variance of distances in a recursive tree. The limiting distribution of path lengths when the number of vertices diverges is provided in [74], and the distance distribution between two random vertices in the tree has been examined in [34].

Mapping to our network model, a delivery path corresponds to a sequence of parent-child infected nodes that are obtained over time with the source node representing the root parent node. Further, the distance between two given nodes would denote the number of hops separating these two nodes.

Consequently, the mean length of the delivery path under our network model would map to the mean distance between the root of a recursive tree of order $N+1$ and a generic vertex v that joins randomly the tree among the remaining set of N vertices. The root maps to the source node in the network model while the random vertex v corresponds to the destination node. Let $X_{i,j}$ denote the distance in the recursive tree between two random vertices labeled i and j of a recursive tree of order n. For $1 \leq i < j \leq n$, the distribution of $X_{i,j}$ does not depend on n [34], and therefore we can consider henceforth $X_{i,n}$. As shown in [77], the expectation of $X_{i,n}$ is shown to be equal to :

$$E[X_{i,n}] = H_i + H_{n-1} - 2 + \frac{1}{i}$$

where $H_k = \sum_{j=1}^{k} j^{-1}$ is the k^{th} harmonic number. Mapping to our network model and conditioning on the possible infection "instants" of the destination node, the mean number of delivery hops between the source node S and the destination node D is equal to :

$$E[X_{SD}] = E[E[X_{SD}/D = n]] = \sum_{n=2}^{N+1} [H_1 + H_{n-1} - 2 + 1].\text{Prob(node D is labeled i)} \quad (5.1)$$

Observe that the probability of the destination node to join the tree after $i-1$ infections is shown in [41, Proposition 1] to depend only on the number of nodes in the network under the assumption of exponentially distributed inter-meeting times. Under the previous network model, this probability is equal to $\frac{1}{N}$. After some algebra calculus, Equation 5.1 translates into :

$$E[X_{SD}] = \frac{1}{N}\sum_{n=1}^{N}[H_n] = H_N - \frac{1}{N}\sum_{i=1}^{N-1}\frac{i}{i+1} \approx log[N] + \gamma - \frac{1}{N}\sum_{i=1}^{N-1}\frac{i}{i+1} \approx \Theta(log[N]) \quad (5.2)$$

where γ is the Euler constant which is equal to 0.5772. This formula implies that the mean number of hops needed to deliver a copy of the packet to the destination under the epidemic protocol is in the order of $\log[N]$ and it is independent of network parameters and contact rates among nodes.

5.3 Validating the diameter expression

To validate the expression of the diameter of the epidemic protocol provided in Equation 5.2 , we proceed through simulations. We consider a squared size area with a side length $L = 10^4\ m$. Nodes move inside the area according to the random waypoint model. In accordance with the previously introduced network model, we are interested in computing the mean number of hops that a packet travels between a source and a destination when the remaining nodes relay the copies according to the epidemic protocol.

Figure 5.1 plots the analytical and empirical values of the diameter versus the number of relay nodes taking values in the following set $\{10, 30, 50, 75, 100\}$. Empirical results refer to three network settings. The first setting (dashdot line) corresponds to a transmission range $r = 100\ m$ and a constant speed $v = 10\ m/s$. The second setting (dashed line) corresponds to $r = 200\ m$ and $v = 10\ m/s$, and the third setting (dotted line) corresponds to $r = 100\ m$ and $v = 20\ m/s$. As the figure illustrates, there is a good matching between Equation 5.2 and simulation results under different network configurations. The small relative errors observed for networks of 75 and 100 relays under a transmission range of 200 m are due to

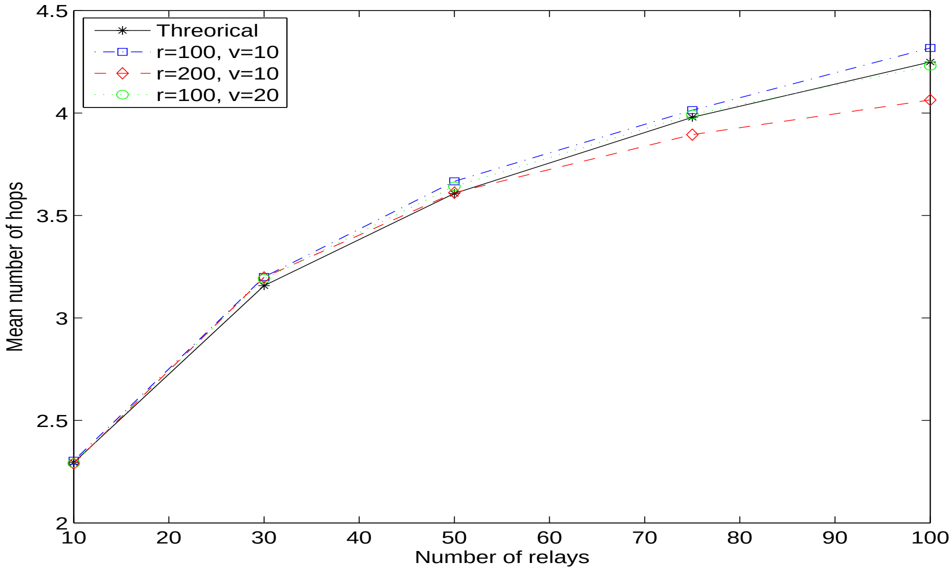

FIG. 5.1 – Network diameter : Analytical and simulated values for various network configurations.

the higher network connectivity so that the exponential assumption of inter-meeting times does no longer hold.

In accordance with Equation 5.2, observe also that the diameter depends only on the number of relay nodes within the network, and it is independent of the contact rate among the nodes. As a side remark, notice that the diameter of the epidemic routing is relatively small with respect to network size and varies between 2 to 4.5 when the relay nodes vary in number from 10 to 100 nodes. This observation has been also reported in [31] on real data traces of human mobility where the authors have as well showed that the length of forwarding paths between a pair of source destination varies around four hops - the number of devices considered in the experiment varied from 37 to 100 devices. Even though this observation implies that the packet needs to travel, on the average, a small number of hops to reach the destination using an epidemic dissemination protocol, it does not give any clue on the number of these paths originating from the source to reach the destination, nor on the number of copies that need to be spread in the network.

In the following section, we investigate the impact of limiting the maximum number of hops that copies can traverse on the performance of the epidemic protocol, in particular when the hops threshold approaches the network diameter.

5.4 Impact of limiting the number of hops on epidemic routing performance

The aim of this section is to examine the impact of limiting the number of hops that copies can traverse on the performance of the epidemic protocol. For instance, we are interested to check how the performance of the network would vary by setting the maximal number of hops under the epidemic protocol equal to the diameter of the epidemic protocol.

First, we consider the same network model as in Section 5.2. In addition, we assume here that each copy that is carried by a node is labeled with a tag indicating the number of hops that the copy had already made starting from the source till reaching the current node. In accordance with that, the source packet has a hop count of zero. In addition, whenever the node passes a copy to another relay node, the latter node increases by one the current hop count of the transferred copy. A node carrying a copy with a hop count of i will stop spreading more copies when the corresponding hop count of the copy is lower by one than the maximum permitted hop count in the network denoted by h_{Max} (i.e. stop spreading if $i = h_{Max} - 1$). The last hop is meant for potential copy delivery to the destination node.

A possible way to compare the performance for different h_{Max} values is through simulations ; however this approach is in some way time consuming since we need to run large number of simulations for each network size and diameter value. A simpler way would be to apply the analytical framework that we introduced in Section 4.3, and to compare numerically the performance metrics obtained under each h_{Max} value.

To apply the framework, we consider the following definition of the states for the associated Markov chain in order to handle the varying values of the maximum permitted number of hops h_{Max}. For a given h_{Max} value noted by H, where H is an integer value, let $\mathbf{A_H(t)} = (I_1(t), I_2(t), \ldots, I_{H-1}(t))$ denote the associated absorbing Markov chain having $H-1$ dimensions. For $n \in \{1, 2, \ldots, H-1\}$, the component $I_n(t) \in \{0, 1, 2, \ldots, N-n\}$ represents the number of infected mobile relays of the n^{th} generation at time t ahead of absorption. The first generation represents the mobile relays who are infected directly by the source. The second generation represents those relays who are infected by relays of the first generation and so on till the $(H-1)^{th}$ generation. Let $\mathcal{E}_{\mathbf{H}} = (\{0, 1, 2, \ldots, N-1\} \times \{0, 1, 2, \ldots, N-2\} \times \cdots \times \{0, 1, 2, \ldots, N-H\}) \cup \{a\}$ denote the state space of $\mathbf{A_H(t)}$.

In accordance with the states definition, the non-zero entries of the generator $\mathbf{Q_H}$ are given by

$$q((i_1, i_2, \ldots, i_{H-1}), (i_1+1, i_2, \ldots, i_{H-1})) = \alpha(i_1, i_2, \ldots, i_{H-1})$$
$$= \lambda(N-1-(i_1+i_2+\cdots+i_{H-1})),$$
$$(i_1, i_2, \ldots, i_{H-1}) \in (I_1(t) - \{N-1\} \times I_2(t) \times \ldots \times I_{H-1}(t))$$

$$q((i_1, i_2, \ldots, i_{H-1}), (i_1, i_2+1, \ldots, i_{H-1})) = \beta(i_1, i_2, \ldots, i_{H-1})$$
$$= \lambda i_1(N-1-(i_1+i_2+\cdots+i_{H-1})),$$
$$(i_1, i_2, \ldots, i_{H-1}) \in (I_1(t) \times I_2(t) - \{N-2\} \times \ldots \times I_{H-1}(t))$$
$$\vdots \qquad \vdots \qquad \vdots$$
$$q((i_1, i_2, \ldots, i_{H-1}), (i_1, i_2, \ldots, i_{H-1}+1)) = \eta(i_1, i_2, \ldots, i_{H-1})$$
$$= \lambda i_{H-2}(N-1-(i_1+i_2+\cdots+i_{H-1})),$$
$$(i_1, i_2, \ldots, i_{H-1}) \in (I_1(t) \times I_2(t) \times \ldots \times I_{H-1}(t) - \{N-H\})$$

$$q((i_1, i_2, \ldots, i_{H-1}), a) = \theta(i_1, i_2, \ldots, i_{H-1})$$
$$= \lambda(1+i_1+i_2+\cdots+i_{H-1}),$$
$$(i_1, i_2, \ldots, i_{H-1}) \in (I_1(t) \times I_2(t) \times \ldots \times I_{H-1}(t))$$

$$q((i_1, i_2, \ldots, i_{H-1}), (i_1, i_2, \ldots, i_{H-1})) =$$
$$-(\alpha(i_1, i_2, \ldots, i_{H-1}) + \beta(i_1, i_2, \ldots, i_{H-1}) + \cdots + \eta(i_1, i_2, \ldots, i_{H-1}) + \theta(i_1, i_2, \ldots, i_{H-1}))$$
$$(i_1, i_2, \ldots, i_{H-1}) \in (I_1(t) \times I_2(t) \times \ldots \times I_{H-1}(t))$$

Recall that our objective is to compute numerically the mean time that the packet or one of its copy takes till reaching the destination under a constrained maximum number of hops. In addition, we aim to compute the mean number of mobile relays who carry a copy of the packet when the packet is being delivered to the destination. We denote these two performance metrics respectively by $\mathbf{E}[T_H]$ and $\mathbf{E}[G_H]$.

To proceed with the computation of $\mathbf{E}[T_H]$ using the framework, we define the matrix $\mathbf{T_H}$ with entry $T_H(\{u\})$ representing the mean sojourn time of $\mathbf{A_H(t)}$ till the absorption starting from state $\{u\}$. Similarly, to compute $\mathbf{E}[G_H]$, we define the matrix $\mathbf{G_H}$ with entry $G_H(\{u\})$ representing the mean number of infected mobile relays, source included, starting from state $\{u\}$ till the absorption time. Matrices $\mathbf{T_H}$ and $\mathbf{G_H}$ are of equal size, and their size, noted as Dim_H, is defined as

$$\begin{aligned} Dim_H &= \mid I_1(t) \mid \times \mid I_2(t) \mid \times \cdots \mid I_{H-1}(t) \mid \\ &= N \times (N-1) \times \ldots (N-n+1) \times \ldots (N-H+2) \end{aligned} \tag{5.3}$$

Once the state space $\mathcal{E}_\mathbf{H}$, the generator $\mathbf{Q_H}$ and matrices $\mathbf{T_H}$ and $\mathbf{G_H}$ are well defined, we proceed by applying the analytical framework to compute $\mathbf{E}[T_H]$ and $\mathbf{E}[G_H]$ under different H values. As an example, **Algorithm 3** illustrates an implementation of the framework on the defined $\mathbf{A_H(t)}$ for the case when the maximum permitted number of hops

H is set equal to 4. Observe here that the associated $\mathbf{A_4(t)}$ has a dimension of three. The only slight difference with **Algorithm 1** is that here we need to take into account the issue that $I_{n+1}(t)$ cannot take non-zero values unless that $I_n(t) > 0$. This is in accordance with our definition of $\mathbf{A_H(t)}$ where infected relay nodes of generation $n+1$ cannot exist unless there is at least one infected relay node belonging to generation n.

Algorithm 3 : Computes $\mathbf{E}[T_H]$ and $\mathbf{E}[G_H]$ when $h_{Max} = 4$

Require: N and λ
Ensure: $\mathbf{E}[T_4]$ and $\mathbf{E}[G_4]$
for $k = N-3$ to 0 **do**
  for $j = N-2-k$ to 1 **do**
    for $i = N-1-k-j$ to 1 **do**
      $\alpha(i,j,k) = \lambda(N-1-i-j-k)$
      $\beta(i,j,k) = \lambda i(N-1-i-j-k)$
      $\gamma(i,j,k) = \lambda j(N-1-i-j-k)$
      $\theta(i,j,k) = (1+i+j+k)\lambda$
      $S(i,j,k) = \alpha(i,j,k) + \beta(i,j,k) + \gamma(i,j,k) + \theta(i,j,k)$
      if $(i+j+k) = (N-1)$ **then**
        $T_4(i,j,k) = \frac{1}{S}$
        $G_4(i,j,k) = N$
      else
        $T_4(i,j,k) = \frac{1}{S} + \frac{\alpha}{S}T_4(i+1,j,k) + \frac{\beta}{S}T_4(i,j+1,k) + \frac{\gamma}{S}T_4(i,j,k+1)$
        $G_4(i,j,k) = (1+i+j+k)\frac{\theta}{S} + \frac{\alpha}{S}G_4(i+1,j,k) + \frac{\beta}{S}G_4(i,j+1,k) + \frac{\gamma}{S}G_4(i,j,k+1)$
      end if
    end for
  end for
end for
for $i = N-1$ to 1 **do**
  $\alpha = \lambda(N-1-i)$
  $\beta = \lambda i(N-1-i)$
  $\theta = \lambda(i+1)$
  $S = \alpha + \beta + \theta$
  if $i = (N-1)$ **then**
    $T_4(i,0,0) = \frac{1}{S}$
    $G_4(i,0,0) = N$
  else
    $T_4(i,0,0) = \frac{1}{S} + \frac{\alpha}{S}T_4(i+1,0,0) + \frac{\beta}{S}T_4(i,1,0)$
    $G_4(i,0,0) = (1+i)\frac{\theta}{S} + \frac{\alpha}{S}G_4(i+1,0,0) + \frac{\beta}{S}G_4(i,1,0)$
  end if
end for
$\mathbf{E}[T_4] = \frac{1}{\lambda N} + \frac{\lambda(N-1)}{\lambda N}T_4(1,0,0)$
$\mathbf{E}[G_4] = \frac{\lambda}{\lambda N} + \frac{\lambda(N-1)}{\lambda N}G_4(1,0,0)$
return $\mathbf{E}[T_4]$ and $\mathbf{E}[G_4]$

Table 5.2 reports the accuracy of **Algorithm 3** for two network sizes ($N = 51$ and

$N = 101$). It compares the analytical and the simulated values of $\mathbf{E}[T_4]$ and $\mathbf{E}[G_4]$ obtained under the scenario summarized in Table 5.1 when h_{Max} is set equal to 4. The analytical results have been computed using **Algorithm 3**.

TAB. 5.1 – Summary of the scenario under consideration.

Mobility model	RWP
Playground	$10 \times 10\ km^2$
Nodes speed	$v = 10$ m/s
Nodes comm. range	$r = 100$ m
Mobile/mobile meeting rate	$\lambda = \frac{8rv\omega}{\pi L^2}$

TAB. 5.2 – h_{Max} H = 4 : Analytical (**Algorithm 3**) versus simulation results.

N	Metric	Computed	Measured	Relative Error %
51	$\mathbf{E}[T_4]$	$2.75\ 10^3$	$2.59\ 10^3$	6.27
	$\mathbf{E}[G_4]$	21.51	23.83	-9.75
101	$\mathbf{E}[T_4]$	$1.68\ 10^3$	$1.58\ 10^3$	6.4
	$\mathbf{E}[G_4]$	38.05	44.21	-13.92

In the following, we apply the analytical framework of Section 4.3.2 on the present Markov chain $\mathbf{A_H(t)}$ in order to evaluate $\mathbf{E}[T_H]$ and $\mathbf{E}[G_H]$ for different values of the maximum permitted travel hop number H. Figure 5.2 displays the theoretical network diameter when the number of relays varies in the range $[5, \ldots, 75]$. The analytical values of network diameter have been obtained using Equation 5.2. With respect to an equivalent network operating under the epidemic protocol, figure 5.3 displays the relative mean delivery delay and number of copy transmissions versus the number of relays for different values of h_{Max} threshold $H \in \{2, 3, 4, 5\}$.

As the results illustrate, there is no clear evidence on performance impact when the h_{Max} threshold approaches the diameter corresponding to the network in hand. However, limiting the maximum number of traveled hops between the source and the destination results in a trade-off associated with lowering the energy consumption at the nodes at the expense of increasing the delivery latency. The latter observation is of interest for some applications which would tolerate relative latency increase at the advantages of a decrease of routing overhead.

5.5 Conclusion

In this chapter, we have studied the mean number of hops of a delivery path in an opportunistic DTN operating under the epidemic protocol. We have considered a network model made of $N+1$ nodes moving according to some random mobility model where meetings

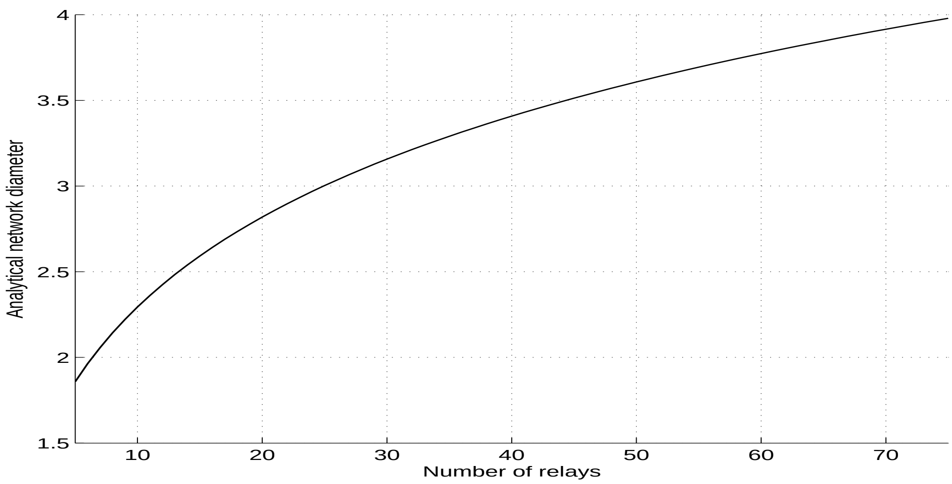

FIG. 5.2 – Analytical network diameter plotted versus the number of relays.

