UNIVERSITÀ DEGLI STUDI DI PARMA

Dottorato di Ricerca in Tecnologie dell'Informazione

XXVIII Ciclo

Efficient Multi-hop Broadcast Data Dissemination in IoT and Smart Cities

Coordinatore:

Chiar.mo Prof. Marco Locatelli

Relatore:

Chiar.mo Prof. Gianluigi Ferrari

Dottorando: *Andrea Gorrieri*

Gennaio 2016

To my family and to Agnese

Contents

List of Figures

List of Tables

Introduction

The Internet of Things (IoT) can be defined as a "*network of networks*" composed by billions of uniquely identified nodes often denoted as Smart Objects (SO). SOs can be items equipped with sensors, consumer devices (e.g., smartphones, tablets, or wearable devices), and enterprise assets which have to be connected both to the Internet and to each other. The birth of the IoT, with its communication paradigms, can be considered as an enabling factor for the creation of the so-called Smart Cities. In Smart Cities, Information and Communication Technologies (ICT) are used to improve the quality, performance and interactivity of urban services, ranging from traffic management and pollution monitoring to government services and energy management.

One of the major challenges in order to build applications in IoT and Smart Cities scenarios is related to the huge number of involved devices (forecasts say that there will be 50 billion connected IoT devices by 2020). Moreover, the extreme degree of node spatial density, together with the high level of heterogeneity of the used communications technologies, lead to an high level of interference. Another significant problem is that IoT nodes are typically constrained in terms of memory, computational capabilities and battery life. In the last years, the scientific community has investigated these problems by creating new communication standards suited for the new scenarios and by integrating and improving already existing solutions.

The main focus of this thesis is related to the efficient multi-hop data dissemination in Smart Cities and IoT environments. In particular, great importance is given to the so-called Mobile Ad-hoc NETworks (MANETs). The main characteristic of

MANETs is that each node may forward data packets associated with multi-hop communications between other pairs of nodes. In other worlds, each node can act as source, router, and destination at the same time. In these networks nodes can move with a high mobility level making the design of routing protocols for MANETs very difficult. For these reasons, routing protocols in MANETs are different with respect to those designed for classical infrastructure-based networks. In particular, they typically rely on control messages which must be propagated from one source to all other nodes in the network via multi-hop broadcasting. This is done for different purposes, for example to send routing tables updates or to search a multi-hop path toward a desired destination node.

A subclass of MANETs which plays a key role in Smart Cities is represented by the so-called Vehicular Ad-hoc NETworks (VANETs). Like MANETs, VANETs do not have fixed infrastructure and nodes which compose the ad-hoc network are represented by vehicles. Due to the different level of mobility, VANETs exhibit characteristics that can be really different from generic MANETs. Moreover, in this kind of scenario, not only Vehicle-to-Vehicle (V2V) communications are considered. In fact, the presence of a so-called Road Side Unit (RSU) allows the support of Vehicle-to-Infrastructure (V2I) communications. Multi-hop data dissemination plays an important role also in this case, as one of the main application areas in VANETs is related to the efficient dissemination of traffic warning information. For example, through multi-hop broadcasting, drivers can be promptly informed about traffic congestions or vehicle collisions.

The high level of heterogeneity of the used communication technologies, together with the large number of involved devices, makes the adoption of classical multi-hop broadcast techniques (like flooding) unacceptable in IoT and Smart Cities scenarios. In fact, flooding generates a large level of redundancy which decreases the efficiency of the multi-hop propagation due to collisions. For these reasons, multi-hop broadcasting should be improved by making a good use of the available bandwidth, saving retransmissions and selecting the best rebroadcasters.

In this thesis, innovative multi-hop dissemination techniques are presented. Mostly based on probabilistic forwarding, these techniques have been used for different pur-

poses: from the improvement of the performance of unicast protocols for MANETs to the efficient data dissemination within VANETs.

Thesis Structure

In Chapter 1, some of the major problems related to multi-hop communications in IoT and Smart Cities scenarios are described. First, in order to understand how interference can influence the performance of wireless multi-hop protocols, a coexistence analysis between IEEE 802.15.4 and IEEE 802.11 technologies is performed. Then, a survey on the main routing protocols for ad-hoc networks is provided. Last, some of the major problems related to multi-hop routing in ad-hoc networks, together with possible solutions proposed in the literature, are described.

In Chapter 2, an innovative unicast reactive routing protocol for MANETs, denoted as irresponsible Ad-Hoc On-demand Distance Vector (iAODV) is presented. The iAODV protocol is obtained starting from one of the most sought routing protocol for MANETs, i.e., the Ad-Hoc On-demand Distance Vector (AODV) protocol. In particular, iAODV is obtained by replacing, in the route discovery process of the AODV protocol, the flooding mechanism with an innovative probabilistic forwarding technique denoted as Irresponsible Forwarding (IF). The performance of the iAODV protocol is analyzed both analytically and trough simulations considering different kinds of scenarios.

In Chapter 3, the concept of Vehicular Sensor Network is introduced. VSNs are a subclass of VANETs in which vehicles continuously gather, process, and share location-relevant sensor data (e.g., road conditions, pollution, etc...). Information collection and dissemination can be performed using inter-vehicular communications and/or relying on the presence of roadside infrastructure. An innovative scheme for data acquisition in clustered VSNs is presented. In particular, the proposed scheme foresees a two-phase communication mechanism. First of all, a *downlink phase* is triggered by a RSU to form a clustered topology. The so-formed clustered VSN is then used, during the (second) *uplink phase*, for data aggregation.

In Chapter 4, a novel probabilistic broadcasting strategy, denoted as Distance-

based Silencing Irresponsible Forwarding (DiSIF) is proposed. This strategy ”stems” from the already cited IF technique and tries to address its inefficiencies through a novel silencing mechanism which effectively selects rebroadcasters by introducing an initial contention phase. In order to investigate the performance of DiSIF, a simulative analysis has been carried out, moreover, analytical considerations are provided.