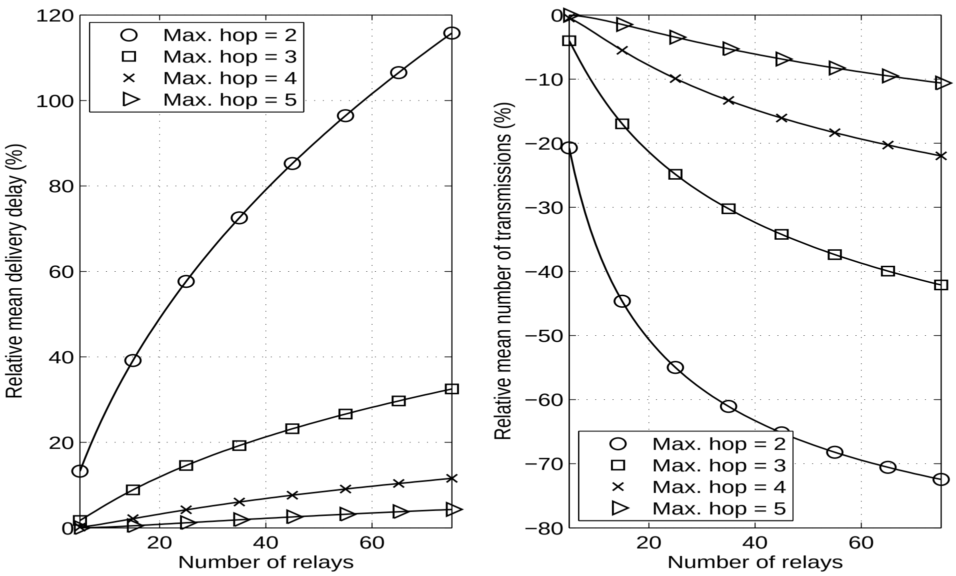

FIG. 5.3 – Relative mean delivery delay (on the left) and number of copy transmissions (on the right) with respect to epidemic protocol versus the number of relays.

between nodes follow a Poisson process. Using the theory of recursive trees, we have derived a closed-form expression for the average length of forwarding path between a source and a destination under such scenario. The latter expression is in the order of $\log(N)$ and it is independent of mobility model, contact rate and network parameters. By applying the analytical framework that we presented in Chapter 4, we have compared numerically the impact of limiting the number of hops on epidemic dissemination protocol, in particular when the number of hops approaches the network diameter. The obtained results showed that limiting the maximum number of hops can be useful for some applications to trade-off an increase in the delivery delay at the advantage of a decrease in power overhead in the network.

Design and analysis of an 802.11 MAC algorithm

Chapitre 6

Adaptive Backoff Algorithm for IEEE 802.11

Contents

This chapter presents an adaptive backoff algorithm for the contention-based Distributed Coordination Function (DCF) of the IEEE 802.11 standard. Our design objectives are to maximize the system throughput while minimizing the packet delay at the transport level. The proposed algorithm, referred to as "adaptive BEB" judiciously sets the size of the minimal contention window to adapt to the level of congestion in the wireless channel. Unlike previous proposals found in the literature, our algorithm relies on on-line measurements of the number of active sources. These measurements are filtered through a simple filter and used to determine the optimal size of the minimal contention window. The performance of our algorithm is evaluated through extensive simulations in `ns-2`. Our algorithm performs remarkably well in all investigated situations. It outperforms the standard in both ad hoc and infrastructure mode of operation, whether the number of active sources is constant, changes smoothly or abruptly.

Note : The material presented in this chapter has appeared in [73].

6.1 Introduction

The IEEE 802.11 standard [2] defines Medium Access Protocol (MAC) and Physical (PHY) frameworks to provide wireless connectivity in Wireless Local Area Network (WLAN). However, since its initial appearance in 1997, the deployment of this standard have engendered several research issues. One of the well-known drawbacks of IEEE 802.11 is the low performance of its MAC protocol in terms of throughput and access delay in congested networks. Consequently, the design of algorithms that maximize these performances has been a challenge since quite some time, and several research studies have proposed solutions to better react to the congestion in the network [15, 26, 67, 83, 4, 36, 104]. However, and as it will be shown latter, the performance of these algorithms in networks with varying topologies, in noisy environments or in the case of nodes with limited buffer size is not satisfactory.

In this chapter, we propose and evaluate through simulations an adaptive backoff algorithm that enhances the performance of the Distributed Coordination Function (DCF) mechanism of the IEEE 802.11 standard under a large variety of conditions. This new algorithm is meant to be simple and as close as possible to the standard. Our objectives are to maximize the system throughput while minimizing the packet delay at the transport level. Considering saturation conditions, i.e. active stations having always packets to transmit, we derive a simple expression relating the minimum contention window size of the DCF backoff to the optimal transmission probability. This expression is at the basis of our algorithm and allow us to dynamically adapt the minimum contention window size after a successful transmission to the changing network conditions. The latter probability has been computed in the past in [14] and an alternative formulation can be derived from [26]. Both formulations require an estimate of the number of contending stations. To carry the estimation, we propose a novel method that relies on counting *signs of life* coming from the other transmitting stations to estimate their number.

Our algorithm has been implemented in the Network Simulator `ns-2` [75] as well as two other algorithms proposed in the literature [83, 4, 36]. We have run multiple simulations with different scenarios in order to evaluate the performance of our algorithm and compare it to the IEEE 802.11 standard and to the two other proposed algorithms. Our proposal outperformed the standard in all the considered scenarios. It outperformed the other adaptive algorithms in the cases of abrupt network changes, infrastructure mode of operation or in the presence of noisy channels.

With respect to the other proposals of adaptive backoff algorithms, our proposed approach presents two main features :

- It is an efficient backoff algorithm for IEEE 802.11 that adaptively adjusts the minimum contention window of the wireless stations after a successful transmission. The choice of the minimum contention window size takes into account the number of active stations in the vicinity of the transmitting station.
- It provides a novel and adequately accurate method to estimate the number of active stations. The method consists in measuring their number over the air then tracking their evolution across time using a weighted filter.

The chapter is organized as follows : Section 6.2 surveys some recent works related to the design of adaptive algorithms for the 802.11 DCF mechanism. The analytical background needed to design our algorithm is presented in Section 6.3. Section 6.4 is devoted to the algorithm itself, motivating every step of the algorithm and describing its design and the implementation. A simple estimator of the number of contending nodes is presented in Section 6.5, and in Section 6.6, the effectiveness of our proposal is analyzed and compared to other approaches through extensive simulations. Section 6.7 studies the fairness of our algorithm, and last Section 6.8 summarizes our results.

6.2 Related work

Several research work have proposed modifications to the backoff algorithm of the IEEE 802.11 standard in order to improve its performance in the case of congested networks. Other studies propose innovative distributed contention-based MAC protocols like the Fast Collision Resolution (FCR) algorithm ([67]) and the General Contention window Adaptation (GCA) algorithm ([104]) as a solution for dynamic MAC algorithms in wireless LANs. In this section, we will present and briefly discuss several algorithms which propose to adapt the backoff window for the 802.11 DCF mechanism. For details about the 802.11 DCF mechanism, the interested readers are referred to Appendix 6.A.

The earliest of these algorithms has been developed by Bianchi *et al.* [15]. The idea of the algorithm is simple and based on a unique contention window CW, whose size is updated after each transmission attempt based on an estimation of the number of active stations. CW represents the optimal contention window size that maximizes the system throughput, and is found to be equal to $\bar{N}\sqrt{2T}$ under the assumption that all stations use the same contention window size. In this expression, T is the total packet transmission time and $\bar{N}$ is an estimation of the current number of contending stations. In order to estimate N, the authors compute $\mathbf{E}[c(B)]$ which is the average number of busy slots observed during the backoff decrease period denoted by B. The current number of contending stations is then inferred as follows

$$\hat{N} = 1 + \frac{\mathbf{E}[c(B)](CW + 1)}{2B}$$

and the estimation of N is achieved by using the following filter

$$\bar{N}(n) = \alpha\bar{N}(n-1) + \frac{(1-\alpha)}{q}\sum_{i=1}^{q}\hat{N}_{n-(i-1)}$$

The parameters α and q are set heuristically to 0.8 and 10, respectively. However, there are several problems inherent to this algorithm. First, the inferred $\hat{N}$ depends on the old value of CW, which itself depends on the estimated $\bar{N}$. Therefore, little deviations in the value of N propagate from one estimate to the following one, making the estimation unstable. To palliate this problem, the authors of [15] introduce another function $s(N) = (1 + h/\sqrt{N})$ using a parameter h set heuristically to 2. The optimal contention window is thus $s(\bar{N})\bar{N}\sqrt{2T}$. Another problem that might arise relates to fairness among contending stations. Clearly, stations having lower CW values, will receive higher bandwidth. Since CW uses an estimation of N which might differ from one station to another, a fairness issue might arise in this algorithm.

Another algorithm based on a unique contention window size is given in [25, 26]. The authors establish an analogy between the p-persistent IEEE 802.11 protocol and the standard protocol. In a p-persistent 802.11 protocol, the backoff interval is sampled geometrically with parameter p. Under the condition of equal average backoff time, the authors infer the average backoff interval that should be adopted in a given network configuration to optimize the throughput of the system. It comes then that $p = 1/(\mathbf{E}[B] + 1)$ where $\mathbf{E}[B]$ is the average size of the backoff interval. By developing an analytical model for the p-persistent 802.11 protocol, the authors derive an expression of the throughput in terms of p and N, the number of contending stations, and compute the optimal value of p that maximizes the system throughput. The optimal p yields the optimal average backoff time in the p-persistent 802.11 protocol. Due to the analogy assumption, the authors propose to apply the obtained results to the standard protocol by using a single backoff window with the size set to the previously obtained optimal average backoff time. Observe that the computation of the optimal value of p requires the estimation of the number of current active users. By exploiting their analytical model, the authors derive the following equation

$$\hat{N} = \frac{1 - p}{p.Idle(p)}$$

which infers the number of users in terms of the previous persistence probability p and the average number of idle slots in a virtual transmission time[1] denoted as $Idle(p)$. As in the previous algorithm, the optimal persistence probability is computed based on $\hat{N}$, which itself depends on p. Thus, this algorithm too has an instability problem. Furthermore, $Idle(p)$ cannot be measured accurately also as a given station cannot know if transmissions to other stations were successful or not. Thus, a further error is introduced in the estimation of N. Last, we expect this algorithm to have the same fairness problem as the previous one.

Another class of algorithms suggest to slowly decrease the contention window size instead of resetting it after a successful transmission. This allows to keep some sort of knowledge on the current congestion level in the network by keeping the contention window size close to the previous one, showed to be convenient for the current congestion level. However, there are some issues related to the slow decrease schemes that will be discussed later on throughout the chapter.

The earliest in this class of algorithms is [83]. It proposes to multiplicatively decrease the contention window after a successful transmission, namely $CW_{\text{new}} = \max(CW_{\min}, D\, CW_{\text{old}})$ where D is a decrease factor chosen in the range $[0, 1]$ that is set constant for all stations in the network. The choice of D greatly influences the performance of the algorithm and constitutes a trade-off between avoiding packet collisions by keeping the value of CW closer to the previous one and risking high access delay where the previous contention window might be unnecessarily large. Specified initially as a power of 2 to meet industrial requirements and to simplify mathematical analysis, the decrease factor D was set equal to $1/2^g$

[1]The virtual transmission time is the time that elapses between the end of two consecutive successful transmissions over the channel (see Section 6.3.1 for further details).

where g is a positive integer number. The authors found later in [4] that maximum system throughput is achieved with a value of $D = 0.9$ at the cost of a high system response time.

In [36], the author proposes to linearly increase, respectively linearly decrease, the contention window size after a successful, respectively an unsuccessful, transmission. More precisely, upon a missed ACK, the current contention window is increased by a constant value ω. Upon receiving an ACK, CW is decreased by ω with probability $1 - \delta$, and with probability δ the contention window is not changed. In the simulation analysis provided in [36], ω and δ are set heuristically to 32 and 0.4943 respectively.

Through the simulation results reported in [4],[36], the last two algorithms showed good performance for fixed or even slowly varying number of users. However, and as we will show latter, their performance degrades greatly over noisy channels or in presence of abrupt changes in the number of active users. This is due to the fact that the control of the contention window in these approaches is done automatically based only on information such as successful or unsuccessful transmissions, regardless of the cause of failure (corruption or collision). In addition, the increase and decrease factors are set heuristically to constant values that are independent of the network state.

To conclude, we would like to briefly mention that [16] proposes a Kalman filter to estimate the number of active users N and relies on channel sensing to measure the collision probability.

6.3 Analytical background

In this section, we build on the analytical context which led to the design of our algorithm. The main results presented in this section have appeared in [14] for the IEEE 802.11 DCF function. In the following, we will first present some already derived performance measures such as the transmission probability of a node, the normalized system throughput and the optimal transmission probability. Last, we will provide some insights on the optimality of the transmission probability.

Our starting point is based on the Markov model that has been proposed by Bianchi in [14] to model the behavior of the binary exponential backoff corresponding to a packet transmission. We recall briefly here the basic assumptions of the considered network model as well as the main performance measures derived in that model.

In [14], it is assumed that the network is constituted by N fixed nodes, all located in the same transmission range of each others (the nodes compose a single cell). Each node is a source for a traffic connection and it is a destination for another different connection; every source node has always packets ready to send, i.e. the network works under saturation conditions. Channel conditions are assumed to be ideal; hidden node and channel capture problems are disregarded. Furthermore, the author assume that the collision probability is constant among all sources and independent of the past. Last, it is assumed that there is no limit on the number of retransmissions of a lost packet.

Based on this network model, [14] introduces a two-dimensional discrete-time Markov chain to model the evolution of the BEB algorithm corresponding to a packet transmission of a given node. Using the previous Markovian model, the author derives the following expression of the stationary probability τ that a node transmits a packet in a randomly chosen slot time :

$$\tau = \frac{2(1-2p)}{(1-2p)(CW_0+1) + p\,CW_0(1-(2p)^m)} \tag{6.1}$$

where $m = \log_2 \frac{CW_{max}}{CW_0}$ and p denotes the collision probability seen by the transmitting source node. Recall from Appendix 6.A that CW_0 (resp. CW_{max}) denotes the minimum (resp. the maximum) size of the contention window. A collision occurs when there is at Least two simultaneous transmissions on the channel, and hence p writes as

$$p = 1 - (1-\tau)^{N-1} \tag{6.2}$$

By studying the events that can occur within a generic slot time, the author provided in addition an expression of the normalized saturation throughput S of the network which is given by

$$S = \frac{P_s \mathbf{E}[\text{payload}]}{P_s T_s + P_c T_c + P_I \sigma} \tag{6.3}$$

and where the various parameters are summarized in Table 6.1.

TAB. 6.1 – Normalized saturation throughput parameters.

Parameter	**Significance**
S	Normalized throughput of the network
$\mathbf{E}$[payload]	Expected time to successfully transmit the payload of a frame
T_s	Expected time taken by a successful transmission
T_c	Expected time taken by an unsuccessful transmission
σ	Slot time length
P_s	Probability of a successful transmission $= N\tau(1-\tau)^{N-1}$
P_I	Probability of an idle slot time $= (1-\tau)^N$
P_c	Probability of a collision $= 1 - P_I - P_s$

Observe that for a given N, τ and p can be computed using a fixed-point approach. Consequently, all the terms of Equation 6.3 are known, and thus, the saturation throughput can then be rewritten as

$$S = \frac{\mathbf{E}[\text{payload}]}{T_s - T_c + \left(\frac{T_c/\sigma - (1-\tau)^N (T_c/\sigma - 1)}{N\tau(1-\tau)^{N-1}}\right)\sigma} \tag{6.4}$$

According to Equation 6.4, maximizing the previous expression of the system throughput is achieved at

$$\tau^* = \frac{\sqrt{(N + 2(N-1)(T_c/\sigma - 1))\,/N} - 1}{(N-1)(T_c/\sigma - 1)} \tag{6.5}$$

which is referred to as the optimal transmission probability, and which under the condition $\tau \ll 1$, can be approximated by

$$\tau^* \approx \frac{1}{N\sqrt{T_c/(2\sigma)}} \tag{6.6}$$

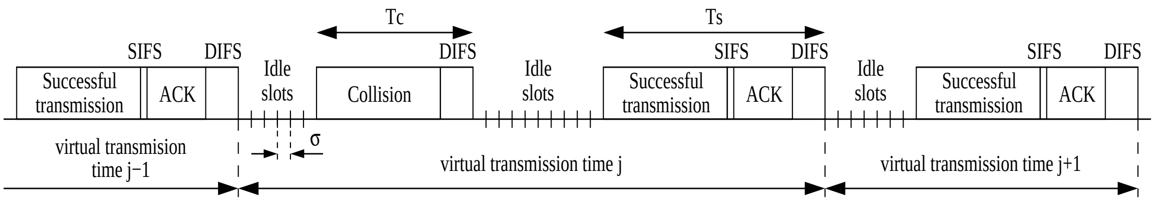

FIG. 6.1 – An illustration of the virtual transmission time.

The latter expression of the optimal transmission probability is of great importance and it is at the basis of our algorithm in the next. But before proceeding with the algorithm, we will provide some insights on the optimality of this transmission probability.

6.3.1 Insights on the optimality of the transmission probability

The optimal transmission probability found in [14] and reported in Equation 6.6 will not only maximize the saturation throughput but will minimize as well the system response time. In other words, the idle time wasted over the wireless channel will be minimal. In this section, we will provide some insights on the optimality of the transmission probability.

The time over the wireless channel can be partitioned into successive *virtual transmission times*. A virtual transmission time initiates just after a successful transmission over the wireless channel, and ends at the end of the next successful transmission as illustrated in Figure 6.1. The wasted time in a virtual transmission time, denoted as $waste_\tau$, is due to collisions and idle slot times and can be written

$$waste_\tau = \mathbf{E}\left[N_c T_c + (N_c+1)\text{Idleness}\right] = \mathbf{E}[N_c]T_c + (\mathbf{E}[N_c]+1)\mathbf{E}[\text{Idleness}] \qquad (6.7)$$

where $\mathbf{E}[N_c]$ and $\mathbf{E}[\text{Idleness}]$ respectively represent the average number of collisions and the average length of an idle period in a virtual transmission time (N_c and Idleness are statistically independent). In (6.7) it is assumed that all packets have the same size, hence T_c is a constant. The first and second terms of (6.7) respectively account for the total time wasted in collisions and the total idle time in a virtual transmission time. Observe that the virtual transmission time is simply $waste_\tau + T_s$, T_s being the expected time taken by a successful transmission, so the normalized system throughput can be expressed as the ratio $\mathbf{E}[\text{payload}]/(waste_\tau + T_s)$.

To derive the expressions for $\mathbf{E}[N_c]$ and $\mathbf{E}[\text{Idleness}]$, we look at the distribution in a virtual transmission time of the number of collisions N_c and the idle period length :

$$P[N_c = j] = \left(\frac{P_c}{P_c+P_s}\right)^j \frac{P_s}{P_c+P_s} \qquad \text{for } j \in \mathbb{N}$$
$$P[\text{idle\_period} = j\sigma] = P_I^j(1-P_I) \qquad \text{for } j \in \mathbb{N}$$

where σ is the slot time duration. Therefore, we can write

$$\mathbf{E}[N_c] = \sum_{j=0}^{\infty} j \left(\frac{P_c}{P_c + P_s} \right)^j \frac{P_s}{P_c + P_s} = \frac{P_c}{P_s} \quad (6.8)$$

$$\mathbf{E}[\text{idle\_period}] = \sigma \sum_{j=0}^{\infty} j P_I^j (1 - P_I) = \frac{\sigma P_I (1 - P_I)}{(1 - P_I)^2}$$
$$= \frac{\sigma P_I}{1 - P_I} = \frac{\sigma (1-\tau)^N}{1 - (1-\tau)^N} \quad (6.9)$$

After some calculus, it comes that the average time wasted in collisions during a virtual transmission time is given by

$$\mathbf{E}[N_c] T_c = T_c \left(\frac{1 - (1-\tau)^N}{N\tau(1-\tau)^{N-1}} - 1 \right) \quad (6.10)$$

while the corresponding average idle time reads as

$$(\mathbf{E}[N_c] + 1)\mathbf{E}[\text{Idleness}] = \frac{\sigma (1-\tau)^N}{N\tau(1-\tau)^{N-1}} \quad (6.11)$$

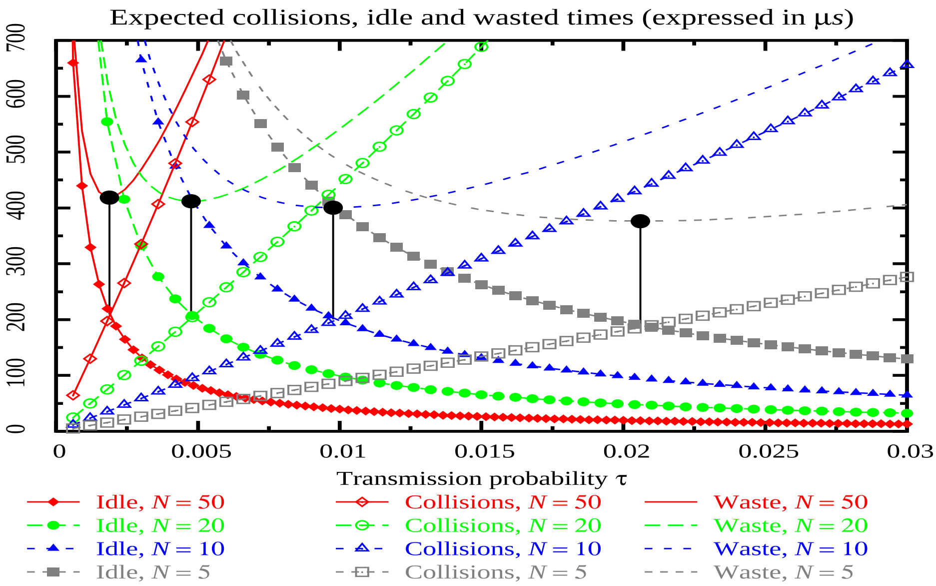

FIG. 6.2 – Expected wasted time, expected time spent in collisions and expected idle time in a virtual transmission time for different values of N ($T_c = 4335\mu s$ and $\sigma = 20\mu s$).

It can easily be proved that $\mathbf{E}[N_c]T_c$ is a monotone increasing function of τ, whereas $(\mathbf{E}[N_c] + 1)\mathbf{E}[\text{Idleness}]$ is a monotone decreasing function of τ. In Figure 6.2, both terms are plotted against τ for different values of the number of contending stations N. Their sum, $waste_\tau$, is also plotted and its minimal value for each N has been marked. The minimal values of $waste_\tau$ naturally correspond to the optimal transmission probability for each N. As seen in Figure 6.2, the minimal values of $waste_\tau$ correspond to the intersections of both collisions and idle times. In other words, the optimal use of the channel is achieved when the wasted time is equally shared between collisions and idleness. This optimal operating point has been identified in [9, p. 243] for the CSMA slotted Aloha protocol and in [27] for the p-persistent IEEE 802.11 protocol.

Observe also that for different values of the number of active stations N, the values on the y-axis of the intersection points are very close to each others, implying that the maximum saturation throughput of the system does not depend much on the number of active stations. It can be achieved in different network topologies by well tuning the transmission probability of the stations to the optimal transmission probability for each network topology. Furthermore, it is note worthy to mention here that having the system response time minimized at the optimal transmission probability does not guarantee that the packet delay is minimized as well. This minimization also relies on the protocol fairness and its ability to equally share the medium among contending sources.

Last, Figure 6.3 depicts the expected number of collisions $\mathbf{E}[N_c]$ as a function of the number of active stations N. The three plotted curves correspond to the standard DCF, the DCF tuned with the optimal transmission probability and its approximation. Clearly, the number of collisions decreases dramatically when τ is optimal and it becomes insensitive to N. The curves corresponding to the DCF tuned with the optimal τ given in Equation 6.5 and the DCF tuned with the approximate optimal τ given in Equation 6.6 are indistinguishable. Hence, the simple formulation of Equation 6.6 will be used in the following as an accurate approximation of the optimal transmission probability.

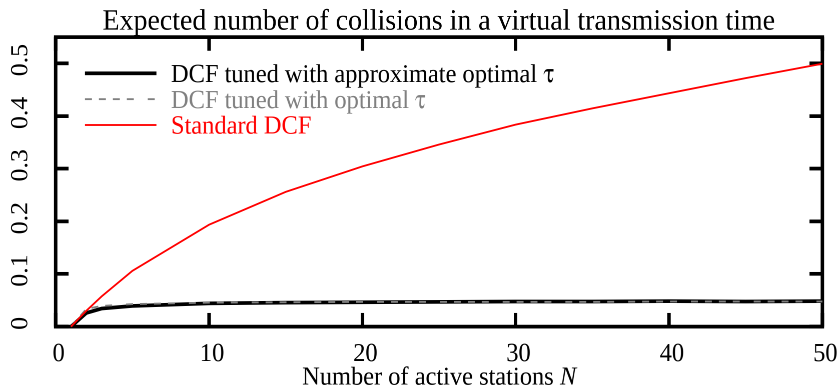

FIG. 6.3 – Analytical expected number of collisions in a virtual transmission time versus the number of stations.

6.4 Adaptive Binary Exponential Backoff

The binary exponential backoff algorithm used in DCF has two main drawbacks. First, the contention window of a transmitting station increases automatically when this station does not receive an acknowledgement of its transmited frame within the *ACK timeout*. Even that an unacknowledged frame could be potentially the result of a collision, it could also due to other factors such as frame corruption resulting from noisy channels or even from a buffer overflow at the destination.

The second shortcoming results from the fact that a transmitting station that observes a successful transmission of its packet will reset directly the current value of its contention window CW to its default initial value CW_0[2]. While starting with CW_0 would be convenient in networks with small number of transmitting nodes, it is highly inefficient in congested networks. In particular, in congested networks, a "convenient" contention window size of a *successfully transmitted frame* is obtained at the cost of several collisions when CW starts from CW_0. The last attained value by CW is a consequence of the current congestion level in the network, and therefore, by resetting CW to its default value CW_0, the transmitting station will forget its knowledge about the current congestion level in the network. Consequently, the next frames will suffer again the same average number of collisions until reaching the appropriate CW, wasting channel bandwidth and penalizing system performance.

Hereafter, we present an algorithm that alleviates the second issue by adjusting the starting window size of the binary exponential backoff mechanism, CW_{min}, to some optimal value that depends on the current number of active stations in the network. The optimality of this value is computed in a way to maximize system throughput and to minimize frames delivery time. Extension of this algorithm to react to the first issue is more complicated and is not considered by the algorithm since the transmitting station needs to know the origin of missed ACKs in order to respond, and which results from various factors.

To proceed, recall from Section 6.3.1 that optimal system performance can be achieved in different network topologies by well tuning the transmission probability of the stations to the optimal transmission probability τ^* given in Equation 6.6. For a given network topology, transmitting at the optimal transmission probability can be achieved by well adjusting the size of the minimum starting contention window $CW_{\min}$. Note that a one-to-one relation between these two parameters can be obtained by inverting Equation 6.1, and $CW_{\min}$ will then be expressed in terms of τ^* as follows :

$$CW_{min} = \frac{(2-\tau^*)(1-2p)}{\tau^*(1-p-p(2p)^m)} \tag{6.12}$$

This latter relation between $CW_{\min}$ and τ^* motivates our algorithm, hereafter referred to as *Adaptive Binary Exponential Backoff* (Adaptive BEB), which can be summarized as follows. After a failed transmission, the behavior is the same as in the standard, i.e. the contention window size is doubled. However, after a successful transmission, the starting window size $CW_{\min}$ of the binary exponential backoff is set to an optimal value mirroring the current network state. To keep the algorithm as close as possible to the standard, we will consider that the minimal contention window can only take value in the set of values defined by the standard $\{2^j CW_0, j=0,\ldots,m\}$.

The algorithm is given in Algorithm 4. Given an estimation of N, the approximate optimal transmission probability is computed using Equation 6.6 (line 1), enabling the computation of the collision probability p (line 2). The optimal minimal size of the contention window size is computed according to Equation 6.12 (line 3). Note that to be as close as possible to the standard, $CW_{\min}$ is imposed to take only those values defined by the standard $\{2^j CW_0, j=0,\ldots,m\}$ (lines 4–5). Last, the value of m is updated (line 6). Observe

[2] CW_0 refers to the size of the minimal contention window that is defined by the standard.

Algorithm 4 Adjustment algorithm.

Require: An estimation $\overline{N}$ of the number of active nodes
Ensure: Sub-optimal values of $CW_{\min}$ and m
1: Set $\tau^* = \frac{1}{\overline{N}\sqrt{T_c/2\sigma}}$
2: Set $p = 1 - (1 - \tau^*)^{\overline{N}-1}$
3: Compute $cw = \frac{(2-\tau^*)(1-2p)}{\tau^*(1-p-p(2p)^m)}$
4: Select j such that $2^j CW_0$ is the closest to cw, $j \in [0, m]$
5: Set $CW_{\min} = 2^j CW_0$
6: Set $m = \log_2 \frac{CW_{\max}}{CW_{\min}}$ $\{CW_{\max}$ is fixed$\}$

that the values of $CW_{\min}$ and m will be sub-optimal due to the fact that we made $CW_{\min}$ discrete. Observe also that the parameter m will still be dependent on the previous value of $CW_{\min}$ as follows $m = \log_2(CW_{\max}/CW_{\min})$. However, this will not impact the new selected value for $CW_{\min}$ since the term $p(2p)^m$ will be negligible with respect to 1. Note that the value of $CW_{\max}$ stays constant while the value of m reflects the size of the starting window.

The analytical throughput achieved by our adaptive algorithm can be obtained by substituting the new values of $CW_{\min}$ and m and the optimal value of p given in line 2 of Algorithm 4 to compute the sub-optimal value of τ, and therefore the achieved throughput by using Equation 6.3.

6.5 Estimation of the number of contending stations

To adjust the value of $CW_{\min}$ in Algorithm 4, we need an estimation of N, hereafter denoted $\overline{N}$, that tracks its time-evolution. However, since the value of $CW_{\min}$ is made discrete, a reactive and adequately accurate estimator is sufficient. For instance, consider the case when $N = 30$. When applying Algorithm 4, we would obtain the correct optimal value $CW_{\min} = 2^4 CW_0$ for $\overline{N} \in [20.45, 40.95]$.

As opposed to previously proposed methods [26, 15, 16] which estimate the current number of sources by measuring the channel activity, we propose to estimate the number of active stations directly by counting as active each station from which the concerned measuring station receives a *sign of life*, i.e. error-free data or RTS packets. The measurement period will be the virtual transmission time of the station, i.e. the time between two consecutive successful transmissions. Specifically, since each active node is always filtering received packets whether they are destined to it or not, it can thus keep trace of a large number of active nodes in a given time interval by counting the number of *distinct* signs of life received in that time interval from these nodes.

Let $\widehat{N}_n$ denote the count of distinct signs of life at the nth measurement period of a given station. For multiple reasons, these measurements cannot be used directly in Algorithm 4 as an estimation of N. First, the count of signs of life cannot account for stations whose transmitted packets are corrupted, either by noise or collisions. Second, variable and relatively small measurement periods result in counting only a variable portion of all active stations. Therefore, based on the measurements $\{\widehat{N}_n\}_{n\in\mathbb{N}}$, an unbiased estimator should be devised. Observe that measurements collected over relatively large measurement times are more "accurate" than those collected over small measurement times and should be preferentially treated. By observing that the expected length of a measurement period is reflected by the value reached by the contention window at the end of the measurement time, one can use this value to proportionally weight the corresponding measurement, so that the following filter $\sum_{i=0}^{q-1} cw_{n-i}\widehat{N}_{n-i}/\sum_{i=0}^{q-1} cw_{n-i}$ can be used. Simulation analysis have shown that low values of the filter order q will suffice for a convenient performance, so hereafter, we set $q = 3$ so as to heighten reactiveness.

However, since corrupted packets cannot contribute to the collected measurements, the previous ratio, henceforth referred to as the observation, underestimates N and needs thus to be corrected. To derive an appropriate correction on the observations, we have proceeded through simulations. We have performed extensive simulations with different values of N, and we have plotted on Figure 6.4(a) the mean values of the observations obtained through simulations against the actual number of sources. As the figure illustrates, a linear correction is needed, resulting in the following estimator

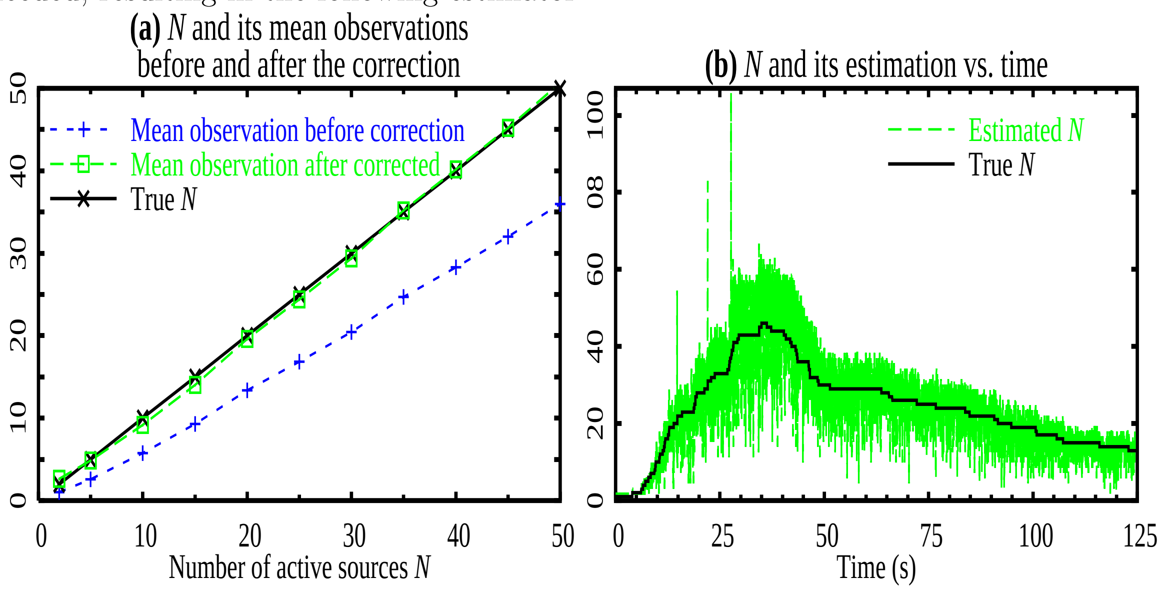

FIG. 6.4 – **(a)** Mean observations of N vs. N. **(b)** Estimation of time-varying N vs. time.

$$\overline{N}_n = a\left(\sum_{i=0}^{q-1} cw_{n-i}\widehat{N}_{n-i}\right) / \left(\sum_{i=0}^{q-1} cw_{n-i}\right) + b \tag{6.13}$$

with $a = 1.35$, $b = 1.75$ for $q = 3$. These latter values need to be adjusted in the case of noise over the channel. Figure 6.4(b) illustrates the estimation of N, $\overline{N}_n$, over a simulation in which nodes arrivals are Poisson and activity times are exponentially distributed. The estimator exhibits good reactiveness to changes in N at the cost of large fluctuations, but these have only a limited impact on the performance of Algorithm 4 because of the discretization process.

TAB. 6.2 – Simulation parameters.

Parameter	Value	Parameter	Value	Parameter	Value
$CW_{\min}$	32	Data rate	2 Mbps	Slot σ	$20\mu s$
$CW_{\max}$	1024	T_c	$4335\mu s$	$N_{\max}$	50

6.6 Simulation Results

In this section, we investigate the effectiveness of our proposal through a series of simulations conducted in `ns-2` [3]. Both system throughput and end-to-end packet delay achieved with our algorithm are compared to those obtained with the standard and with the slow decrease algorithms under various simulations scenarios. In the following, we will refer to the additive increase decrease algorithm [36] and the multiplicative slow decrease algorithm [83, 4] as "Additive" and "Multiplicative". The implementations of the latter two algorithms consider the values of the parameters which maximize the algorithms performance.

All simulations presented hereafter consider the basic access mechanism with a packet payload set to 1050 bytes and channel data rate of 2 $Mbps$. Simulated nodes, at the maximum number of 50 are uniformly distributed in a 100×100 m^2 square area and the power transmission is sufficiently high so that all the nodes are within communication range. The nodes work in saturation conditions and all the sources generate Constant Bit Rate (CBR) traffic. The protocol parameters are as follows : $CW_0 = 32$, $CW_{\max} = 1024$, and the physical characteristics are those for the DSSS technique (see [2]). Table 6.2 summarizes the values of the parameters used in the simulations. The expected collision duration T_c used in Equation 6.6 is set to a constant value corresponding to the *Maximum Packet Data Unit* (MPDU) delivered by the MAC layer to the PHY layer.

6.6.1 Comparing analytical and empirical throughput

In this section, we investigate on the validity of the analytical model. To this end, we will compare the analytical results with the ones obtained from simulations. In fact, Equation 6.3 allows the derivation of the optimal analytical throughput for a given network topology by substituting the value of the transmission probability τ in that equation with its optimal value for that topology. Recall from Section 6.4 that the analytical throughput of our algorithm, noted hereafter as the sub-optimal throughput, can be computed by using the sub-optimal value of τ obtained by using the new value of CW_{min}.

Figure 6.5 plots the optimal analytical throughput, the sub-optimal analytical throughput and the empirical throughput of the algorithm against the number of active nodes N. Note that, for each value of N, we have ran four simulations each with a duration of 200 seconds. For each simulation, we considered a fixed number of nodes set to N, and then we averaged the values of the system throughput obtained in these four runs. Observe that the difference between the optimal and the sub-optimal analytical throughput is small. On the other hand, the average difference between the sub-optimal throughput and the simulative throughput is 5.15%.