Chapter 1

Background and Related work

1.1 Introduction

Internet of Things (IoT) can be defined as a "*network of networks*" composed by billions of uniquely identified physical Smart Objects (SO), organized in an Internet-like structure. Smart Objects can be items equipped with sensors, consumer devices (e.g., smartphones, tablets, or wearable devices), and enterprise assets that are connected both to the Internet and to each others. The birth of the IoT, with its communications paradigms, can be considered as an enabling factor for the creation of the so-called Smart Cities. A Smart City uses Information and Communication Technologies (ICT) to enhance quality, performance and interactivity of urban services, ranging from traffic management and pollution monitoring to government services and energy management.

The major challenges in order to build applications in IoT and Smart Cities scenarios are the following: (i) the involved communication technologies are strongly heterogeneous and can interfere with each other; (ii) the huge number of involved devices create serious problems of deployment costs and depletion of the available bandwidth; (iii) IoT nodes are typically constrained in terms of memory, computational capabilities and battery life. In these years, the scientific community has tackled these problems by creating new communication standards suited for the new sce-

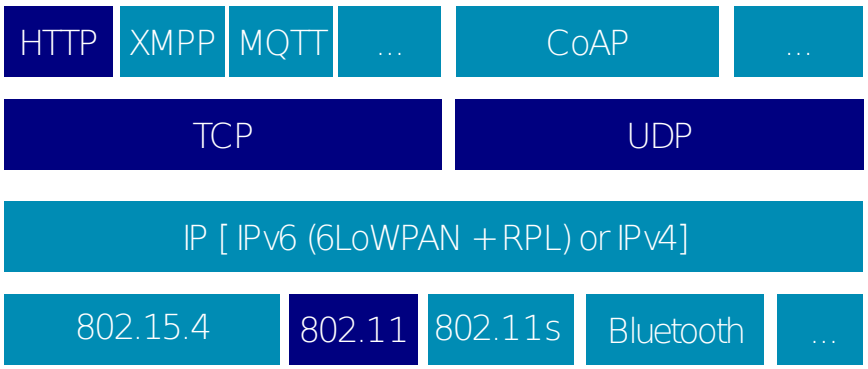


Figure 1.1: IoT protocol stack. In light blue the new protocol proposed in order to cope with the typical problems of constrained nodes.

narios and by integrating and improving already existing solutions. In particular, in Figure 1.1 a typical IoT protocol stack is shown where the newly adopted solutions are represented in light blue. At the application layer, two of the most sought protocols for the IoT are represented by the Constrained Application Protocol (CoAP) [2] and the MQ Telemetry Transport (MQTT) [3] protocol. Regarding the MAC layer, different solutions such as IEEE 802.15.4 [4], IEEE 802.11s [5] and Bluetooth Low Energy (BLE) [6] have been proposed. At the network layer, the IP protocol is considered as the contact point between the new IoT networks and the classical Internet. In fact, each smart object must be uniquely identified with an IP address in order to be fully integrated in the existing Internet. Given the number of involved devices, the use of IPv4 is not an option and IPv6, which has a much larger addressing space, is required. However, IPv6 is not suited for constrained nodes (e.g., it uses a large packet size) and, for this reason, is introduced in combination with the 6LowPAN [7] protocol which defines encapsulation and header compression mechanisms.

The main focus of this thesis is related to the efficient multi-hop data dissemination in Smart Cities and IoT environments. In the following, some of the major problems related to multi-hop communications in IoT and Smart Cities scenarios are described. First, in order to understand how interference can influence the performance of wireless multi-hop protocols, a coexistence analysis between IEEE 802.15.4 and

IEEE 802.11 [8] technologies is performed. Then, a survey on the main routing protocols for ad-hoc networks is provided. Last, some of the major problems related to multi-hop routing in ad-hoc networks, together with some possible solutions proposed in the literature, are described.

1.2 Coexistence Analysis between IEEE 802.15.4 and IEEE 802.11

In order to demonstrate how the heterogeneity of the involved technologies can become a problem in IoT and Smart Cities scenarios, in this section the problem of the interference between the IEEE 802.11 and IEEE 802.15.4 technologies is considered. It is worth noting that these two wireless communication technologies are representative since they are very common in IoT and Smart Cities environments and they both work in the Industrial, Scientific, Medical (ISM) band (i.e., 2.4 GHz). In the following, a brief introduction to the standards is given. Later, some existing studies about IEEE 802.15.4 and IEEE 802.11 interference will be presented by showing how communications are affected and what possible countermeasures can be adopted.

1.2.1 IEEE 802.15.4

IEEE 802.15.4 is a low-power and low-cost standard used in wireless sensor networks. It specifies the phisical layer and the MAC layer and it is maintained by the IEEE 802.15 working group. Some technical standards like ZigBee [9], are built on the IEEE 802.15.4 protocol.

The standard define two types of device: a Full Functon Device (FFD), and a Reduced Function Device (RFD). An RFD can only communicate with an FFD, whereas an FFD can communicate with FFD and RFD. Networks can be built as either peer-to-peer or star topology networks. Star topology means that the communications are made between devices and a single central controller called PAN coordinator. The peer-to-peer topology also has a PAN coordinator, however the difference from the star topology is that any device can communicate with any other device as long as

they are in range of one another.

The coordinator of a PAN can optionally bound its transmission time using a superframe structure. A superframe starts and ends with the transmission of a beacon frame and it has an active portion and an inactive portion. In a beacon-enabled network when a device has to transfer data to a coordinator, it first listen to network beacon. When the beacon is found, the device synchronizes with the superframe structure and transmit its data according to the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) protocol—every transmission must end before the arrival of the second beacon.

Regarding the PHY layer, IEEE 802.15.4 foresee two frequency bands. The low band 868/915 MHz uses binay phase shift keying (BPSK) modulation whereas the high band 2.4 GHz uses Offset Quadrature Phase Shift Keying (O-QPSK).