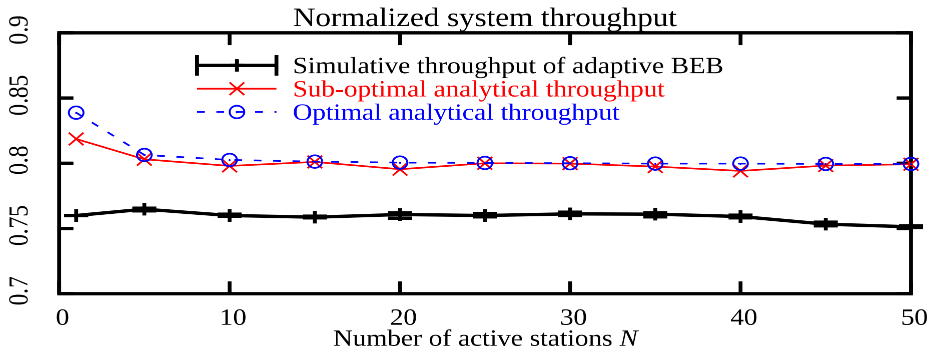

FIG. 6.5 – Analytical and simulative throughput of the adaptive BEB algorithm.

6.6.2 Ad hoc mode

In this section, we report the results of simulations performed in ad hoc mode having all nodes within range of communication. Two scenarios have been considered and will be discussed separately in the subsequent paragraphs.

6.6.2.1 Abrupt change in the number of nodes

We consider here an abruptly changing configuration where 3 different groups of stations are activated and idled at different instants. Specifically, a first group of 5 nodes start transmitting at time $10s$; at time $25s$, 25 other nodes start transmitting ; the third and last group of nodes starts transmitting at time $45s$. At times $65s$ and $80s$, the second and third groups are idled, and at time $110s$, the first group stops transmitting. The objective of this scenario is to investigate on the performance reactiveness of the adaptive BEB and to compare with the other proposed algorithms.

Figures 6.6(a) and 6.6(c) illustrate respectively the system throughput and the packet delay over time obtained when using the standard and the three adaptive algorithms : Adaptive BEB, Additive and Multiplicative. Observe from the curves of Figures 6.6(a) and 6.6(c) that, upon the arrivals of the stations, the adaptive BEB is the fastest to react to network change. This can be explained by the fact that a large CW_{min} is rapidly selected, consequently avoiding a certain number of collisions. This is not the case of both slow decrease approaches who witness a sharp deterioration in their performance at these instants. Considering that the buffers at the different nodes take some time to get empty after the sources have idled, the effective number of contending nodes decreases smoothly. Therefore, the size of the contention window decreases smoothly upon stations idleness, and hence no sharp transient period is observed at nodes idleness instants.

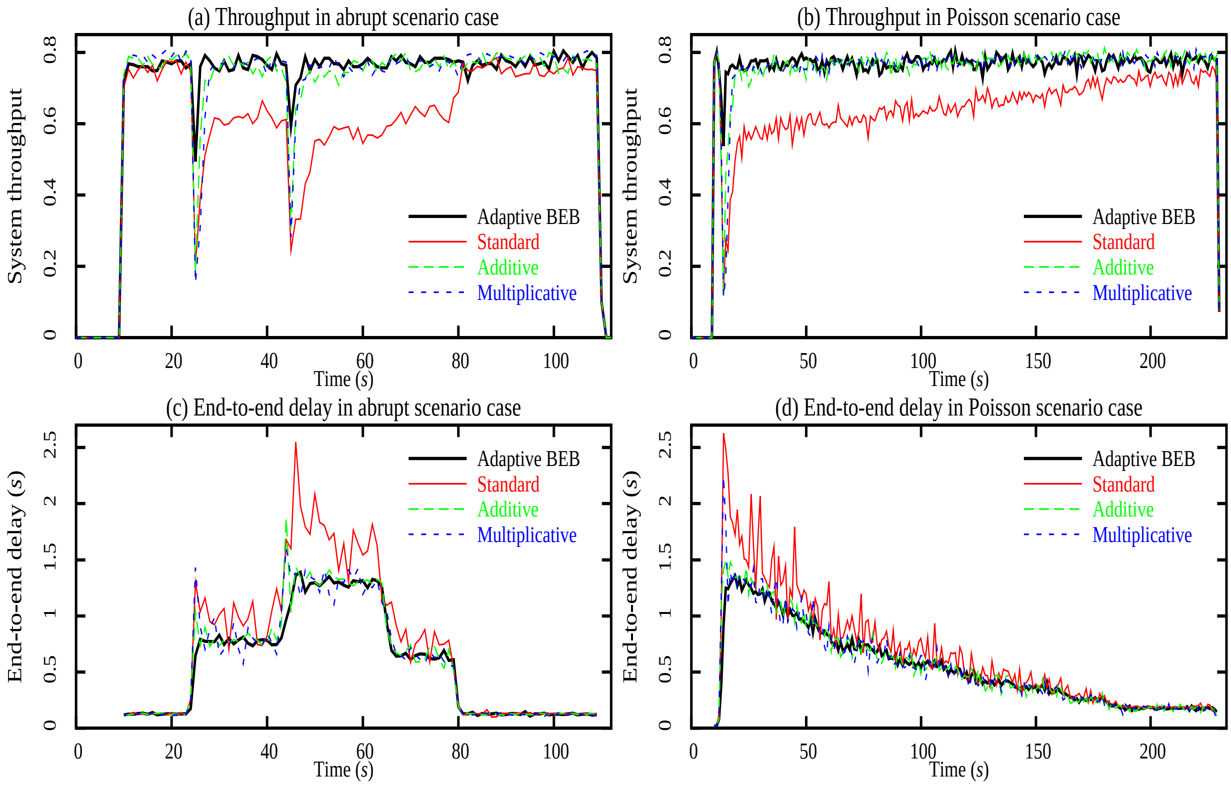

FIG. 6.6 – Simulations results in **ad hoc** mode of operation : Overall throughput and end-to-end delay over time in abrupt and Poisson scenario cases.

6.6.2.2 Poisson arrivals with exponential activity times

In this scenario, we consider a Poisson process with an arrival rate of 5 stations per second and an exponentially distributed activity time with a mean of $100s$. Figures 6.6(b) and 6.6(d) respectively illustrate the system throughput and the packet delay over time, obtained when using the standard and the three adaptive algorithms. As in the previous scenario, we observe a sharp transient period at the beginning of the simulation for the standard, but also for the two slow decrease approaches. This is due to the high stations arrival rate and the slow response time of the standard and the slow decrease approaches to adequately control their contention window sizes. On the other hand, adaptive BEB reacts very well. The small transient period observed at the beginning of the simulations is due to the time needed for all stations to estimate the number of active nodes and consequently to tune their contention window accordingly. Looking at Figure 6.6(b), we can see that the throughput achieved by the IEEE 802.11 DCF function increases with time. As a matter of fact, in the Poisson scenario, all the 50 nodes are activated in the first tens of seconds of simulation, so the number of active nodes can only decrease after that. The minimal contention window of the standard is equal to 32, a value that is optimal when there are 5 nodes. Hence, as time goes on, nodes end their activity period and the standard achieves an increased performance over time.

6.6.3 Infrastructure mode

In this section, we address the simulations performed in infrastructure mode of operation with traffic being uploaded to a single access point. We consider in this mode the abrupt and Poisson scenarios introduced in Sections 6.6.2.1 and 6.6.2.2 respectively. Results of the system throughput and packet delay obtained for each one of these scenarios are illustrated in Figure 6.7.

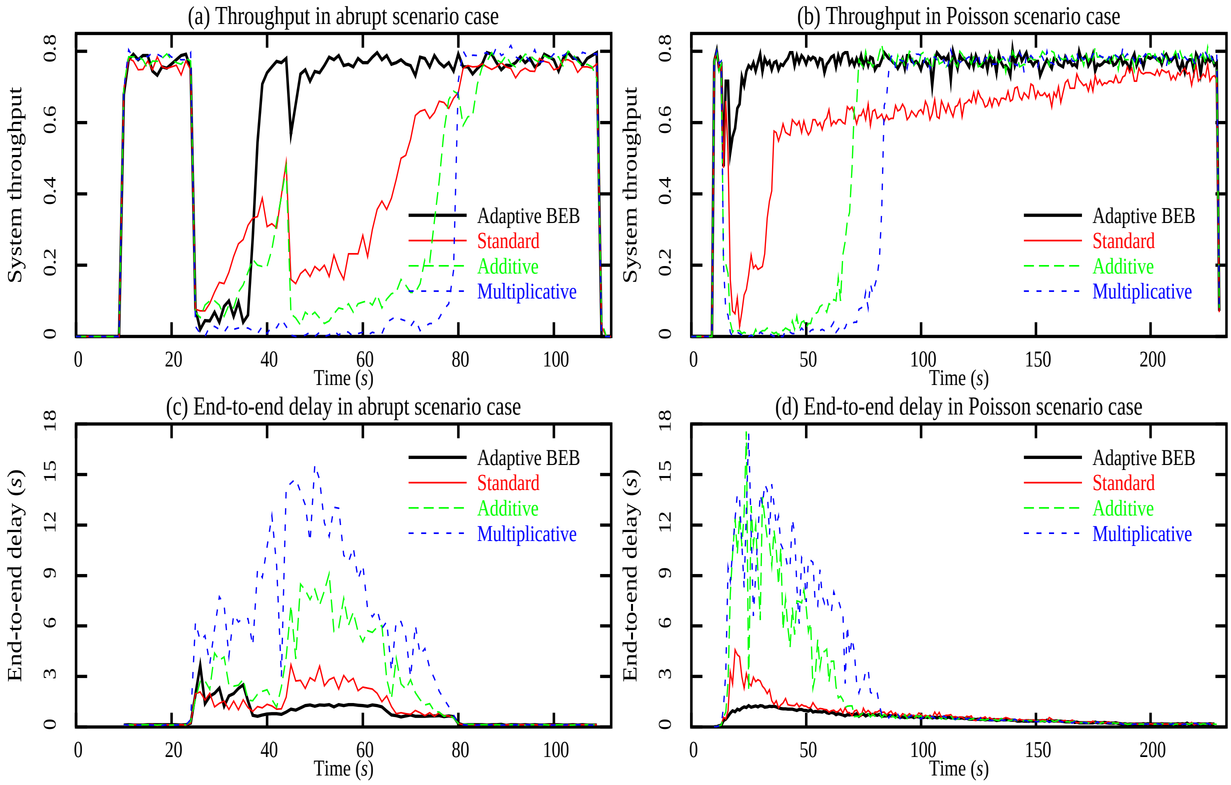

FIG. 6.7 – Simulations results in **infrastructure** mode of operation : Overall throughput and end-to-end delay over time in abrupt and Poisson scenario cases.

The simulation results in the case of infrastructure mode are essentially different from the ones obtained in ad hoc mode. All considered algorithms are more or less affected by changes in the number of active stations. Also, in contrast with the results obtained for the slow decrease approaches in ad hoc mode, the results obtained in the case of infrastructure mode show, in some cases, performances below the standard. On the other hand, the adaptive BEB keeps, to a large degree, a uniform and steady performance in both modes, and under the two scenarios. The sharp drop seen in the performance of the algorithms can be explained by the fact that in the case of operation in infrastructure mode, multiple flows share the access point buffer, which is of limited size. When that buffer fills up, the receiving station starts dropping packets, which leads to lost packets and subsequent missed ACKs. Lost packets result in a deteriorated performance of the system which will be seen by all algorithms. Missed ACKs are severely penalizing the slow decrease approaches, since each missed ACK is considered as a result of a packet collision, and hence the size of the contention window is consequently increased to large values extending considerably the assumed optimal values for the current network state. In contrast, the adaptive BEB will not suffer from missed ACKs since the increase of the minimal contention window is based on an estimation of the number of contending stations ; only lost packets will decrease the system performance.

6.6.4 Performance in noisy environments

To evaluate the performance of our proposed algorithm in noisy environments, we have considered two error models, namely the Bernoulli and Gilbert models. These two models capture, to a large extent, the occurrence of the errors on the channel in a random and in a burst manner, respectively. Specifically, in the Bernoulli model, errors occur independently on the channel with a certain Packet Error Rate (PER), while in the Gilbert model, the channel is modeled as a 2-state Markov process, the states being named *Good* and *Bad* [38]. When the channel is in the *Good* state, the error rate is 0, and it is high in the *Bad* state. The Gilbert model permits to specify the mean time spent in each state, namely T_G and T_B, and the transition rates between these states. In our simulations, the parameters for the Gilbert model have been set as follows :

- $\text{PER}_G = 0$ and $\text{PER}_B = 0.4$;
- The transition rates are set equal to 1 ;
- T_B is 0.5 seconds, and T_G is set in such a way that the expected PER obtained in the Gilbert model is equal to the PER value in the Bernoulli model.

We have considered PER values of $0.05, 0.1, 0.2, 0.3$ and 0.4 for both error models. Figures 6.8 and 6.9 plot the system throughput and the packet delay for $5, 10$ and 15 sources versus different values of PER in the case of Bernoulli and Gilbert error models, respectively. Note that each value on the curves is averaged over four simulation runs, each of $200s$ duration.

As the figures illustrate, the adaptive BEB exhibits robustness and stable behavior when either error models are considered, even in the case of small number of sources. The drop observed in the throughput is mainly due to the loss of corrupted packets. Corrupted packets and subsequent missed ACKs make the algorithm unnecessarily double the size of the contention window, leading to a decrease in the system throughput. However, the effect of channel errors will have less impact in the case of large number of active nodes. In such cases, CW_{min} will already be large and channel errors will then have lower influence on the performance of the Adaptive BEB.

As for the two slow decrease algorithms, we observe a stronger degradation in their performance. The fact that these algorithms keep the value of CW close to the previous value is highly penalizing since that value does not reflect in these situations the optimal value of the congestion level in the network. Hence, the contention window will be often high even in the case of small number of stations.

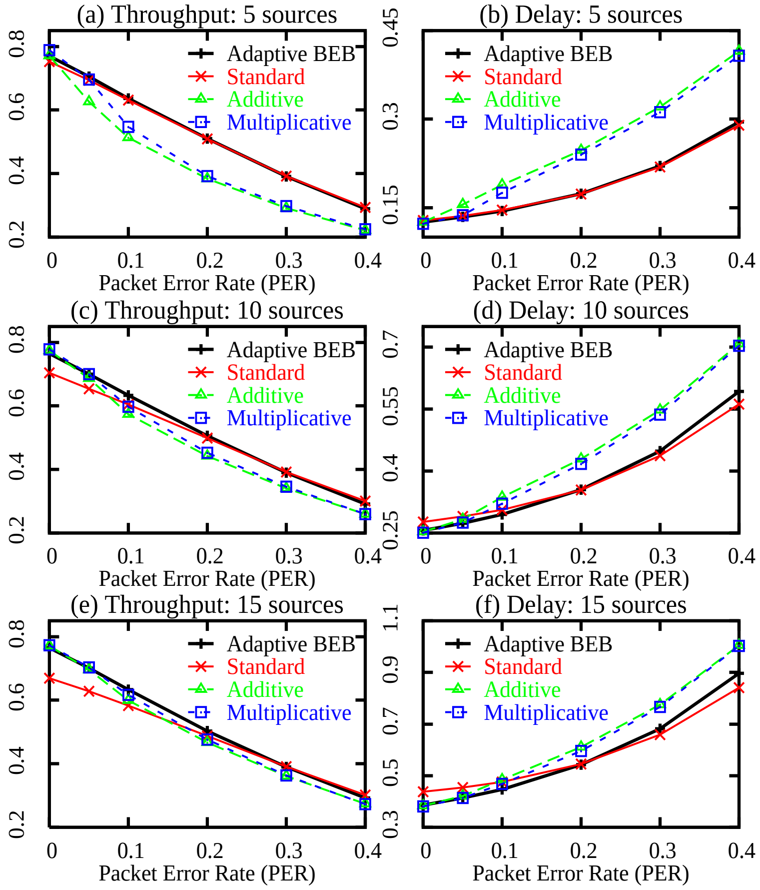

FIG. 6.8 – Impact of noisy channel on system throughput and packet delay (Bernoulli losses).

6.7 Fairness Analysis

Fairness describes the MAC protocol capability to distribute the available resources equally among communicating terminals [72]. There are a variety of fairness definitions intended to support different QoS and service differentiation. In our study, fairness is simply the equal partitioning of the bandwidth among all flows at hand.

In the following, we compare for the fairness in terms of channel access of the adaptive BEB and the standard. To proceed, we have considered a network that is constituted by a fixed number of N active stations working in ad hoc mode. At time $10s$, all the N CBR sources are activated and at time $100s$ they stop their transmissions. We have run this scenario in the case of $N = 10, 30$ and 50 sources. For each case, we compute the index of fairness of Jain *et al.* [53] which is given by

$$f(x_1, x_2, \ldots, x_I) = \frac{\left(\sum_{i=1}^{I} x_i\right)^2}{I \sum_{i=1}^{I} x_i^2}$$

where x_i denotes the allocation for user i, and I the number of users sharing the bandwidth. The index f is between 0 and 1. If the bandwidth is equally partitioned, the index of fairness is 1. If only $k\%$ of the flows are treated fairly while the others are not, then the index of

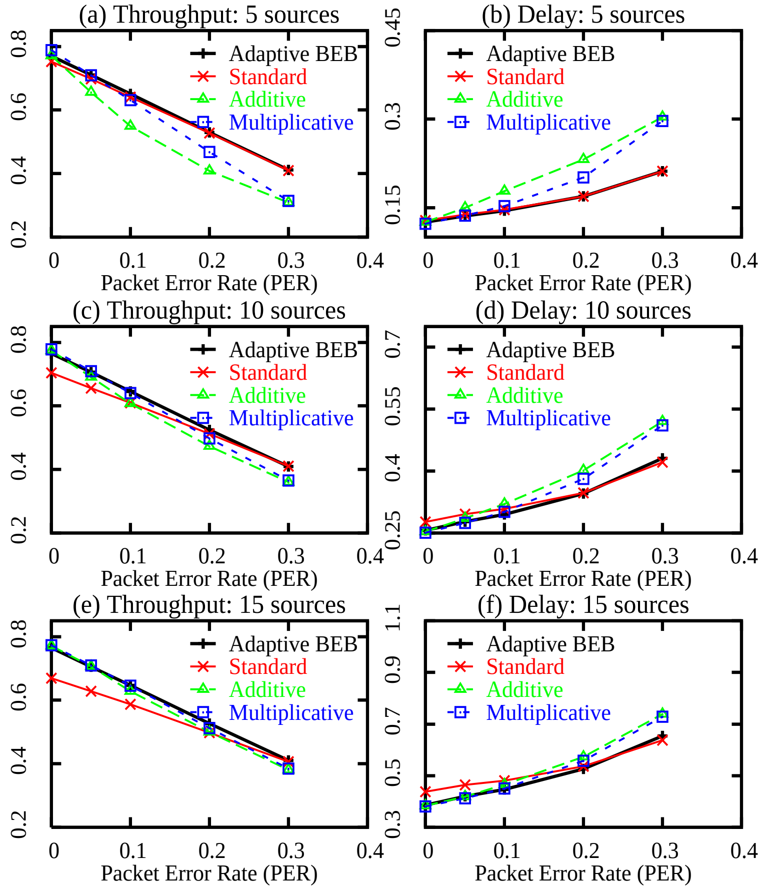

FIG. 6.9 – Impact of noisy channel on system throughput and packet delay (Gilbert losses).

fairness is equal to $k\%$.