1.2.2 IEEE 802.11

IEEE 802.11 is a set of MAC and PHY layer specifications for implementing Wireless Local Area Network (WLAN) communications, they are created and maintained by the IEEE LAN/MAN Standards Committee (IEEE 802 [10]). The most popular are those defined by the IEEE 802.11b [8] and IEEE 802.11g [8] standards. IEEE 802.11 b/g work in the 2.4 GHz band and has 11 channels. IEEE 802.11b uses Direct Sequence Spread Spectrum (DSSS) and a complementary Code Keying Keying (CCK) modulation, IEEE 802.11g is based on a Orthogonal Frequency Division Multiplexing (OFDM) scheme and the same CCK modulation of the 802.11b standard.

1.2.3 Experimental Measurements

There are a lot of parameters that can affect the network performances such as modulation scheme, error correction, spread spectrum, frequency hopping rate, packet size, offered load, transmission power, and network topology. Except for the modulation scheme and error correction, all these parameters affect the interactions between the interferer system and the victim system. This is why coexistence is an important issue in wireless network. In order to study different coexistence parameters is

very important to create a lot of different scenarios and see how the interference influences the communication between two or more devices. In this Subsection some useful measurements-based studies are summarized.

Authors in [11] underline that there is a difference between *uplink* interference and *downlink* interference. In fact, two kinds of devices are defined in IEEE 802.11: Access Points (AP) and terminals. *Uplink* refers to the transmission of packets from terminal to access point and *downlink* refers to the transmission of packets from access point to terminals. The authors propose two kinds of scenario: the first is an *uplink* scenario in which the interferer IEEE 802.15.4 nodes are placed near to the access point, the second is a *downlink* scenario in which the interferer are placed near to the 802.11 terminal. Experimental results in [11] show that interference is stronger in the first case: the reason could be related to the different transmission powers between access point and terminal. This concept is easy to understand but is representative of the complexity of the interference problem.

In [12], coexistence problems between an IEEE 802.11b WLAN and an IEEE 802.15.4 Wireless Sensor Network (WSN) are experimentally analyzed. Experiments have been carried out by using the testbed sketched in Figure 1.2. Although there is only one 802.11 transmitting terminal, this scheme is representative of situations with more than two WLAN nodes: in fact, according to the CSMA/CA mechanism, only one WLAN station can transmit at a time. The 802.11 behavior under the effects of 802.15.4 interference is investigated both in the case of AP transmitting to Terminal Station (ST) and vice versa, thus in the *downlink* and *uplink* cases described in [11]. First of all, the configuration where the AP transmits to ST is considered: this case emulates the situation in which the 802.11 transmitter senses an interference power level that is below its CSMA/CA threshold. In other words, the channel is sensed free and the transmission are enabled: this can lead to a large amount of collisions at the receiver side. The results are summarized in Figure 1.3, where the Packet Loss Ratio (PLR) is shown as a function of the 802.11 duty cycle λ_{wlan} when the 802.15.4 network is enabled (WSN *on*-solid line) and when it is disabled (WSN *off*-solid line). PLR is defined as $1 - N_{\text{rp}}/N_{\text{tp}}$ where N_{rp} is the number of correctly received packets and N_{tp} is the number of packet delivered by the source. In absence of 802.15.4 inter-

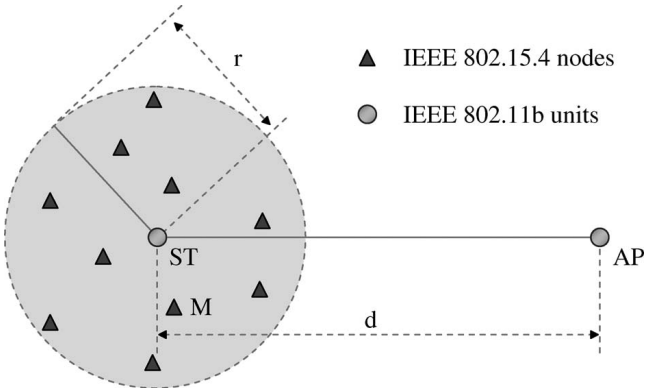


Figure 1.2: Testbed for the coexistence problem between IEEE 802.15.4 and IEEE 802.11b.

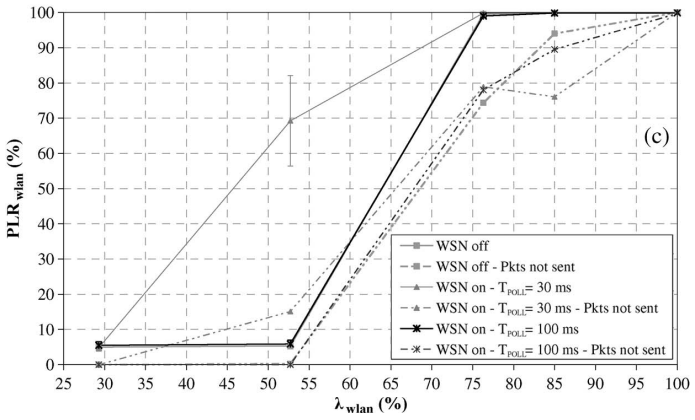


Figure 1.3: Transmission from AP to terminal: downlink mode.