To compute the index of Jain, we have used the Sliding Window Method (SWM) introduced in [65]. Specifically, considering a trace of channel accesses, the fraction of channel accesses per flow is computed within a sliding window. For each position of the sliding window, the index of fairness is computed by setting x_i equal to the fraction of channel accesses achieved by flow i. As we slide the window one element at a time, we obtain a series of values of f that are averaged at the end to return the final value of the index of fairness for this particular size of the sliding window. A small sliding window size allows the study of the short-term fairness whereas large sliding window size enables the long-term fairness one. The results, reported in Figure 6.10, show that our algorithm achieves better fairness (higher index value) than the standard in the three considered cases. Indeed, the fact that our algorithm assigns the contending nodes optimal and relatively equal CW_{min} allows on one hand to minimize the number of collisions, and consequently the risk of having largely different backoff intervals, and on the other hand, it permits to obtain equal opportunities of accessing the channel for the various nodes.

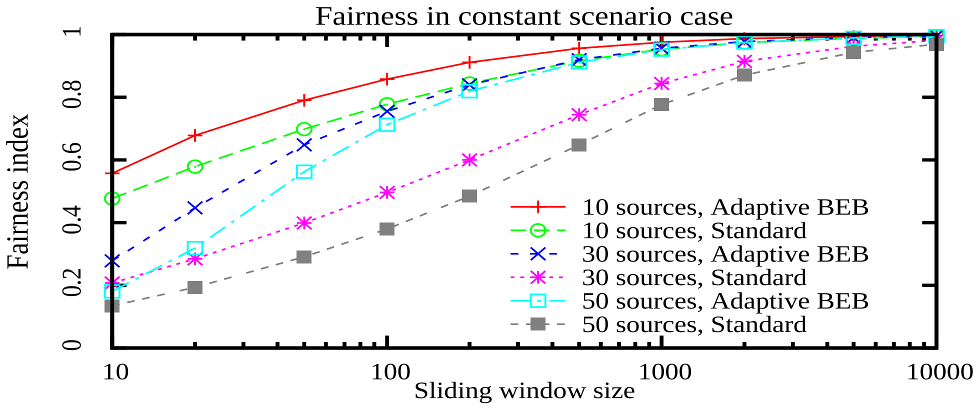

FIG. 6.10 – Index of fairness versus the size of the sliding window.

6.8 Conclusion

In this chapter, we have proposed and evaluated a simple, efficient and robust adaptive mechanism that can be easily incorporated within the IEEE 802.11 DCF protocol in order to enhance its performance. This mechanism relies on on-line counts of *signs of life* to estimate the number of active nodes in the wireless network. The minimal contention window size is adjusted accordingly after each successful transmission to optimize the transmission probability. The performance of our algorithm, termed Adaptive BEB, has been evaluated through extensive simulations in `ns-2`. The proposed algorithm performs remarkably well in all investigated situations. It outperforms the standard in both ad hoc and infrastructure mode of operation, whether the number of active sources is constant, changes smoothly or abruptly. It also outperforms other proposed algorithms in infrastructure mode of operation and in noisy environments. The fact that the control of the contention window in our proposed algorithm is done based on a estimation of the number of contending stations results in these enhanced performances. The estimation allows to capture the real state of the network, and hence allows to adapt the window size to the actual congestion level in the network. The index of fairness has been also computed under various network settings showing that the adaptive BEB algorithm achieves better fairness than the IEEE 802.11 standard.

6.A Overview of IEEE 802.11 and DCF

Released in 1997 by the Institute of Electrical and Electronics Engineers (IEEE), the 802.11 protocol specifies a standardized architecture to support data transmission over wireless links. In analogy with previous IEEE 802 standards, the 802.11 standards focus on the bottom two levels of the ISO model [99], the physical layer and the data link layer. The physical layer defines three physical technologies to interact with the data link layer, which are :

– Frequency Hopping Spread Spectrun (FHSS) in the 2.4 GHz band.
– Direct Sequence Spread Spectrum (DSSS) in the 2.4 GHz band.
– InfraRed.

The data link layer consists of two sublayers : Logical Link Control (LLC) and Media Access Control (MAC) [2]. 802.11 uses the same LLC sublayer and 48-bit addressing as other 802 LANs, thus allowing for simple bridging from wireless to IEEE wired networks. Similarly to the 802.3 MAC protocol of wired network, the 802.11 MAC protocol is designed to regulate the access of multiple users on a shared medium, and to handle collisions that occur when two or more stations try to transmit simultaneously.

The specifications of 802.11 outline two possible modes of operation : infrastructure mode and ad hoc mode [2]. In the infrastructure mode, nodes, either fixed or mobile, communicate with fixed base stations called access points (APs). These access points are connected to wired networks hence bridging wireless nodes to other wired nodes. In the ad hoc mode, there is no base station, and nodes can communicate directly with each others as in a peer-to-peer network. Two different medium access coordination functions have been defined in 802.11 : *Distributed Coordination Function* (DCF) and *Point Coordination Function* (PCF).

The Distributed Coordination Function (DCF) is the primary access protocol in 802.11 for sharing the wireless medium between active stations, mostly fitted for delay-insensitive applications. It is the only possible way of medium access in the case of ad hoc networks, while it can be employed with the Point Coordination Function (PCF) in the case of infrastructured networks. The PCF is a contention-free access method, intended to support real-time transmissions by using a centralized polling mechanism. More specifically, the medium access is controlled by a polling coordinator, usually situated at the AP, that polls the participating nodes in the cell eliminating hence the contention between the nodes.

The DCF is based on the Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) protocol, with discrete time scale that is slotted into constant intervals called *Time Slot*. The carrier sense activity consists in listening to the channel before transmitting : if the channel is found to be idle for a time interval greater than the *Distributed InterFrame Space* (DIFS) period, the station transmits directly, otherwise it differs its transmission time until the ongoing transmission terminates. When the channel becomes idle again for a DIFS, the station enters the collision avoidance phase by selecting a random number of time slots called *backoff interval* that is used to initialize a *backoff counter* [2].

The backoff interval is an integer number of time slots where the value is randomly chosen in the interval $[0, CW-1]$. Here CW refers to the current size of the *Contention*

Window and gives the number of possible values of the backoff counter. Considering a particular frame transmission, CW starts with an initial value CW_0 called the minimum contention window size. After each non-acknowledged transmission of the packet, the size of CW is doubled until reaching a maximum value noted by $CW_{max} = 2^m CW_0$ where m denotes the maximum backoff stage. According to the standard, a maximum number of seven retransmissions are allowed before the frame is dropped. Backoff counter is a timer that is decreased as long as the channel is sensed idle, stopped when a transmission is in progress, and reactivated when the channel is sensed idle again for more than DIFS. When the backoff counter reaches zero, the station transmits directly, then waits for an *Acknowledgement* (ACK) from the destination station.

If the source station has not received an ACK within a specified *ACK timeout*, it will assume that the transmitted packet was lost due to a collision, and it will start the backoff procedure again after *doubling* its CW. On the other hand, if the packet is well received by the destination station, the latter will send an ACK to the source station after a period of time, much shorter than a DIFS, called *Short InterFrame Space* (SIFS). Upon receiving the ACK, the source station will reset its contention window size to the minimal value CW_0 and proceed to the next frame in the buffer. For the remaining frames and in order to avoid channel capture, the source station waits for a random backoff time even if the medium is sensed idle for more than a DIFS. It is worth to mention here that the values CW_0 and CW_{max} are physical layer dependent. Figure 6.11 illustrates the operation of the DCF access mechanism.

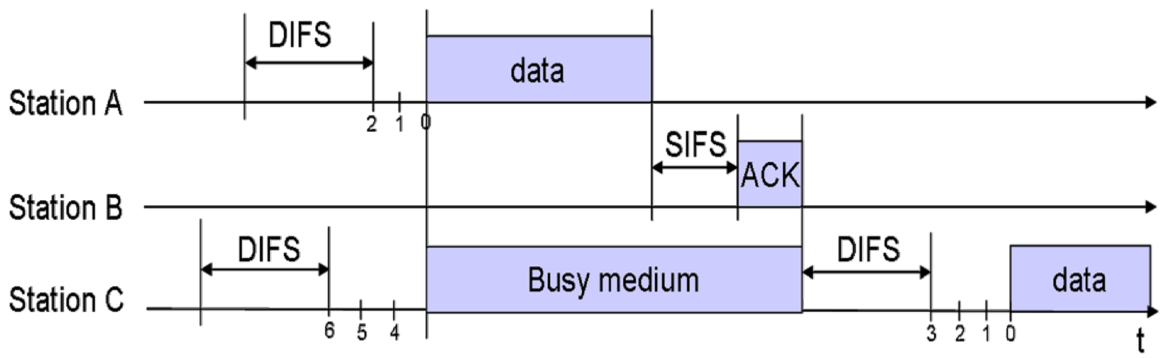

FIG. 6.11 – Basic access DCF mechanism.

In addition to the two-way handshaking technique described above and named basic access mechanism, DCF specifies another optional four-way handshaking technique called *Request To Send / Clear To Send* (RTS/CTS). RTS/CTS constitutes an extension of the basic access technique as it provides a solution for the so-called hidden node problem by reserving the channel prior to a transmission. Hidden node problem occurs when two transmitting stations cannot hear each but their simultaneous transmissions collide at a thrid station. The RTS/CTS technique works as follows : a station willing to transmit will first listen to the channel for a DIFS, follow the backoff procedure before transmitting a short control frame, the RTS frame, which will include both source and destination addresses and the duration of the transmission. If the destination station senses the medium free, it will respond after a SIFS with a Clear To Send (CTS) frame, which will include the same duration information. All stations receiving either the RTS or the CTS frames, will initialize an indicator called *Network Allocation Vector* (NAV) to the duration of the transmission and will use this vector to delay their transmissions, therefore reducing potential collisions.

RTS/CTS access mechanism is well suited for large packet transmissions as it greatly reduces the time spent in packet collision and in collision detection due to the small RTS frame size. However, its performance degrades in the case of small sized packets or in the case of sub-optimal channel usage.

Bibliographie

[1] Delay tolerant networks (DTN) research group. http ://www.dtnrg.org.

[2] *IEEE Std. 802.11b, Part 11 : Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification*, 1999.

[3] The network simulator, version 2.27, 2005. http ://www.isi.edu/nsnam/ns/.

[4] I. Aad, Q. Ni, C. Barakat, and T. Turletti. Enhancing IEEE 802.11 MAC in congested environments. In *Proc. of IEEE ASWN, Boston, Massachusetts*, August 2004.

[5] A. Al Hanbali, A. A. Kherani, and P. Nain. Simple models for the performance evaluation of a class of two-hop relay protocols. In *Proceedings IFIP Networking*, Atlanta, GA, USA, May 2007.

[6] A. Al Hanbali, P. Nain, and E. Altman. Performance of ad hoc networks with two-hop relay routing and limited packet lifetime. In *Proc. IEEE/ACM ValueTools*, Pisa, Italy, Oct. 2006.

[7] N. Banerjee, M. D. Corner, and B. N. Levine. An energy-efficient architecture for DTN throwboxes. In *To appear in Proc. of IEEE INFOCOM 2007*, Anchorage , Alaska , USA, May 6-12 2007.

[8] N. Bansal and Z. Liu. Capacity, delay and mobility in wireless ad-hoc networks. In *Proc. of IEEE Infocom*, Mar. 2003.

[9] D. Bertsekas and R. Gallager. *Data Network.* Prentice Hall, 1992.

[10] C. Bettstetter. Mobility modeling in wireless networks : Categorization, smooth movement, border effects. *ACM Mobile Computing and Communications Review*, 5(3) :55–67, July 2001.

[11] Christian Bettstetter, Hannes Hartenstein, and Xavier Pérez-Costa. Stochastic properties of the random waypoint mobility model. *ACM/Kluwer Wireless Networks, Special Issue on Modeling and Analysis of Mobile Networks*, 10(5) :555–567, Sept. 2004.

[12] Christian Bettstetter, Giovanni Resta, and Paolo Santi. The node distribution of the random waypoint mobility model for wireless ad hoc networks. *IEEE Transactions on Mobile Computing*, 2(3) :257–269, July 2003.

[13] Christian Bettstetter and Christian Wagner. The spatial node distribution of the random waypoint mobility model. In *Proc. German Workshop on Mobile Ad-Hoc Networks (WMAN), GI Lecture Notes in Informatics*, Ulm, Germany, Mar. 2002.

[14] G. Bianchi. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications*, 18(3) :535–547, March 2000.

[15] G. Bianchi, L. Fratta, and M. Oliveri. Performance evaluation and enhancement of the CSMA/CA MAC protocol for 802.11 wireless LANs. In *Proc. of PIMRC, Taipei, Taiwan*, pages 392–396, October 1996.

[16] G. Bianchi and I. Tinnirello. Kalman filter estimation of the number of competing terminals in an IEEE 802.11 network. In *Proc. of IEEE INFOCOM, San Francisco, California*, March 2003.

[17] E. Brewer, M. Demmer, B. Du, M. Ho, M. Kam, S. Nedevschi, J. Pal, R. Patra, S. Surana, and K. Fall. The case for technology in developing regions. *IEEE Computer*, 38 :pp. 25–38, May 2005.

[18] J. Broch, A.D. Maltz, B.D. Johnson, Y.-C. Yu, and J. Jetcheva. A performance comparison of multi-hop wireless ad hoc network routing protocols. In *Proc. ACM Mobicom*, pages 85–97, Dallas, TX USA, October 1998.

[19] R. Bruno, M. Conti, and E. Gregori. Mesh networks : commodity multihop ad hoc networks. *Communications Magazine, IEEE*, 43(3) :123–131, 2005.

[20] John Burgess, Brian Gallagher, David Jensen, and Brian N. Levine. MaxProp : Routing for Vehicle-Based Disruption-Tolerant Networks. In *Proceedings of IEEE INFOCOM*, April 2006.

[21] John Burgess, Brian Gallagher, David Jensen, and Brian Neil Levine. MaxProp : Routing for Vehicle-Based Disruption-Tolerant Networks. In *Proc. of IEEE infocom*, Barcelona, Spain, April 2006.

[22] S. Burleigh, A. Hooke, L. Torgerson, K. Fall, V. Cerf, B. Durst, K. Scott, and H. Weiss. Delay tolerant networking : an approach to interplanetary internet. *IEEE Communications Magazine*, 41 :pp. 128–136, June 2003.

[23] S. Burleigh, L. Torgerson, K. Fall, V. Cerf, B. Durst, K. Scott, and H. Weiss. Delay-tolerant networking : an approach to interplanetary internet. *IEEE Comm. Mag.*, 41(6) :128–136, 2003.

[24] B. Burns, O. Brock, and B. Levine. Mv routing and capacity building in disruption tolerant networks. In *Proc. IEEE INFOCOM*, March 2005.

[25] F. Cali, M. Conti, and E. Gregori. 802.11 wireless LAN : Capacity analysis and protocol enhancement. In *Proc. of IEEE INFOCOM, San Francisco, California*, volume 1, pages 142–149, March 1998.

[26] F. Cali, M. Conti, and E. Gregori. Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit. *IEEE/ACM Trans. on Networking*, 8(6) :785–799, 2000.

[27] F. Cali, M. Conti, and E. Gregori. IEEE 802.11 protocol : design and performance evaluation of an adaptive backoff mechanism. *IEEE Journal on Selected Areas in Communications*, 18(9) :1774–1786, 2000.

[28] T. Camp, J. Boleng, and V. Davies. A survey of mobility models for ad hoc network research. *Wireless Communication and Mobile Computing (WCMC) : Special Issue on Mobile Ad Hoc Networking Research, Trends and Applications*, 2(5) :483–502, 2002.

[29] V. Cerf, S. Burleigh, A. Hooke, L. Torgerson, R. Durst, K. Scott, K. Fall, and H. Weiss. Delay-tolerant network architecture, IETF RFC 4838. http ://www.ietf.org/rfc/rfc4838.txt, April 2007.

[30] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott. Impact of human mobility on the design of opportunistic forwarding algorithms. In *Proc. of IEEE INFOCOM*, Barcelona, Spain, April 2006.

[31] A. Chaintreau, A. Mtibaa, L. Massoulié, and C. Diot. Diameter of opportunistic mobile networks. In *Proceedings of ACM Sigcomm CoNext*, December 2007.

[32] Imrich Chlamtac, Marco Conti, and Jennifer J. Liu. Mobile ad hoc networking : imperatives and challenges. *Ad Hoc Networks*, 1(1) :13–64, July 2003.

[33] J. Davis, A. Fagg, and B. Levine. Wearable computers as packet transport mechanisms in highly partitioned ad hoc networks. In *Proc. of 5 IEEE Intl. Symp. on Wearable Computers*, Zurich, Switzerland, 2001.

[34] Robert P. Dobrow. On the distribution of distances in recursive trees. *Journal of Applied Probability*, 33(3) :749–757, Sep. 1996.

[35] K. Fall. A delay tolerant network architecture for challenged internets. In *Proc. of ACM Sigcomm*, Karlsruhe, Germany, Aug. 2003.

[36] J. Galtier. Optimizing the IEEE 802.11b performance using slow congestion window decrease. In *Proc. of 16th ITC Specialist Seminar, Anvers, Belgique*, Septembre 2004.

[37] D. P. Gaver, P. A. Jacobs, and G. Latouche. Finite birth-and-death models in randomly changing environments. *Advances in Applied Probability*, 16 :715–731, 1984.

[38] E. N. Gilbert. Capacity of a burst-noise channel. *Bell System Technical Journal*, 39(5) :1253–1265, September 1960.

[39] D.J. Goodman, J. Borras, N.B. Mandayam, and R.D. Yates. Infostations : a new system model for data and messaging services. In *Proc. of IEEE VTC*, pages 969–973, Phoenix, AZ, USA, May 1997.

[40] R. Groenevelt. *Stochastic Models in Mobile Ad Hoc Networks.* PhD thesis, University of Nice-Sophia Antipolis, INRIA, Apr. 2005.

[41] R. Groenevelt, P. Nain, and G. Koole. The message delay in mobile ad hoc networks. *Performance Evaluation*, 62(1-4) :210–228, Oct. 2005.

[42] M. Grossglauser and D. Tse. Mobility increases the capacity of ad hoc wireless networks. *ACM/IEEE Transactions on Networking*, 10(4) :477–486, Aug. 2002.

[43] R.A. Guérin. Channel occupancy time distribution in a cellular radio system. *IEEE Transactions on Vehicular Technology*, 36 :89–99, August 1987.

[44] P. Gupta and P.R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, IT-46(2) :388–404, March 2000.

[45] Z. J. Haas. The routing algorithm for the reconfigurable wireless networks. In *Proc. of ICUPC'97*, San Diego, CA, Oct. 1997.

[46] Y. Han, R. J. La, A. M. Makowski, and S. Lee. Distribution of path durations in mobile ad-hoc networks – palm's theorem to the rescue. *Special Issue of Computer Networks Journal on Network Modeling and Simulation*, 50(12) :1887–1900, Aug. 2006.

[47] K. Harras, K. Almeroth, and E. Belding-Royer. Delay tolerant mobile networks (dtmns) : Controlled flooding schemes in sparse mobile networks. In *Proceedings of IFIP Networking*, 2005.

[48] D. Hong and S.S. Rappaport. Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures. *IEEE Transactions on Vehicular Technology*, 35(3) :77–92, Aug. 1986.

[49] Pan Hui, Augustin Chaintreau, James Scott, Richard Gass, Jon Crowcroft, and Christophe Diot. Pocket switched networks and the consequences of human mobility in conference environments. In *Proc. ACM SIGCOMM workshop on Delay Tolerant Networking and Related Topics (WDTN-05)*, Philadelphia, USA, August 2005.