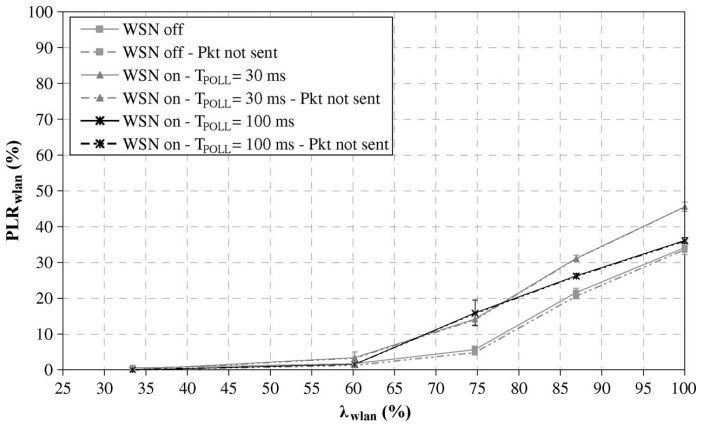


Figure 1.4: Transmission from ST to AP: uplink mode.

ference, a performance degradation is visible for increasing values of λ_{wlan}. However, in the presence of 802.15.4 interference the performance decreases further. After this analysis, the configuration in which the ST is the transmitter and the AP is the receiver is considered. This configuration emulates the typical case in which the WLAN transmitter senses a high level of the WSN signal beyond its CSMA/CA threshold. In this case, the presence of the WSN signal saturates the WLAN channel, and the ST, according to the CSMA/CA protocol, is forced to defer the transmission. Consequently, data packets are delivered from ST to the AP with a lot of delay, and increased values of PLR are experimented, the results related to this latter scenario are shown in Figure 1.4. With respect to Figure 1.3, the interference is not so detrimental this is due to the fact that in this configuration the delay is the metric that is mostly influenced by the interfering network.

In [13], the authors measure the interference with IEEE 802.15.4 by IEEE 802.11g and IEEE 802.11n. The IEEE 802.11n standard uses several technologies to improve throughput, including using wider channel bandwidth. Compared with IEEE 802.11b/g which uses 22 MHz channel, IEEE 802.11n supports 40 MHz channel, so that the measurements show that one IEEE 802.11n working channel interferes with IEEE 802.15.4 more seriously than IEEE 802.11g. A significant measurement

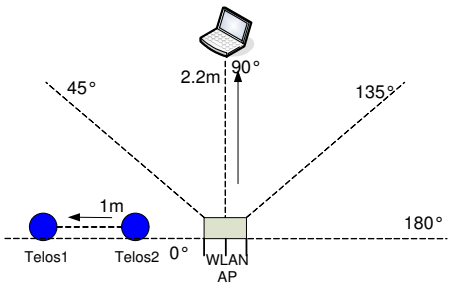


Figure 1.5: Mesurements scenario for evaluate the effect of different orientations between nodes.

in [13] is the effect of different orientations of IEEE 802.11n transmission on IEEE 802.15.4 devices. As shown in Figure 1.5, the IEEE 802.11 terminal is the only device changing the position. The mesurements results depicted in Figure 1.6 show the packet delivery ratio (PDR) as a function of the angle between the 802.15.4 (TelosB platforms [14]) and the 802.11 Line-of-sight communications. When IEEE 802.11n traffic is at 90^o , IEEE 802.15.4 has the best performance. The second best performance of IEEE 802.15.4 can be observed when IEEE 802.11n traffic is at 180^o . When the IEEE 802.11n nodes are in line with IEEE 802.15.4 nodes, namely at 0^o , almost all of the overlapping IEEE 802.15.4 channels are heavily affected by the IEEE 802.11n.

1.2.4 How to Mitigate the Interference Problem

There are many solutions which are able to statically mitigate the interference. Most of these statical solutions have already been included into the standards and perform independently of interference. Among the set of all the static solution are included: DSSS, use of multiple channels through Frequency Division Multiple Access (FDMA), CSMA, Time Division Multiple Access (TDMA), short packet length, channel scan, channel hopping, minimum overhead...etc. In the following, we describe other solutions that are performed dynamically according to the type of interference. After the detection of a high level of interference, a wireless device must

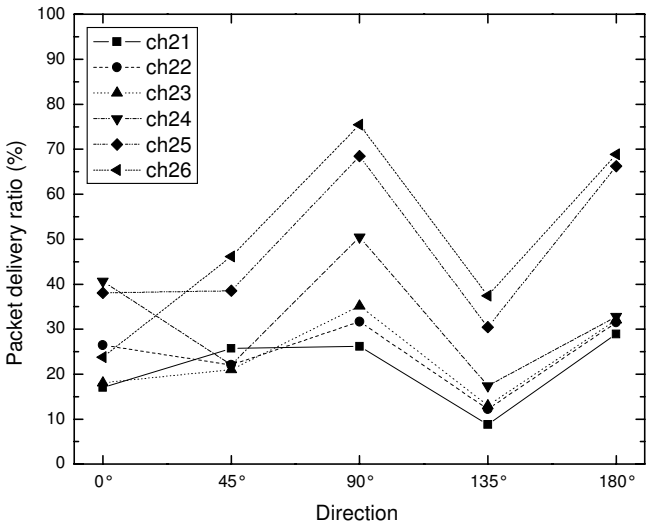
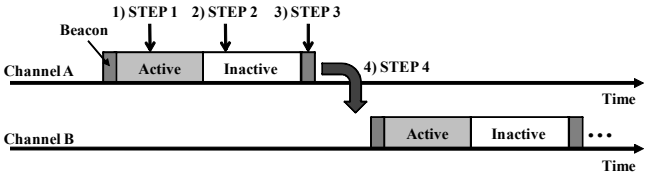


Figure 1.6: Effect of different orientations of IEEE 802.11n transmission on IEEE 802.15.4 devices.

take actions in order to mitigate the interference. Actually, these “on-demand” methods must perform some operations that are very similar to that performed by the static methods, making sure that the conflicting resources (time, frequency and space) are correctly shared among the largest possible number of devices.

The most common resource that can be shared is the frequency: therefore a set of possible solution is related to sharing the frequency which means that victim or interferer changes to a clean channel when they detect interference. This frequency change can be imposed by a central controller or in a distributed fashion.