[50] E. Hyytiä, P. Lassila, and J. Virtamo. Spatial node distribution of the random waypoint mobility model with applications. *IEEE Transactions on Mobile Computing*, 5(6) :680–694, Jun. 2006.

[51] Esa Hyytiä and Jorma Virtamo. Random waypoint model in n-dimensional space. Technical Report TD(04)032, COST279, Sep. 2004.

[52] Mouhamad Ibrahim, Ahmad Al Hanbali, and Philippe Nain. Delay and resource analysis in manets in presence of throwboxes. *Performance Evaluation*, 64 :933–947, 2007. Issues 9-12.

[53] R. Jain, W. Hawe, and D. Chiu. Quantitative measure of fairness and discrimination for resource allocation in shared computer systems. Technical Report DEC-TR-301, Digital Equipment Corporation, September 1984.

[54] S. Jain, M. Demmer, R. Patra, and K. Fall. Using redundancy to cope with failures in a delay tolerant network. In *Proceedings of ACM Sigcomm*, Oct. 2005.

[55] S. Jain, K. Fall, and R. Patra. Routing in a delay tolerant networking. In *Proceedings of ACM Sigcomm*, Aug 2004.

[56] S. Jain, R. C. Shah, W. Brunette, G. Borriello, and S. Roy. Exploiting mobility for energy-efficient data collection in sensor networks. In *IEEE WiOpt*, 2004.

[57] David B. Johnson. Routing in ad hoc networks of mobile hosts. In *Proc. of the Workshop on Mobile Computing Systems and Applications*, Santa Cruz, CA, Dec. 1994.

[58] David B. Johnson and David A. Maltz. Dynamic source routing in ad hoc wireless networks. Mobile Computing, edited by Tomasz Imielinski and Hank Korth, Chapter 5, pages 153-181, Kluwer Academic Publishers, 1996.

[59] Evan P.C. Jones, Lily Li, and Paul A.S. Ward. Practical routing in delay-tolerant networks. In *Proc. of ACM SIGCOMM Workshop on Delay Tolerant Networking*, 2005.

[60] Evan P.C. Jones and Paul A.S. Ward. Routing strategies for delay-tolerant networks. *Submitted to Computer Communication Review.*

[61] Philo Juang, Hidekazu Oki, Yong Wang, Margaret Martonosi, Li-Shiuan Peh, and Daniel Rubenstein. Energy-efficient computing for wildlife tracking : Design tradeoffs and early experiences with zebranet. In *the Tenth International Conference on Architectural Support for Programming Languages and Operating Systems ASPLOS-X*, San Jose, CA, Oct. 2002.

[62] Thomas Karagiannis, Jean-Yves Le Boudec, and Milan Vojnovic. Power law and exponential decay of inter contact times between mobile devices. In *Proceedings of ACM Mobicom*, 2007.

[63] S. Karlin. *A first course in stochastic processes*. Academic Press, 1968.

[64] Alexander Klemm, Christoph Lindemann, and Oliver P. Waldhorst. A special-purpose peer-to-peer file sharing system for mobile ad hoc networks. In *VTC Fall*, Oct. 2003.

[65] C. E. Koksal, H. Kassab, and H. Balakrishnan. An analysis of short-term fairness in wireless media access protocols (extended abstract). *Performance Evaluation Review, (Proc. of ACM SIGMETRICS, Santa Clara, California)*, 28(1) :118–119, June 2000.

[66] T. G. Kurtz. Solutions of ordinary differential equations as limits of pure jump markov processes. *Applied Probility*, 7 :49–58, 1970.

[67] Y. Kwon, Y. Fang, and H. Latchman. A novel MAC protocol with fast collision resolution for wireless LANs. In *Proc. of IEEE INFOCOM, San Francisco, California*, March 2003.

[68] J.-Y. Le Boudec. Understanding the simulation of mobility models with palm calculus. *Performance Evaluation*, 64(2) :126–147, February 2007.

[69] J.-Y. Le Boudec and M. Vojnovic. Perfect simulation and stationarity of a class of mobility models. In *Proc. of IEEE INFOCOM*, Miami, FL, Mar. 2005.

[70] Jérémie Leguay, Timur Friedman, and Vania Conan. Evaluating mobility pattern space routing for dtns. In *Proceedings of IEEE INFOCOM*, Barcelona, Spain, Apr. 2006.

[71] A. Lindgren, A. Doria, and O. Scheln. Probabilistic routing in intermittently connected networks. In *Proceedings of ACM MobiHoc (poster session)*, Maryland, USA, June 2003.

[72] S. Lu, V. Bharghavan, and R. Srikant. Fair scheduling in wireless packet networks. *IEEE/ACM Trans. on Networking*, 7(4) :473–489, (August 1999.

[73] S. Alouf M. Ibrahim. Design and analysis of an adaptive backoff algorithm for ieee 802.11 dcf mechanism. In *Proceedings of Networking 2006 Conference*, Coimbra, Portugal, May 2006.

[74] M. Mahmoud. Limiting distributions for path lengths in recursive trees. Probability in the Engineering and Informational Sciences, 5, 53-59., 1991.

[75] S. McCanne and S. Floyd. *The manual page of Network Simulator ns version 2, http ://www-mash.cs.berkeley.edu/ns/ns-man.html.*

[76] M. Mitzenmacher. Digital fountains : A survey and look forward. In *Proc. of IEEE Information Theory Workshop*, Texas, USA, Oct. 2004.

[77] J. W. Moon. The distance between nodes in recursive trees. London Math. Soc. Lecture Notes 13, 125-132, 1974.

[78] P. Nain, D. Towsley, B. Liu, and Z. Liu. Properties of random direction models. In *Proc. of IEEE INFOCOM*, Miami, FL, Mar. 2005.

[79] W. Navid and T. Camp. Stationary distributions for the random waypoint model. *IEEE Transactions on Mobile Computing*, 3(1) :99–108, 2004.

[80] W. Navidi, T. Camp, and N. Bauer. Improving the accuracy of random waypoint simulations through steady-state initialization. In *Proc. of the 15th International Conference on Modeling and Simulation*, 319-326, Mar. 2004.

[81] G. Neglia and X. Zhang. Optimal delay-power tradeoff in sparse delay tolerant networks : a preliminary study. In *ACM Sigcomm workshop on Challenged Networks*, Pisa, Italy, Sep. 2006.

[82] M. Neuts. *Matrix-Geometric Solutions in Stochastic Models : An Algorithmic Approach.* Johns Hopkins University Press, 1981.

[83] Q. Ni, I. Aad, C. Barakat, and T. Turletti. Modeling and analysis of slow CW decrease for IEEE 802.11 WLAN. In *Proc. of IEEE PIMRC, Beijing, China*, September 2003.

[84] PRISMS Lab, University of Massachusetts in Amherst. Diverse Outdoor Mobile Environment (DOME) Project : http ://prisms.cs.umass.edu/dome/.

[85] National Institute of Standards and Technology (NIST). http ://www.antd.nist.gov/.

[86] J. Ott and D. Kutscher. Drive-thru internet : IEEE 802.11b for automobile users. In *Proc. of IEEE infocom*, Hong Kong, Mar. 2004.

[87] S. PalChaudhuri, J. Le Boudec, and M. Vojnovic. Perfect simulations for random trip mobility models. In *Proc. of 38th Annual Simulation Symposium*, San Diego, California, April 2005.

[88] C. Perkins, E. Belding-Royer, and S. Das. Ad hoc on-demand distance vector (AODV) routing. RFC 3561, July 2003.

[89] Charles E. Perkins and Pravin Bhagwat. Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers. In *Proceedings of ACM Sigcomm*, pages 234–244, 1994.

[90] N. Sadagopan, F. Bai, B. Krishnamachari, and A. Helmy. Paths : analysis of path duration statistics and their impact on reactive manet routing protocols. In *ACM MobiHoc*, June 2003.

[91] L. Santalo. *Integral Geometry and Geometric Probability.* Cambridge Mathenatical Library, 2004.

[92] N. Sarafijanovic-Djukic, M. Piorkowski, and M. Grossglauser. Island hopping : Efficient mobility-assisted forwarding in partitioned networks. In *IEEE SECON 2006*, Reston, VA, Sept. 2006.

[93] K. Scott and S. Burleigh. Bundle protocol specification, IETF RFC 5050. http ://www.ietf.org/rfc/rfc5050.txt, Nov. 2007.

[94] T. Small and Z. J. Haas. The shared wireless infostation model : A new ad hoc networking paradigm. In *Proc. ACM MobiHoc*, Anapolis, MD, USA, June 2003.

[95] Tara Small and Zygmunt J. Haas. Resource and performance tradeoffs in delay-tolerant wireless networks. In *Proceedings of ACM Sigcomm workshop on Delay Tolerant Networking*, Aug. 2005.

[96] Thrasyvoulos Spyropoulos, Konstantinos Psounis, and Cauligi Raghavendra. Spray and wait : An efficient routing scheme for intermittently connected mobile networks. In *Proceedings of ACM Sigcomm workshop on Delay Tolerant Networking*, Aug. 2005.

[97] Thrasyvoulos Spyropoulos, Konstantinos Psounis, and Cauligi Raghavendra. Performance analysis of mobility-assisted routing. In *Proceedings of ACM/IEEE MOBIHOC*, May 2006.

[98] Richard S. Sutton. Learning to predict by the methods of temporal differences. *Machine Learning*, 3 :9–44, 1988.

[99] A. Tanenbaum. *Computer Computer Networks, Fourth Edition.* Prentice Hall PTR, 2003.

[100] A. Vahdat and D. Becker. Epidemic routing for partially connected ad hoc networks, 2000.

[101] Y. Wang, S. Jain, M. Martonosi, and K. Fall. Erasure-coding based routing for opportunistic networks. In *Proceedings of the ACM Sigcomm Workshop on Delay-Tolerant Networking*, Aug. 2005.

[102] Forrest Warthman. Delay-tolerant networks (DTNs) : A tutorial. http ://www.dtnrg.org/wiki/Docs, Mar. 2003.

[103] J. Widmer and J.-Y. Le Boudec. Network coding for efficient communication in extreme networks. In *Proceedings of the ACM Sigcomm Workshop on Delay-Tolerant Networking*, Aug. 2005.

[104] Y. Yang, J. Wang, and R. Kravets. Distributed optimal contention window control for elastic traffic in wireless LANs. In *Proc. of IEEE INFOCOM, Miami, Florida*, March 2005.

[105] J. Yoon, M. Liu, and B. Noble. Random waypoint considered harmful. In *Proc. of IEEE INFOCOM*, San Francisco, California, March 2003.

[106] E. Zhang, G. Neglia, J. Kurose, and D. Towsley. Performance modeling of epidemic routing. In *Proc. of Networking*, pages 827–839, Coimbra, Portugal, May 2006.

[107] Xiaolan Zhang, Jim Kurose, Brian Neil Levine, Don Towsley, and Honggang Zhang. Study of a Bus-Based Disruption Tolerant Network : Mobility Modeling and Impact on Routing. In *Proceedings of ACM Mobicom*, Sep. 2007.

[108] Z. Zhang. Routing in intermittently connected mobile ad hoc networks and delay tolerant networks : Overview and challenges. *IEEE Communications Surveys and Tutorials*, Jan 2006.

[109] W. Zhao, Y. Chen, M. Ammar, M. Corner, B. N. Levine, and E. Zegura. Capacity enhancement using throwboxes in DTNs . In *Proc. of IEEE Int. Conf. on Mobile Ad hoc and Sensor Systems*, Oct 2006.

[110] Wenrui Zhao and Mostafa Ammar. Message ferrying : Proactive routing in highly-partitioned wireless ad hoc networks. In *Proc. of the IEEE Workshop on Futrure Trends in Distributed Computing Systems*, Puerto Rico, May 2003.

[111] Wenrui Zhao, Mostafa Ammar, and Ellen Zegura. A message ferrying approach for data delivery in sparse mobile ad hoc networks. In *Proceedings of ACM Mobihoc*, Tokyo Japan, May 2004.

[112] Wenrui Zhao, Mostafa Ammar, and Ellen Zegura. Controlling the mobility of multiple data transport ferries in a delay-tolerant network. In *Proceedings of IEEE INFOCOM*, Miami Florida, 2005.

[113] Wenrui Zhao, Yang Chen, Mostafa Ammar, Mark D. Corner, Brian Neil Levine, and Ellen Zegura. Capacity Enhancement using Throwboxes in DTNs. In *Proc. IEEE MASS*, pages 31–40, Oct 2006.

Appendix A

Présentation des Travaux de Thèse

A.1 Introduction

Depuis leur émergence sur les marchés il y a dix ans, les réseaux sans fil ont permis *d'émanciper* l'accès aux réseaux, autrefois limité aux connections à travers des stations fixes, favorisant ainsi leur apparition dans de nombreux secteurs de la vie quotidienne. Par exemple des hots spots dans les aéroports, les gares, les cafés, les endroits publics, etc.

Parallèlement à leur prolifération, le déploiement et l'utilisation des réseaux sans fil ont fait apparaître de nouveaux défis techniques jadis inexistants dans les réseaux fixes traditionnels. L'un des problèmes majeurs concerne la conception d'un protocole efficace pour l'accès au canal qui prend en compte les caractéristiques particulières du milieu sans fil. Par conséquent, plusieurs variantes de protocoles d'accès ont été proposées, que ce soit dans le milieu académique ou bien à travers des produits commerciaux.

Aujourd'hui, le standard IEEE 802.11 [2] constitue le protocole le plus utilisé sur le marché et dont la version initiale a fait son entrée en 1997. Le IEEE 802.11 définit un moyen standardisé au niveau de la couche physique ainsi qu'au niveau de la couche d'accès au canal (MAC) dans les réseaux sans fil. Il définit aussi deux modes possibles d'opération pour les réseaux constitués d'un seul saut (c.-à-d. toutes les stations sont dans la même portée de transmission l'une de l'autre) : mode infrastructure et mode ad hoc. Dans le mode infrastructure, le réseau englobe une station de base fixe qui constitue le lien entre le réseau local et le monde extérieur et qui joue le rôle d'un routeur pour le réseau local. Dans le mode ad hoc, le réseau ne nécessite pas de station de base, et les divers noeuds, qu'ils soient fixes ou mobiles, peuvent établir des communications directes l'un avec l'autre d'une façon décentralisée. Malgré sa popularité, ce protocole s'avère inefficace dans certaines situations où sa performance se dégrade brusquement, notamment dans le cas où le réseau inclut un grand nombre de stations actives.

Au fil du temps, et pour supporter de nouvelles applications de plus en plus sophistiquées, les réseaux ad hoc sans fil multi-sauts (wireless multihop networks) ont fait leur émergence. Plus spécifiquement, dans ces réseaux, un noeud peut communiquer directement avec un autre noeud situé dans sa portée de transmission selon le mode ad hoc. Pour communiquer avec d'autres noeuds situés au-delà de sa portée, un noeud source fait usage des autres noeuds intermédiaires pour acheminer ses paquets en utilisant un certain protocole de routage. Les réseaux sans fil multi-sauts incluent les réseaux sans fil maillés (mesh networks) ainsi que les réseaux ad hoc mobiles (mobile ad hoc networks MANETs). Tandis que les réseaux sans fil maillés sont construits à partir de stations fixes et mobiles, les réseaux MANETs sont constitués seulement par des stations mobiles.

Contrairement aux réseaux maillés, le déploiement des MANETs était plutôt médiocre malgré les efforts significatifs de recherche dans ce domaine. En effet, conçues initialement pour des fins militaires, l'usage civil des MANETs a été restreint pour certaine applications spécialisées (réseaux ah hoc véhiculaires, réseaux de réhabilitation après un désastre, réseaux pour le contrôle de l'environnement et des habitas,...). La principale raison est due au fait que dans beaucoup de ces applications, la grande mobilité des noeuds rend difficile l'établissement et le maintien des routes à plusieurs sauts. Dans un autre cas, un chemin de bout en bout entre deux noeuds est inexistant, une conséquence directe de la mobilité des noeuds ou du partitionnement du réseau. Ce dernier cas des réseaux MANETs est désigné réseaux tolérants aux perturbations (delay tolerant networks DTNs) ou bien réseaux à connectivité intermittente. En l'occurrence, les protocoles de routage classiques conçus initialement pour les réseaux MANETs sont voués à l'échec.

Face à ces défis, une nouvelle approche de routage a émergé qui exploite la mobilité des noeuds pour mener le routage. Cette approche, désigné par "store-carry-and-forward" prend l'avantage de la mobilité des noeuds intermédiaires pour relayer les paquets aux autres noeuds [61, 55, 21]. En particulier, lors d'un contact, les noeuds intermédiaires stockent les paquets des autres noeuds, puis les transmettent lors des futurs contacts à leurs destinations respectives ou bien aux autres noeuds relais. Cette approche a été adoptée dans le réseau *ZebraNet* [61] qui est un DTN conçu pour supporter le suivi scientifique des zèbres dans la nature. Le réseau est constitué par des collier-capteurs qui ont été attachés à un groupe de zèbres, et par des stations de base de chercheurs qui ont été installées sur des véhicules qui se déplacent sporadiquement aux alentours. Le but de ce réseau est d'acheminer vers les stations de base les traces des données qui ont été collectées par les zèbres aux cours de leurs déplacements. Vu le manque de connectivité, le réseau se sert de la mobilité des zèbres pour effectuer le routage des traces collectées. En particulier, quand deux zèbres se rencontrent, leur capteurs échangent leurs donnés respectives, augmentant ainsi leur probabilité de livraison à une des stations de base.

Récemment, plusieurs études ont proposé différentes approches qui se basent sur la mobilité des noeuds pour assister dans le routage des paquets (voir Chapitre 1). Vu la dépendance de ces approches sur la mobilité des noeuds relais, leurs performances dépendent largement des caractéristiques des opportunités de contact entre ces noeuds comme par exemple la fréquence, la durée, ou bien la séquence de ces contacts. D'autre part, ces approches doivent tenir compte des caractéristiques particulières des DTNs. En effet, bien que les applications conçues pour les DTNs tolèrent relativement de grands délais (échange des emails, diffusion des informations, etc,) ces applications bénéficient également des petits délais. Par ailleurs, les noeuds constituant ces réseaux sont généralement caractérisés par des ressources limitées en termes d'énergie et de capacité de stockage [49, 61, 94]. Par conséquent, les diverses techniques proposées ont eu pour objectifs de maximiser le taux de livraison des paquets, de minimiser le délai de livraison et la consommation des ressources, ou bien de trouver un compromis entre ces critères.

A.1.1 Motivations et contributions de la thèse

La plupart des travaux qui ont porté sur les protocoles de routage pour les réseaux ad hoc mobiles à connectivité intermittente ont considéré généralement des réseaux constitués uniquement de noeuds relais mobiles. Dans cette thèse, on s'intéresse à évaluer l'amélioration potentielle de ces protocoles de routage dans un tel contexte lorsqu'on augmente le réseau par des relais fixes qu'on désigne comme boîtes. Les boîtes sont simplement des noeuds sans fil stationnaires qui jouent le rôle de relais fixes et qui ont des fonctionnalités améliorées en termes d'énergie et de capacité de stockage. Les boîtes peuvent être déployées d'une façon ad hoc dans le réseau, en utilisant des batteries pour un usage sur le court terme ou bien des panneaux solaires pour un usage sur le long terme. Par ailleurs, vu leur bas coût, elles peuvent être déployées en grand nombre dans le réseau afin d'augmenter les opportunités de contact entre les noeuds mobiles.