In [15], a cognitive radio based channel switching mechanism to avoid the interference between IEEE 802.15.4 and other networks is proposed. A beacon-enabled network with star topology is considered for the analysis. The switching mechanism consists of 4 steps and is depicted in Figure 1.7. In STEP 1,the PAN coordinator can guess whether or not networks operate in the same channel by using Link Quality Indication (LQI) which represents a IEEE 802.15.4 physical layer function which measures the SNR for each channel. In STEP 2, the PAN coordinator tries again to



STEP 1: Link quality indication
STEP 2: Other user detection
STEP 3: Channel switching notification
STEP 4: Channel switching

Figure 1.7: Radio based channel switching mechanism.

detect the presence of another network in the same channel. Since all devices interact with others only during an active period of the super-frame, if the PAN coordinator perceives other signals during inactive period other networks exist in the same channel. Once the coordinator knows that other networks exist in the same channel it must notify in STEP 3 the other devices in order to do channel switching. In STEP 4, the device switches to another candidate channel. Note that the PAN coordinator has to select the next candidate free channel, thus it has to periodically control other channels to know if they are free or not.

The scheme proposed in [15] is a centralized solution and, for this reason, it restricts the scalability and robustness of IEEE 802.15.4 networks. Moreover, it does not take into account considerations on the network power consumption. In [16], the authors propose a distributed channel selection algorithm that allows the nodes to dynamically and independently adapt their transmission channel when they detect interference. The energy cost of the proposed algorithm is taken into account. This solution makes sure that the devices select the channel by themselves when interference occurs, so that it cannot prevent that a group of adjacent devices would move to the same channel.

The studies proposed above attest that the interference between IEEE 802.11 and IEEE 802.15.4 is a problem which must be addressed in the design of IoT and Smart Cities applications. In fact, interference can heavily affect the performance and, obviously, the problem is exacerbated with a large number of involved nodes deployed

with a high node spatial density.

1.3 Routing Protocols for Ad-hoc Networks

1.3.1 Mobile Ad-hoc Networks (MANETs)

A Mobile Ad-hoc NETwork (MANET) consists of a set of mobile devices that communicate dynamically without the need of a pre-existing network infrastructure. In a MANET, nodes share a common wireless channel: thus, they have to cope with typical problems of noise, evanescence, and interference (from both other nodes and other communication technologies). The main characteristic of a MANET is related to the fact that each node may forward data packets associated with multi-hop communications between other pairs of nodes. In other words each node can act as source, rebroadcaster, and destination at the same time. Nodes can move with a highly heterogeneous mobility level, thus the network topology is strongly dynamic. Moreover, they are typically constrained in terms of transmission power, battery life, and computational capabilities.

MANETs can be useful in all situations where networks need to be deployed very quickly and fixed network infrastructures are not available. Examples include, but are not limited to, battlefield and emergency scenarios, disaster area or conferences and events. Due to the growing interest in Smart Cities and IoT applications [17], in the last years subclasses of MANETs, such as Vehicular Ad-hoc NETworks (VANETs) [18] and Opportunistic (such as pedestrian or Machine-to-Machine, M2M) ad-hoc networks [19], have been intensely investigated.

The scientific community has tackled the design of multi-hop communication protocols in MANETs very intensely in the past decade, and many routing protocols for ad-hoc networks have been proposed and analyzed [20, 21].

1.3.2 Ad-hoc Routing: Principles

According to the previous subsection, the routing problem in MANETs should be addressed differently with respect to the classical infrastructure-based networks. In par-

ticular, one of the major challenges in the design of a routing protocol for MANETs is related to the fact that each source node should know the multi-hop path toward its corresponding destination node before transmitting. This problem is exacerbated the higher are the number and level of mobility of the involved devices.

In general, routing protocols for ad-hoc networks can be divided into two categories:

- Proactive protocols: with this kind of protocol, each node in the MANET knows in every moment the multi-hop path related to all other nodes. This is done through a regular exchange of control messages.
- Reactive protocols: with this kind of protocol, a source node tries to find a multi-hop route to a desired destination node only when it has packets to transmit. No control or routing information is generated by nodes that are not involved in a communication.

With the proactive approach, no initial delay is required in order to discover the route to the destination. However, in order to maintain up-to-date routing informations from each node to every other node a constant overhead of routing traffic is generated. Conversely, reactive protocols generate less overhead but the initial delay needed to discover the multi-hop path toward the destination could be potentially high.

Another significant classification of routing protocol for MANET is related to the goal of a single communication: *unicast* protocols allow the routing of packets from one source to one destination, conversely, with *multicast/broadcast* protocols a single transmitted packet is intended to be received by multiple destination nodes.

1.3.3 Unicast Routing Protocols

In this subsection, a brief description of unicast routing protocols is provided—proactive and reactive approaches will be considered.

One of the most sought proactive protocols for MANETs is the Destination Sequenced Distance Vector (DSDV) protocol [22]. With this protocol, each node keeps a classical routing table in which each entry contains the Destination ID, the next-hop ID, a hop count and a Sequence Number (SN) for that destination. The SN helps

nodes to maintain a up-to-date routing information to the destination and avoid loops. To deal with the frequently changing network topology, nodes periodically broadcast routing table updates over the network. When a node receives a route-update packet, it changes its routing table entries only if the SN of the destination indicated in the packet is higher than that one in its routing table (in case the SN is equal the route with the smallest hop count is selected). This kind of routing informations update could strongly increase the network traffic, thus, in order to reduce the network overhead, the DSDV protocol provides two types of updates: *full dump* and *incremental*. A full dump update contains all the routing table entries of a node while the incremental update contains only the routing table entries which are changed since the last full dump. The DSDV protocol is one of the first proposed protocol for ad-hoc networks and obtains reasonable performance when used in MANETs with a limited number of nodes.

Another proactive protocol for MANET is the Wireless Routing Protocol (WRP) [23]. Like the DSDV protocol, WRP provides that nodes inform each other of link changes through the use of update messages. An update message is sent only between neighboring nodes and contains a list of updates, as well as an indication of which nodes should acknowledge (ACK) the update message. Although the WRP is very similar to the DSDV protocol, it exploits different routing table maintenance procedures. While the DSDV maintains only one topology table, WRP uses a set of four tables to maintain more accurate informations: (i) Distance Table (DT), (ii) Routing Table (RT), (iii) Link Cost Table (LCT), and (iv) Message Retransmission List (MRL). The DT contains the network view of the neighbors of a node. In particular it contains a matrix where the rows index i indicates a particular node in the network and the column index j indicates a particular neighbor node. Each element in the matrix contains the distance (in terms of number of hops) and the next hop, toward a particular destination node i reported by a neighbor j. The RT contains the updated network view for all known destinations. Each entry contains some parameters such as the shortest distance, the next hop, and a flag indicating the status of the path. The LCT contains the cost (e.g., the required number of hops to reach the destination) related to each multi-hop path —the cost of a broken link is considered equal to in-

finity. Each entry of the MRL contains the sequence number of an update message, a retransmission counter, an acknowledgment-required flag vector with one entry per neighbor, and a list of updates sent in the update message. The MRL protocol records which updates in an update message need to be retransmitted by neighbors and which neighbors should acknowledge the retransmission.

Regarding the reactive routing protocols, a first example is represented by the Dynamic Source Routing (DSR) [24] protocol. The main characteristic of this protocol is the use of the so-called *source routing*: with this technique a source node knows the complete multi-hop path toward the destination before the transmission of a packet. The complete path is inserted in the header of the packet to be transmitted so that each relay node does not need to rely on classical routing tables. Since DSR is a reactive protocol, when a source node has to transmit a packet toward a destination node it must first discover the correct multi-hop path. This is done through the so-called *route discovery* process. The route discovery process requires that a source node s which wants to discover the multi-hop path toward a particular destination node d first sends a Route REQuest (RREQ) packet to all its neighbors with its own address in the header. The RREQ is uniquely identified through a broadcast id, which is composed by the combination of a sequence number and address of the sender. When a node receives a RREQ packet it first checks if it has been already received earlier, if this is not the case the RREQ is rebroadcasted appending in the header its own address. In this way, when the destination node d (or an intermediate node which knows the path toward d) receives a RREQ, it knows the multi-hop path toward the source and can reply back with a Route REPly (RREP) packet. If, due to mobility, a link present in a multi-hop path break, the source node is notified through a Route ERRor (RERR) packet. In this case a new procedure of route discovery is required.

One of the most important and sought reactive routing protocol for MANETs is the Ad-hoc On-demand Distance Vector (AODV) [25] protocol. In particular, it is used in the Zigbee standard and various variants of this protocol, such as MAD-HOC, Kernel-AODV, AODV-UU, AODV-UCSB and AODV-UIUC [26], have been proposed in the last years. The AODV protocol exploits the same route discovery mechanism of the DSR protocol, however, it does not use source routing and packets

are redirected by nodes using classical routing tables. This means that, when a node retransmit a RREQ packet it must keep track of the node from which the RREQ packet comes with a so-called *backward pointer*. In this way, following the backward pointers, the RREP packet can be delivered back to the source. When a relay node retransmits a RREP packet toward the source it set a *forward pointer* and in this way the bidirectional multi-hop path is set and the actual data transmission (from source to destination) can start. Like DSDV, the AODV protocol make use of SN in order to maintain up-to-date routing informations in the routing tables. One important feature of the AODV protocol is related to the fact that, if a route entry is not used for a certain amount of time, it is deleted. Since nodes keep track also of the predecessor node in the multi-path propagation, when an entry is deleted this predecessor is alerted with a RERR packet. This procedure repeats recursively, thus all nodes in the multi-hop path are notified. Another important property of the AODV protocol is the management of the local connectivity. If a node does not send a message to any of its neighbors within a *hello interval* (dimension: [sec]), it broadcasts a special RREP, denoted as *hello message*, containing its identity. The *hello message* is not further rebroadcasted by the node's neighbors because its field *Time To Live* (TTL) is set to 1. Hello messages are used to detect a route break as follows: if a relay node between a source and a destination fails to receive a minimum number, denoted as *allowed hello loss*, of hello messages from the next hop in the path, a notification of link failure is sent to the source. In this case, another route discovery process is required and a new "wave" of RREQ messages floods the network. For this reason, hello messages and, in particular, the values of the parameters *hello interval* and *allowed hello loss* have a significant impact on the total number of transmitted RREQ packets.

1.3.4 Multicast Routing Protocols

With multicast routing protocols, data packets are sent by a single sender to multiple receivers on the network, in other words a source transmits a single message to a selected group of recipients. For example, multicast approaches are exploited in streaming video scenarios, in which many megabytes of data are sent over the network. The major advantage of multicast is that it saves bandwidth and resources,

moreover, data can still be delivered to the destination on alternative paths even if some route breaks.

One of the most important multicast routing protocol for ad-hoc networks is the On Demand Multicast Routing Protocol (ODMRP) [27]. With the OMRP protocol, a source node which has to start a multicast communication must first create a so-called fowarding group, i.e. a group of nodes which will rebroadcast packets sent by the source in order to reach all the multicast destinations. The fowarding group is created through a procedure similar to the route discovery of the AODV and DSR protocols and is composed by request phase and a reply phase. In particular, a source node which want to create a forwarding group, first sends a Join Query (JQ) message to all its neighbors. When a node receives a not-duplicate JQ, it rebroadcast the packet to all its neighbors appending its own address in the header. When one of the multicast destination nodes receives a JQ, by reading the header, it creates a table, denoted as Members Table (MT), in which the complete multi-hop path toward the source is saved. Then the MT is send to all the multicast destination neighbors through a join reply (JR) packet. When a node receives a JR, it checks if its address is contained in the MT and, in this case, it enables the Forwarding Group Flag (FG-FLAG), understanding to be part of the forwarding group. In Figure 1.8, a representative example of forwarding group creation is shown.