L'introduction des boîtes dans cette classe de réseaux est une nouvelle approche prometttant, essentiellement en termes des améliorations qui peuvent être apportées sur les performances des protocoles de routage, et ainsi sur les performances des applications derrières. Par exemple, Banerjee et autres dans [7] ont reporté des résultats collectés sur le réseau expérimental *UmassDieselNet* dans le cas où ce dernier a été augmenté par des boîtes. Ils ont montré que l'addition d'une seule boîte à ce réseau augmente le taux de livraison des paquets de 37% et réduit le délai de livraison des messages de 10%.

Cette thèse s'intéresse à la conception et à l'évaluation des protocoles de routage et d'accès au canal pour les réseaux sans fils. Dans la première partie, on se focalise principalement sur l'évaluation et sur la conception des protocoles de routage pour les réseaux à connectivité intermittente incluant des boîtes. Plus spécifiquement, on considère des réseaux ayant des contacts opportunistes où les noeuds mobiles se déplacent aléatoirement selon un certain modèle de mobilité et où les boîtes sont déployées selon une certaine distribution. La notion de contacts opportunistes signifie que les noeuds n'ont aucune information concernant leurs précédents ou futurs contacts avec d'autres noeuds. Ayant considéré un protocole de routage donné, nous cherchons à évaluer ses performances dans un tel réseau à travers deux critères. Le premier critère consiste à évaluer sa performance du point de vue de l'utilisateur qui serait reflété par le délai moyen de livraison des messages. Le deuxième critère consiste à évaluer sa performance du point de vue de l'opérateur qui serait reflété par la charge induite dans le réseau et qui est traduite par le nombre de copies générées.

Considérant le modèle précédent du réseau, nous avons montré à travers des simulations approfondies que le processus de contact entre une boîte et un noeud mobile qui se déplace selon un modèle de mobilité aléatoire (modèle Random Wapoint et modèle Random Direction) peut être bien approximé par un processus de Poisson. Nous avons également fourni une approximation précise de paramètre de ce processus ; nous avons montré qu'il dépend des paramètres des noeuds et du réseau, ainsi que de la densité de probabilité stationnaire de position du modèle de mobilité et de la densité de probabilité de distribution des boîtes. Par conséquent, ce résultat nous a permis de fournir des constatations concernant le déploiement optimal des boîtes dans le cas d'un modèle de mobilité dont la densité stationnaire spatiale est connue.

En se basant sur la précédente propriété des laps de temps exponentiel, et sur une autre propriété similaire pour les temps de contacts entre deux noeuds mobiles, nous avons introduit un modèle markovien à deux dimensions. Le but était d'évaluer les performances de deux protocoles de routage largement étudiés, à savoir le protocole de routage épidémique et le protocole de relais à deux sauts, dans un réseau MANET épars incluant des boîtes. Pour les deux protocoles, nous avons fourni des expressions closes pour la distribution du temps de livraison d'un paquet entre un noeud source mobile et un autre noeud destination mobile, ainsi que pour la distribution du nombre total des copies générées. L'analyse a été menée pour les deux cas où les boîtes sont connectées en maille ou bien totalement déconnectées.

Nous avons aussi proposé et évalué numériquement cinq stratégies de routage pour accomplir le routage des paquets dans un réseau MANET épars incluant des boîtes. Nous avons aussi introduit une plateforme stochastique basée sur des équations récursives afin de calculer numériquement et de comparer les performances de chacune de ces stratégies. La plateforme proposée est assez générale et elle peut être facilement généralisée pour prendre en compte d'autres métriques de performance ou bien d'autres approches de routage.

Nous avons aussi examiné la longueur moyenne du chemin de livraison en terme de nombre de sauts dans un réseau utilisant le protocole épidémique. En assumant que les laps de temps séparant deux rencontres successives entre deux noeuds suivent une distribution exponentielle, nous avons fourni une expression close pour cette métrique, désignée par le diamètre du protocole épidémique.

Dans la deuxième partie de la thèse, nous avons proposé un nouveaul mécanisme destiné à améliorer les performances de la fonction de coordination distribuée (Distributed Coordination Function (DCF)) de l'algorithme backoff du protocole IEEE 802.11. Le mécanisme proposé est simple et facile à intégrer avec le mécanisme du standard. Il présente deux nouvelles caractéristiques majeures. En particulier, après une transmission avec succès sur le canal, la station remet la valeur initiale de sa fenêtre de contention à une valeur optimale qui dépend du nombre de stations actives aux alentours. Pour estimer le nombre de stations actives, la station émettrice détecte leur présence en comptant le nombre de signaux émis durant un intervalle de temps, puis elle utilise ces échantillons comme entrées pour un filtre à moyenne mobile.

A.1.2 Organisation de la thèse

La thèse est divisée en deux parties. La première partie est constituée des chapitres 1-5, et se concentre principalement sur l'évaluation et la conception de protocoles de routage dans les réseaux tolérants aux perturbations quand ces réseaux incluent des noeuds relais fixes, appelés boîtes. La deuxième partie, constituée uniquement du chapitre 6, se concentre sur la conception et l'évaluation d'un algorithme de backoff dynamique pour le standard IEEE 802.11.

A.2 Chapitre 1 : Etat de l'art des techniques de routage dans les réseaux tolérants aux perturbations

Les réseaux tolérants aux délais et aux perturbations (DTN for *Delay/Disruption Tolerant Networks*), aussi appelés des réseaux à connectivité intermittente, sont des réseaux ah hoc mobiles sans fil où souvent un chemin de communication entre une station source et une station destination n'existe pas, que se soit directement ou bien à traves des routes constituées par des stations intermédiaires. Cette situation apparaît dans le cas où le réseau est épars (suite à une mobilité élevée, une densité faible, etc.) et/ou lorsque le réseau s'étend sur de longues distances. Dans ces cas là, les protocoles traditionnels de routage qui ont été développés pour les réseaux ad hoc mobiles se montrent inefficaces puisqu'ils nécessitent ou supposent l'existence d'un réseau connecté pour pouvoir acheminer les paquets de contrôle ou bien de donné. Pour pallier à ce problème, les protocoles de routage qui ont été proposés pour ce nouveau type de réseaux se servent essentiellement de la mobilité des stations. Plus précisément, ils se basent sur une approche intitulée en anglais "*store-carry-and-forward*" où la source peut transmettre son paquet, ou éventuellement des copies, à d'autres stations intermédiaires lorsque la source passe dans leur portée. Ces stations se comportent comme des relais permettant ainsi de livrer le paquet ultérieurement à sa destination. Dans ce chapitre, nous dressons l'état de l'art des différentes techniques de routage dans les réseaux tolérants aux délais et aux perturbations. Nous proposons également une classification de ces techniques, puis nous discutons les particularités de chacune d'elles.

Les DTNs possèdent certaines caractéristiques qui leurs sont propres et qui sont différentes de celles des réseaux ad hoc mobiles classiques. Ces particularités incluent principalement des connectivités intermittentes, des temps de livraison relativement larges, des ressources limitées aux nœuds. De ce fait, la plupart des études qui se sont intéressées aux techniques de routage dans les DTNs ont considéré principalement comme critère de performance la maximisation du taux de livraison des paquets, la minimisation du temps moyen de délai de livraison, la minimisation de la consommation des ressources utiles (énergie, capacité de stockage,. . .), ou bien un compromis entre ces différents critères.

D'autre part, vu que ces techniques s'appuient sur la mobilité des nœuds intermédiaires afin de pouvoir diffuser les paquets, leurs performances dépendent largement des caractéristiques du modèle de mobilité de ces nœuds. Ce dernier définit le temps qui sépare un contact entre deux nœuds, la séquence de ces contacts ainsi que la durée moyenne d'un contact. Les caractéristiques statistiques du temps séparant deux contacts entre un nœud mobile et un nœud fixe fera l'objet du Chapitre 2.

Dans la littérature publiée autour des DTNs, deux études se sont penchées sur les spécificités des différentes approches du routage dans les DTNs [60, 108]. Dans ce chapitre, nous décrivons et classifions ces techniques selon le degré de connaissance qu'un nœud dispose concernant ses futurs contacts potentiels avec d'autres nœuds dans le réseau. En particulier, selon que ces futurs contacts sont planifiés, contrôlés, pronostiqués ou bien opportunistes, on a pu identifier quatre classes différentes de telles techniques. Nous décrivons alors brièvement chacune de ces classes.

Dans le cas où les futurs contacts sont planifiés, il y a trois informations importantes qu'un nœud devrait acquérir afin de pouvoir choisir le chemin optimal pour le transfert de paquets [55]. Ces informations concernent les temps d'occurrence de ces prochains contacts, la taille de la fille d'attente de l'autre nœud lors d'un contact, et la charge du trafic dans le réseau. La connaissance complète de ces trois informations permet à un nœud de prévoir au mieux la dynamique du réseau et ainsi d'améliorer la performance du protocole de routage. Pourtant, en pratique, les nœuds disposent en général seulement d'une ou de deux de ces informations.

Les techniques qui sont basées sur des contacts contrôlés ont recours en général à des nœuds mobiles dédiés dont le seul but est d'assurer un grand nombre des opportunités de transfert entre les nœuds mobiles du réseau, et ainsi, d'améliorer la performance générale du réseau [56, 110, 111]. Ces noeuds additionnels ont généralement des déplacements contrôlés. Leur trajet pourrait suivre un chemin fixe et prédéterminé qui a été sélectionné afin de permettre un grand nombre possible de contacts avec les autres nœuds. D'autre part, ce trajet pourrait être ajusté dynamiquement afin de l'adapter aux flux du trafic dans le réseau. Les techniques basées sur les contacts contrôlés permettent de limiter le temps de livraison des paquets ainsi que de borner d'autres métriques de performance dans le réseau.

Les techniques basées sur des contacts pronostiqués essayent d'exploiter certaines connaissances utiles concernant le modèle de mobilité ou bien certains aspects de mouvements répétitifs des nœuds [33, 92, 70]. En se basant sur une estimation de cette connaissance, la station va décider si elle va transférer le paquet qu'elle détient ou bien le garder pour une meilleure opportunité. Chaque station se fait attribuer un ensemble des métriques représentant sa vraisemblance à délivrer le paquet à sa destination. Quand un nœud portant un paquet rencontre un autre nœud qui possède de meilleures métriques pour la destination, le premier nœud va transférer au deuxième son paquet augmentant ainsi la probabilité de livraison du paquet. Selon la nature des connaissances pronostiquées, on a pu subdiviser cette classe en deux autres sous-classes selon que la connaissance estimée est basée sur des caractéristiques particulières de la mobilité des nœuds ou bien sur l'historique de leurs anciens transferts ou contacts.

Enfin, les techniques basées sur des contacts opportunistes sont généralement caractérisées par des contacts aléatoires entre les nœuds, suivis par un échange de données entre les nœuds concernés [100, 94, 41]. Etant donné que sous de telles conditions, une rencontre entre deux nœuds, et par conséquent l'échange des paquets, ne pourrait pas être prévue à l'avance, les techniques de routage correspondantes ont recours à la diffusion de plusieurs copies du paquet dans le réseau afin d'augmenter les chances de livraison via l'une de ces copies. On a pu également subdiviser cette classe en deux autres sous-classes selon que la technique de diffusion est basée sur un mécanisme épidémique ou bien sur un codage du paquet de la source.

A.3 Chapitre 2 : Caractérisation du temps séparant deux contacts entre un noeud mobile et un noeud fixe

Dans ce chapitre, on s'intéresse aux caractéristiques statistiques des intervalles de temps séparant deux contacts consécutifs dans un réseau à deux dimensions contenant un nœud mobile et un ou N nœuds fixes. Le nœud mobile se déplace dans le réseau selon un certain modèle de mobilité aléatoire (e.g. modèle *Radom Waypoint* ou bien modèle *Radom Direction*). Les oeuds fixes, nommés boîtes (en anglais "*throwboxes*"), servent comme des relais fixes dans le réseau : un nœud mobile qui passe à portée d'une boîte va lui passer une copie de son paquet ; dans un temps ultérieur, lorsqu'un autre nœud mobile passe à portée de la boîte, la boîte va lui passer le paquet ou bien une copie (cela dépend de la spécificité du protocole de routage). Les boîtes sont supposées être dispersées dans le réseau selon une certaine densité de probabilité aléatoire.

Face à ce scénario, nous avons procédé par simulations afin de pouvoir analyser la densité de probabilité des intervalles de temps séparant deux contacts consécutifs entre le mobile et une boîte dans un premier temps, puis entre le mobile et N boîtes dans un deuxième temps. A travers des simulations approfondies, on a pu montrer que les densités de probabilité de ces deux variables peuvent être approchées par des lois exponentielles. De plus, le paramètre de la loi exponentielle entre le mobile et les N boîtes équivaut à N fois celui de la loi exponentielle entre le mobile et une boîte. Cela implique que le temps de rencontre entre un mobile et N boîtes est le minimum de N variables aléatoires exponentielles ayant le même paramètre. Notons que cette observation reste valide dans la limite où le réseau demeure épars, par exemple en raison d'une faible densité des nœuds ou bien de portées de transmission des nœuds relativement petites par rapport aux dimensions du réseau.

Ensuite, nous avons procédé au calcul du paramètre de la loi exponentielle correspondant à un mobile et à une boîte. Nous avons obtenu une approximation précise de ce paramètre, qu'on note par μ, et qui dépend des paramètres des nœuds et du réseau, ainsi que de la densité de probabilité stationnaire spatiale du modèle de mobilité considéré, et de la densité de probabilité spatiale des boîtes dans le réseau. Nous avons obtenu également une expression de la durée du temps de contact d'un mobile et d'une boîte.

Les expressions obtenues pour le paramètre μ, dont les exactitudes ont été validées par des simulations, permettent de constater les éléments suivants : Dans le cas où les nœuds mobiles se déplacent selon le modèle de mobilité *Random Waypoint*, nous avons intérêt à répartir les boîtes vers le centre de la surface du réseau afin d'augmenter le taux de rencontre entre les nœuds mobiles. Cette répartition pourrait être obtenue en utilisant par exemple la densité de probabilité stationnaire spatiale de ce même modèle de mobilité. Dans le cas où les nœuds se déplacent selon le modèle de mobilité *Random Direction*, l'expression du paramètre μ est indépendante de la répartition spatiale des boîtes dans le réseau. Cela est en partie peut être expliqué par la densité de probabilité stationnaire spatiale de ce modèle de mobilité qui est uniforme, impliquant qu'à un instant aléatoire t, un nœud mobile pourrait se trouver aléatoirement à n'importe quel endroit du réseau.

Les résultats obtenus dans ce chapitre seront utilisés dans les deux chapitres suivants.

A.4 Chapitre 3 : Evaluation des performances des protocoles de relais à deux sauts et de routage épidémique dans les DTNs incluant des boîtes

Ce chapitre porte sur l'analyse des performances de deux protocoles de routage dans les réseaux mobiles ad hoc épars incluant des boîtes. Rappelons qu'une boîte est une station fixe qui joue le rôle d'un relais stationnaire pour les nœuds mobiles et dont le but est d'augmenter le nombre de contacts entre ces nœuds. Le modèle de réseau considéré est le suivant. Les nœuds mobiles bougent indépendamment les uns des autres dans le réseau selon un modèle de mobilité aléatoire (Random Wapoint model ou bien Random Direction model). Les boîtes sont réparties dans le réseau selon une certaine densité de probabilité aléatoire. Nous considérons deux cas où les boîtes sont totalement déconnectées les unes des autres ou bien totalement connectées (sous forme d'une maille). Tous les nœuds sont identiques (i.e. ont la même portée de transmission) et les interférences entre les nœuds sont ignorées. Le réseau fonctionne dans des conditions de connectivité intermittente (conséquence d'une faible densité des nœuds, d'une grande mobilité, etc.), et le routage des paquets s'effectue selon l'approche de "*store-carry-and-forward*" (i.e. retenir-porter-transférer) pour les nœuds mobiles et selon l'approche "*store-and-forward*" (i.e. retenir-transférer) pour les boîtes.

En considérant ce scénario, nous nous sommes intéressés à évaluer les performances de deux protocoles de routage fréquemment considérés dans les DTNs : le protocole de relais à deux sauts [41] (Multicopy two-hop relay protocol (MTR)) et le protocole de routage épidémique [100] (Epidemic routing (ER)). Dans le MTR, quand une station source rencontre une autre station mobile (destination incluse) ou bien une boîte, elle lui passe une copie de son paquet ; quand une station mobile ou bien une boîte portant une copie du paquet rencontre la destination ultérieurement, elle lui livre sa copie. Ainsi, avec le protocole MTR le paquet peut faire au maximum deux sauts entre la source et la destination, d'où le nom du protocole. Le protocole de routage épidémique est un protocole d'inondation de paquets ou chaque station, qu'elle soit fixe ou mobile, qui porte une copie du paquet, transmettra une copie à toute autre station non porteuse qu'elle rencontre sur son trajet[3].

Les performances de ces deux protocoles sont examinées en termes de temps moyen de délai de livraison d'un paquet et de nombre moyen de copies engendrées sur les nœuds

[3]Une station qui porte une copie d'un paquet est dénommée station infectée par analogie à la diffusion d'une épidémie. Une station non porteuse d'une copie est dénommée non-infectée.

mobiles et les boîtes au moment de la livraison d'un paquet. Le délai de livraison est défini comme étant le premier instant ultérieur à $t = 0$ où la destination reçoit le paquet ou bien l'une de ses copies – $t = 0$ correspond au temps de génération du paquet à la source. Le nombre de copies engendrées sur les différents nœuds est lié à l'énergie nécessaire pour livrer le paquet à la destination et sera considéré comme une référence de la charge induite par le protocole de routage.

En utilisant les résultats du Chapitre 2 dans le contexte du modèle de réseau considéré, nous avons approché les intervalles de temps séparant deux contacts consécutifs entre un mobile et une boîte par une loi exponentielle avec un paramètre μ. De même, en utilisant les résultats soulignés en [41], nous avons approché également les intervalles de temps séparant deux contacts consécutifs entre deux mobiles par une autre loi exponentielle ayant un paramètre λ. Les valeurs des deux paramètres μ et λ peuvent être obtenues à travers les formules dérivées dans 2 et [41].

En se reposant sur les lois exponentielles des intervalles de temps écoulés entre les contacts, nous avons proposé un modèle stochastique basé sur une chaîne de Markov absorbante à temps continu et à nombre d'états fini. L'état du système à un instant t, défini comme ayant deux dimensions, spécifie le nombre de nœuds mobiles et de boîtes porteurs d'une copie de paquet à l'instant t.

A travers ce modèle, nous avons dérivé des expressions pour les fonctions de répartition du délai de livraison et du nombre de copies engendrées dans le réseau au moment de la livraison. Ces expressions sont obtenues pour chacun des deux protocoles dans les deux cas où les boîtes sont totalement déconnectées ou connectées. En utilisant des équations différentielles, nous avons obtenues également des résultats asymptotiques pour les espérances de délai de livraison et de nombre de copies engendrées sous le protocole MTR quand le nombre de nœuds mobiles et/ou de boîtes devient grand.

Tous les résultats théoriques obtenus ont été validé par des simulations. Ces résultats ont montré qu'il existe un grand potentiel pour le déploiement des boîtes lorsqu'elles sont connectées ou bien lorsqu'elles opèrent d'une façon épidémique.