1.3.5 The Broadcast Storm Problem

Broadcast multi-hop routing protocols are a subclass of the multicast multi-hop roting protocol. With this kind of protocols a source transmits a single message to all other nodes in the network, in other words a single source node disseminates informations within an entire an ad-hoc network in a multi-hop fashion. This kind of protocols assumes particular importance in the dissemination of alert messages or to transmit control messages in unicast/multicast routing protocols. The most simple protocol which performs this kind of communication is the so-called *flooding*. According to this protocol, each node is required to retransmit packets when received for the first time. The flooding strategy is very simple and is widely used by routing protocols for ad-hoc networks in order to propagate control messages. For example, as mentioned

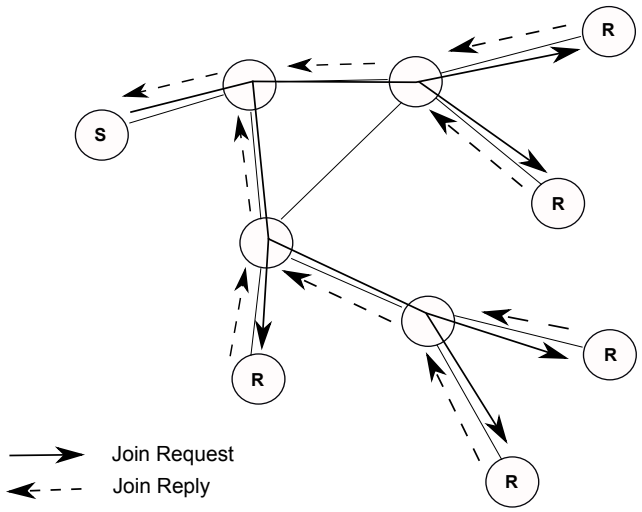


Figure 1.8: Representative example of forwarding group creation with the ODMPR protocol.

in Subsection 1.3.3, the route discovery process of the AODV and DSR protocols is based on flooding of RREQ packets. Although the flooding strategy is really simple and easy to implement, it generates a lot of redundant transmissions which can lead to serious inefficiencies related to the high level of channel contention generated. In particular, the use of flooding heavily increases the level of interference which, in turn, creates collisions degrading the performances of routing protocols. This problem is even more evident in IoT and Smart Cities scenarios which are typically characterized by lack of bandwidth and high level of interference from different technologies (as already described in Section 1.2). This issue, denoted as the Broadcast Storm Problem (BSP) [28], has been largely investigated by the research community in the last years and several methods have been proposed in order to mitigate it.

One possible solution is related to the so-called counter-based method. With this technique, if a node receives a packet at time $t = t_0$, it initializes a counter c to $c = 1$ in order to decide if the packet has to be retransmitted. Then, the node starts to wait

for a certain amount of time denoted as t_{rad} (where RAD stands for Random Assessment Delay) during which each time the same packet is received the counter is incremented. When $t = t_0 + t_{\text{rad}}$, if $c < C$ (where C is a threshold value to be properly optimized) the packet is retransmitted, otherwise the retransmission is inhibited. In this way, the node is prevented from rebroadcast when the expected additional coverage which it can provide is too low.

Another significant approach is related to the distance-based methods. With this kind of propagation technique, the decision about the retransmission is taken by the retransmitter on the basis of its distance from the transmitting node. Suppose that node h receives a broadcast message from source node s and say d the distance between h and s. If d is large, a rebroadcast from node h will result in a large additional coverage. Conversely, if d is small the additional coverage which the rebroadcast of node h is able to provide is lower. At the limit, if $d = 0$, the additional coverage will be 0 too. Thus, the distance d can be used, by node h, as a metric to determine whether to rebroadcast or not. More in detail, fixed a threshold distance value D, the distance-based method provide that if a rebroadcaster node is at a distance $d < D$ from the transmitter, the transmission is inhibited, otherwise the additional coverage provided is sufficiently high and the packet is forwarded. It is worth noting that, in order to use this kind of multi-hop propagation, each node should be equipped with a Global Positioning System (GPS) [88]. In this way, the source node can add its position (geographical coordinates) in the header of the transmitted packet and the receiver, by knowing its own position, is able to compute the exact distance. However, assuming the presence of a GPS transceiver on each device may not be an option in scenarios in which nodes are energy-constrained. In these cases the GPS transceiver could be replaced with less accurate positioning estimation techniques—for example, the inter-node distance d could be estimated using the Received Signal Strength Indicator (RSSI) [81, 82].

One of the most sought solution used to mitigate the BSP is represented by the so-called probabilistic broadcasting approach. According to this technique, a potential rebroadcaster node retransmits a packet with probability p and, consequently, inhibits the rebroadcast with probability $1 - p$. Obviously, the selection of the value of p is

crucial and should depend on fundamental network parameters (e.g., node spatial density, packet size, channel characteristics, transmission power etc...). In particular, when p is fixed and equal for all nodes in the network the probabilistic forwarding is often denotes as GOSSIP. An example of application of the GOSSIP technique is [30], where authors explore the so-called "phase transition" phenomenon as a basis for defining the value of p. More precisely, in a probabilistic forwarding scheme a phase transition occurs when the transmission probability p exceeds a certain threshold value p_{th}, denoted as critical probability, radically changing the overall behavior of the network. The phase transition phenomenon for a probabilistic broadcast technique is analyzed also in [31]. In this case, authors show that replacing flooding with GOSSIP in the route discovery process of the AODV protocol brings a significant performance improvement. In particular, for large networks, the number of control messages used by the modified AODV protocol, denoted as AODV+G, reduces by 35% with respect to that of the AODV protocol. In both [30] and [31], the concept of percolation [32, 33] is used to characterize the phase transition. Percolation studies the behavior of the flow of fluid in porous random media and some results obtained in this context have been shown to be very useful in order to understand the propagation of a multi-hop broadcast communication in simple scenarios [33].