A.5 Chapitre 4 : Une plate-forme pour l'évaluation des performances des réseaux DTNs incluants des boîtes

Le but de ce chapitre est de fournir des indications sur la conception et l'évaluation des techniques de routage dans les réseaux opportunistes DTNs incluant des boîtes. On reprend le même modèle de réseau que dans le Chapitre 3, i.e. on considère un réseau constitué de $N+1$ nœuds mobiles et M boîtes. Le temps de rencontre entre deux nœuds mobiles suit une loi exponentielle avec un paramètre λ, et celui entre un nœud mobile et une boîte suit une autre loi exponentielle avec un paramètre μ. Les nœuds mobiles se déplacent aléatoirement dans le réseau et les contacts entre les différents nœuds sont par conséquent opportunistes. Dans ce contexte, nous nous sommes intéressés au cas où il y a seul paquet dans le réseau qui a été généré par une station mobile qu'on note la source. Ce paquet est destiné à être livré à une autre station mobile qu'on note la destination. Le reste des nœuds mobiles ($N-1$ noeuds) avec les boîtes se comportent comme des relais intermédiaires.

Par rapport à ce réseau, notre objectif est de proposer et d'évaluer des techniques de routage à plusieurs copies qui exploitent la présence des boîtes pour accélérer la dissémination du paquet dans le réseau afin d'en améliorer la performance. Comme critères de performance, nous optons pour minimiser le délai de livraison des paquets et le nombre de transmissions de copies à effectuer par les nœuds mobiles seulement. Les boîtes, qui sont des nœuds extérieurs au réseau et qui ont été ajoutées pour augmenter sa performance, sont supposées posséder de meilleures performances en termes d'énergie et de capacité de stockage que les nœuds mobiles. Ainsi, le but d'une technique de routage est d'exploiter les capacités des boîtes tout en limitant la consommation des ressources sur les nœuds mobiles.

La procédure de conception et d'évaluation est organisée en trois étapes consécutives.

Tout d'abord, nous avons décrit les différents défis auxquels il s'agit de faire face lors de la conception d'un tel réseau. Plus précisément, en illustrant divers aspects liés à l'utilisation de ce type de réseau, nous avons exhibé les différents critères de performance exigés par des applications conçues pour les DTNs.

Dans la deuxième étape, nous avons présenté une plate-forme stochastique qui permet d'évaluer un ensemble de métriques de performances pour un grand nombre de protocoles de routage opérants dans un réseau similaire sans avoir besoin de réaliser des simulations. Cette plate-forme est basée sur une chaîne de Markov absorbante, dont l'état du système est décrit par trois dimensions. La première dimension spécifie le nombre de nœuds mobiles, source incluse, qui ont déjà reçus leurs copies directement de la part de la source. La deuxième dimension spécifie le nombre de nœuds mobiles qui ont déjà reçus leurs copies émanant d'une boîte porteuse d'une copie. Et la troisième dimension spécifie tout simplement le nombre de boîtes porteuses d'une copie.

La plate-forme repose sur des équations récursives et nécessite comme entrées les paramètres N, M, λ, μ, ainsi que le générateur $Q(t)$ de la chaîne de Markov sous-jacente correspondant à la technique de routage considérée. Avec cette plate-forme, l'espérance de délai de livraison du paquet ainsi que l'espérance de nombre de nœuds mobiles porteurs de copies se transposent en des variables aléatoires dont les valeurs peuvent être obtenues facilement à travers la plate-forme.

En troisième étape, nous avons proposé cinq techniques de routage pour les réseaux DTNs incluant des boîtes. Chaque technique de routage représente une combinaison des possibilités de transfert des copies entre la source, la destination, les boîtes et les autres nœuds mobiles. En commun pour ces cinq techniques, la source peut passer une copie de son paquet à la destination ou bien à une boîte. De plus, une boîte "infectée" passera une copie à tout autre nœud mobile qu'elle rencontra, ou bien à la destination. Par ailleurs, ces techniques se différencient par les restrictions imposées à un nœud mobile infecté (possibilité ou non de passer une copie à une boîte ou bien à la destination).

A travers la plate-forme que nous avons introduite, nous avons pu comparer analytiquement l'espérance du délai de livraison du paquet ainsi que l'espérance du nombre des transmissions effectuées par les nœuds mobiles à travers chacune des techniques envisagées. Les résultats obtenus montrent qu'il n'y a pas une technique unique qui achève les meilleures performances par rapport aux deux mesures considérées. Chaque technique représente un compromis entre un délai de livraison court et un nombre de transmissions faible. Cependant, le fait de se servir également des interactions entre la source et les nœuds mobiles, ou bien entre les nœuds mobiles et la destination réduit considérablement le délai de livraison du paquet.

A.6 Chapitre 5 : Le diamètre du protocole de routage épidémique

Dans ce chapitre, nous nous sommes intéressés à analyser la longueur moyenne en termes de nombre de sauts du chemin séparant la source de la destination dans un réseau opportuniste DTN utilisant le protocole de routage épidémique. Cette métrique est nommée le diamètre du protocole de routage épidémique, et elle est définie comme le nombre moyen de sauts à effectuer par un paquet entre la source et la destination sous ce protocole.

L'objectif de cette étude est de caractériser analytiquement le diamètre du protocole épidémique, et donc, d'examiner les conséquences d'une limitation du nombre de sauts sur la performance d'un protocole de diffusion épidémique des paquets. En particulier, nous avons considéré un réseau DTN constitué de $N+1$ nœuds mobiles qui se déplacent selon un certain modèle de mobilité aléatoire (Random Waypoint model ou bien Random Direction model). Sous ce modèle de mobilité, les laps de temps écoulés entre deux rencontres consécutives de deux nœuds sont supposés suivre une loi exponentielle avec un paramètre λ [41]. Les nœuds utilisent le protocole de routage épidémique, et le but est d'acheminer un paquet généré par un certain nœud source à un autre nœud destination.

Pour réaliser cette analyse, nous avons établi une analogie entre l'évolution temporelle du processus d'infection des nœuds sous le protocole épidémique et l'évolution temporelle d'un arbre récursive d'ordre $N+1$. Un arbre récursive d'ordre N, noté par RT_n, est définie comme un arbre ordinaire sans cycle établi sur un graphe connecté. Par ailleurs dans un arbre récursif, les sommets sont étiquetés dans l'ordre croissant de leurs instants d'adhésion à l'arbre. Plus précisément, un arbre récursive d'ordre N, RT_n, évolue d'un arbre récursive d'ordre $N-1$, RT_{n-1}, en joignant aléatoirement le nouveau sommet étiqueté n à un autre sommet de l'arbre RT_{n-1}. Par conséquent, en partant de la racine de l'arbre, les étiquettes des sommets forment une suite croissante sur tous les chemins reliant la racine à tout autre sommet [77, 34].

En transposant ceci à notre modèle de réseau, la longueur moyenne du chemin de livraison sous le protocole épidémique correspondra donc à la distance moyenne entre la racine d'un arbre récursive d'ordre $N+1$ et un sommet v qui joindra l'arbre aléatoirement parmi l'ensemble des autres N sommets. La racine de l'arbre est associée à la station source dans notre réseau tandis que le sommet v est associé à la station destination. Sachant que la valeur moyenne de la distance séparant deux sommets aléatoires i et j est connue, le diamètre du protocole épidémique est obtenue en fixant le sommet i comme la racine de l'arbre et en conditionnant les différentes valeurs possibles du sommet j. L'expression obtenue montre que le diamètre dépend seulement du nombre des stations actives N. Il est indépendant des paramètres du réseau et du modèle de mobilité des stations. De plus, lorsque le nombre de stations actives N devient grand, le diamètre peut être bien approximé par une fonction logarithmique.

Ce résultat montre que le diamètre est relativement petit par rapport aux tailles des réseaux et varie généralement entre 2 et 4.5 quand le nombre des stations N varie entre 10 et 100. Cette observation s'accorde avec une autre étude [31] qui a analysé empiriquement le diamètre à partir des traces réelles d'un modèle de mobilité humain. A travers des simulations effectuées sur différents scenarios, nous avons validé l'expression analytique du diamètre.

Nous avons aussi étudié l'effet d'une limitation du nombre de saut qu'une copie peut effectuer sous le protocole épidémique et ses conséquences sur le délai et le nombre de copies générées dans le réseau. En particulier, nous nous sommes intéressés au cas où cette limite est égale au diamètre du réseau. En utilisant la plateforme développée au chapitre 4, nous avons comparé numériquement l'impact de varier le maximum nombre possible des sauts sur les performances du réseau. Les résultats obtenus ont montré qu'en limitant le maximum nombre de sauts, le délai moyen augmente tandis que le nombre des copies générées diminue inversement.

A.7 Chapitre 6 : Conception et analyse d'un algorithme de backoff pour le standard IEEE 802.11

Dans ce chapitre, nous proposons et évaluons un algorithme de backoff dynamique dont le but est d'améliorer les performances de la fonction de coordination distribuée (Distributed Coordination Function (DCF)) du standard IEEE 802.11 [2] dans une multitude de scenarios. En particulier, dans le cas où le réseau est congestionné. Notre objectif est de maximiser le débit du système tout en minimisant parallèlement le délai moyen de livraison des paquets. Par ailleurs, nous choisissons aussi de garder notre algorithme simple et le plus proche possible du standard.

Pour cela, nous nous sommes basés tout d'abord sur un modèle markovien qui a été développé par Bianchi [14]. Dans ce modèle, l'auteur a modélisé l'évolution de la fenêtre de contention de l'algorithme backoff du standard correspondant à la transmission d'un paquet par un nœud. Le modèle de réseau considéré, et celui qu'on a adopté ensuite, est le suivant. Le réseau est constitué de N nœuds fixes, tous situés dans la même portée de transmission l'un de l'autre (le réseau forme une seule cellule). Chaque nœud est une source de trafic pour une connexion, et la destination pour une autre connexion. Les nœuds fonctionnent dans des conditions de saturation, i.e. chacun des nœuds a toujours des paquets prêts à être envoyés. L'état du canal est supposé idéal (pas de bruit sur le canal), et le problème du nœud caché est supposé inexistant. En outre, la probabilité de collision est supposée constante pour chacun des nœuds et indépendante du passé.

En utilisant ce modèle, l'auteur a fourni des expressions pour la probabilité de transmission d'un paquet sur le canal τ, la probabilité de collision P_c ainsi que la probabilité de transmission optimale τ^* qui maximise le débit. Ces expressions sont fonctions, entre autres, du nombre de stations actives N ainsi que de la valeur minimale par défaut de la fenêtre de contention du standard CW_0. A travers une analyse que nous avons menée, nous avons montré que la probabilité de transmission optimale τ^* correspond aussi à la probabilité de transmission qui minimise le nombre moyen de collisions sur le canal. Par conséquent, le délai moyen de bout-en-bout observé au niveau de la couche transport sera aussi minimisé avec cette probabilité de transmission.

En reprenant ces équations, nous avons dérivé une expression reliant la valeur minimale de la fenêtre de contention, notée par CW_{min}, à la probabilité optimale de transmission τ^*. La valeur CW_{min} traduit la valeur initiale de la fenêtre de contention que la station devrait adopter afin de transmettre à la probabilité optimale τ^*. Par contre, le calcul de la valeur de τ^* nécessite l'estimation du nombre de stations actives N. D'où le besoin d'avoir un estimateur en temps réel de ce nombre noté par $\overline{N}$.

Pour estimer le nombre des stations actives autour d'elle, chaque station va procéder à une mesure directe en comptant comme active chaque station pour laquelle elle reçoit un signe de vie (i.e. un paquet de donnée ou de contrôle) durant un certain intervalle du temps. L'intervalle de mesure sera le laps de temps séparant deux transmissions avec succès sur le canal pour la station "mesurante". Vu la variabilité de ces intervalles de temps, les échantillons d'observation, noté par $\widehat{N}_n$, ne peuvent pas être utilisés directement comme un bon estimateur de N. D'autre part, en observant que la taille de ces intervalles est bien reflétée par la valeur de la fenêtre de contention au moment de la transmission du paquet, nous avons utilisé cette valeur comme un poids pour pondérer ces échantillons. A travers des simulations, nous avons établie qu'une petite valeur de l'ordre de filtre est suffisante pour garantir une bonne réactivité du filtre. En revanche, les échantillons mesurés ne tiennent pas compte des stations dont les transmissions ont été perdues à cause d'une collision durant les intervalles de mesure. Par conséquent, ces mesures sous-estiment la valeur réelle de N et nécessitent une correction qui s'avère être linéaire.

Ainsi, le fonctionnement de l'algorithme backoff proposé est le suivant. Si la transmission résulte en une collision, la station double la taille de sa fenêtre de contention comme dans le cas du standard. Si la transmission s'avère réussi, la station remet la valeur minimale de sa fenêtre CW_{min} à une valeur optimale qui tient compte du niveau de congestion dans le réseau à travers une estimation en temps réel du nombre de stations actives.

Nous avons implémenté notre algorithme en `ns-2` [3] et nous avons comparé ses performances, notamment le débit et le délai, avec celles obtenues par le standard et deux autres protocoles qui ont été proposés. Dans tous les scenarios considérés, notre protocole a présenté de meilleures performances par rapport au standard et aux deux protocoles ainsi qu'une meilleure réactivité par rapport à la variation du nombre de stations actives. En utilisant l'index de Jain [53], nous avons aussi évalué l'équité d'accès au canal résultant avec notre protocole et nous avons montré que notre algorithme permet également d'avoir une meilleure équité d'accès par rapport au standard.

List of Acronyms

ACK	Acknowledgment
ACF	Autocorrelation Function
BEB	Binary Exponential Backoff
CCDF	Complementary Cumulative Distribution Function
CDF	Cumulative Distribution Function
CSMA/CA	Carrier Sense Multiple Access/Collision Avoidance
CSMA/CD	Carrier Sense Multiple Access/Collision Detection
CTS	Clear To Send
CW	Contention Window Size
DCF	Distributed Coordination Function
DTN	Disruption/Delay Tolerant Network
et al.	and others
e.g.	for example
ER	Epidemic Routing
i.e.	id est
i.i.d.	independent and identically distributed
MAC	Medium Access Control
MANET	Mobile Ad Hoc Network
MTR	Multicopies Two-hop Relay
NAV	Network Allocation Vector
NS-2	Network Simulator 2
ODE	Ordinary Differential Equation
REF	Random Direction with Reflection
RD	Random Direction
RTS	Request To Send
r.v.	random variable
RWP	Random Waypoint
VANET	Vehicular Ad hoc Network
WLAN	Wireless local area network

Résumé

Cette thèse s'intéresse à la conception et l'évaluation des protocoles de routage et d'accès au canal pour les réseaux sans fils. La première partie de la thèse focalise principalement sur l'évaluation de protocoles de routage dans les réseaux tolérants aux perturbations quand ces réseaux incluent des noeuds relais fixes, appelées boîtes. Les réseaux tolérants aux perturbations sont des réseaux ad hoc sans fil mobiles qui sont caractérisés par de fortes partitions et où le routage s'effectue généralement selon le paradigme "store-carry-forward". Les boîtes servent de relais pour stocker et transmettre les paquets aux noeuds mobiles du réseau.

Dans un premier temps, nous avons dressé l'état-de-l'art des techniques de routage dans ces réseaux. Puis nous avons montré au travers des campagnes de simulations que les instants successifs de rencontre entre une boîte et un noeud mobile qui se déplace selon un modèle de mobilité aléatoire, sont bien approximés par un processus de Poisson. Nous donnons une formule explicite approchée pour l'intensité de ce processus de Poisson qui dépend notamment de la densité de probabilité spatiale du modèle de mobilité considérée, ainsi que de la densité de probabilité spatiale des boîtes. Dans un deuxième temps, nous étudions l'impact d'ajouter des boîtes sur les performances de deux protocoles de routage classiques, le protocole épidémique et le protocole de routage à deux sauts. Nous développons des expressions explicites pour quantifier la distribution et la moyenne du délai de livraison d'un paquet, ainsi que le nombre des copies générées lors de cette transmission. Finalement, nous proposons cinq stratégies qui s'appuient essentiellement sur la présence des boîtes pour réaliser le routage des copies. Par ailleurs, nous avons introduit une plateforme basée sur un modèle markovien qui permet de calculer et de comparer analytiquement les diverses métriques de performance pour ces cinq stratégies. Cette plateforme s'avère être un outil efficace pour l'évaluation des techniques de routage destinées aux réseaux sans fil mobiles incluant des boîtes.

Dans la deuxième partie de la thèse, nous nous sommes intéressés à l'algorithme de "backoff" du standard IEEE 802.11 pour les communications radio. Nous avons proposé une extension de cet algorithme dont l'objectif est d'améliorer ses performances dans le cas où le réseau possède un grand nombre d'utilisateurs. Cette variante du protocole standard est basée principalement sur un estimateur en temps-réel du nombre de stations actives. L'estimation s'effectue en mesurant "les signaux de vie" (i.e. les paquets de données et de contrôle) émis par ces stations durant un certain intervalle de temps. Au travers des simulations en NS-2, nous avons montré que notre algorithme donne de meilleures performances (débit efficace) que la version standard du protocole et ceci dans un grand nombre de scenarios.

Mots-clés — Evaluation de performance, Réseaux Ad hoc mobiles, Chaîne de Markov, Processus stochastiques, Modèles de mobilité, IEEE 802.11, Protocole de routage épidémique, Réseaux tolérants aux perturbations.

Abstract

This thesis concerns with the design and the evaluation of routing and medium access control protocols for wireless networks. The first part of the thesis focuses on the performance evaluation of routing protocols for disruption tolerant networks (DTNs) when these networks are augmented by throwboxes. Disruption tolerant networks are mobile ad hoc networks which are mainly characterized by lack of connectivity and where routing is generally handled according to the store-carry-and-forward paradigm. Throwboxes are fixed relays that are used to store and forward the packets to the mobile nodes.

As a first step, we review the current state of art of routing approaches in delay tolerant networks. Next, we show through extensive simulations that successive meeting instants between a throwbox and a mobile node moving according to usual random mobility model can be well modeled by a Poisson process. We provide an explicit formula for the rate of the Poisson process, and we show that this rate depends, among others, on the stationary spatial probability density function of the mobility model as well as on the probability density function of throwboxes spatial distribution. As a second step, we study the impact of adding throwboxes on the performance of two common routing protocols, namely the epidemic and the multicopy two-hop relay protocols. We develop explicit expressions for the distribution and the mean of the delivery time of a packet, as well as for the number of generated copies. Following that, we propose five routing strategies that rely on the presence of throwboxes to perform packet forwarding. We therefore introduce an analytical framework based on a markovian model to compute and to compare analytically diverse performance metrics of these strategies. The proposed framework shows to be a useful tool to evaluate the performance of routing techniques destined throwboxes augmented DTNs.

In the second part of the thesis, we deal with the design and the evaluation of an adaptive mechanism to enhance the performance of the backoff algorithm of the IEEE 802.11 standard. We propose an extension of the backoff algorithm where the objective is to enhance the throughput and the delay performances in the case of high number of active users. This variant is mainly based on a real time estimation of the number of active stations around. The estimation is used to feedback a mechanism that sets the starting size of the contention window to an optimal value corresponding to the given network size. Through simulations conducted in NS-2, we show that our algorithm outperforms the standard one in different scenarios.

Keywords — Performance evaluation, Mobile Ad hoc networks, Markov chain, Stochastic processes, Mobility models, IEEE 802.11, Epidemic routing protocol, Disruption tolerant networks.