Another way toward the mitigation of the BSP is related to the contention-based methods. With this methods, control packets are generated before the actual packet dissemination in order to elect the best forwarding nodes. A possible example is described in [70], where the authors present a family of adaptive protocols, denoted as Sensor Protocols for Information via Negotiation (SPIN). These protocols are designed to disseminate individual sensor observations to all sensors in energy-constrained WSNs. In particular, by using an initial negotiation phase, SPIN protocols ensure that only useful information will be transferred.

Another representative example of contention-based protocols is shown in [71, 72]. In this case, the authors propose a multi-hop forwarding technique denoted as Geographic Random Forwarding (GeRaF). According to GeRaF, a node, which wants to transmit a packet, broadcasts a message which is received by all its neighbors. Each receiving node then determines its distance from the final destination and evaluates

its own suitability as a relay. This is done by first dividing the coverage area in two regions: (i) the relay region, which contains all points closer to the final destination than to the transmitting node; and (ii) the non-relay region, which contains all other points. Nodes in the non-relay region are never selected as relays. The relay region is then sliced into "priority regions," on the basis of the distance from the destination. Nodes in the priority region closest to the destination contend first and one of them rebroadcasts the packet. It is worth noting that GeRaF is oriented to unicast communications and requires the knowledge, by each node, of the position of the final destination of each communication route it may belong to. This is realistic for a static wireless network.

In [73], an optimized Broadcast Protocol for Sensor network (BPS) is proposed. In order to broadcast a packet over a network, BPS allows only a few strategically selected nodes to rebroadcast. First of all, the area to be covered is partitioned into hexagons, the source node being at the center of one of these hexagons. The vertices of an hexagon are denoted as "strategic locations." When a node receives a packet, it computes its distance l from the nearest strategic location and delays the packet rebroadcast by $d = l/R$, where R is the node transmission range. The nodes closer to the strategic position will rebroadcast first and silence the other nodes (potential rebroadcasters) in their vicinity.

As already mentioned, the redundant rebroadcast created by flooding represent a problem also in VANETs. In fact, the dissemination of safety or alert messages through multi-hop communications is a key aspect in this kind of scenario. In [75], a multi-hop broadcast protocol for VANETs, denoted as RObust and Fast Forwarding (ROFF) protocol, is proposed. This protocol tries to avoid collisions due to redundant rebroadcasts by assigning to each candidate forwarder a waiting time (before packet retransmission) which is inversely proportional to its forwarding priority. The forwarding priority is assigned to forwarders depending on the distribution of empty spaces between vehicles. Through extensive simulations, ROFF is shown to make broadcasting faster and more reliable than other existing protocols. Another example of protocol which tries to reliably disseminate information in VANETs is described in [76]. In particular, this solution provides that a dynamically generated backbone of

vehicles is first created and, then, used to disseminate broadcast packets. The backbone is created taking into account vehicle movement dynamics and link quality.

In [79], the authors present an analytical framework which models the one-hop broadcast traffic generated by the 802.11p/WAVE protocol [80]. In particular, the periodic transmission of beacon packets and WAVE Service Advertisement (WSA) packets on the Common CHannel (CCH), used for signaling and safety-critical data exchange, is analytically modeled and the performance of the considered protocol, in terms of successful frame delivery probability, is evaluated. The proposed model is particularly accurate and takes into account several aspects related to 802.11p/WAVE networks such as the access priorities assigned to different packet types, as suggested by the Enhanced Distributed Channel Access (EDCA). The proposed model is validated through extensive simulations and the impact of the protocol's parameters on the obtained performance is investigated.

The problem of data dissemination is even more complicated in the so-called cognitive radio ad-hoc networks [77]. In this kind of networks, different unlicensed users (often denoted as secondary users, SUs) exploit available channels which are not used by the licensed users (i.e., primary users, PUs) which have prioritized channel access. For the SUs, the channel availability is not homogeneous across the network, since it strongly depends on the particular positions of PUs: this makes the broadcasting a complicated task for SUs.

1.4 Discussion

Multi-hop communications represent a powerful tool which can be used in IoT and Smart Cities scenarios in order to disseminate informations. For example, through this communication paradigm, safety/alert messages can be delivered to all vehicles of a VANETs or control packets in routing protocol for ad-hoc networks can be efficiently propagated. In these scenarios, the different communication technologies involved together with the high level of node spatial density create an high level of interference. The adoption of multi-hop broadcast techniques like flooding, generates a large level of redundancy which further exacerbates the interference problem. For

this reasons, save retransmissions making a good use of the available bandwidth assumes a fundamental importance and allows to efficiently cope with the BSP. In this chapter, some approaches proposed in the literature in order to solve this problem have been classified and briefly described.

In the next chapters, some innovative approaches able to mitigate the BSP, mostly based on probabilistic forwarding, are presented. This techniques have been used for different purposes: from the improvement of unicast ad-hoc protocols performance to the efficient data dissemination within VANETs.

Chapter 2

Irresponsible AODV routing

2.1 Introduction

A MANET can be defined as a set of mobile devices that communicate dynamically without the need of a pre-existing network infrastructure. In a MANET, each node may forward data packets associated with multi-hop communications between other pairs of nodes, so that each node can act as source, rebroadcaster, and destination at the same time. In this type of networks, nodes share a common channel and can be highly mobile, thus making the design of routing protocols very challenging.

The scientific community has tackled the design of multi-hop communication protocols very intensely in the past decade, and many routing protocols for ad-hoc networks have been proposed and analyzed. One of the most sought protocol for MANETs is the AODV protocol. As said in Section 1.3, this protocol relies on flooding mechanisms in order to perform the so-called route discovery phase. The flooding strategy can be very inefficient in MANETs and, because of highly redundant transmissions, can lead to serious inefficiencies, caused by collisions and interference. This problem, referred to as BSP, is more exacerbated the higher are the node spatial density, the node mobility level, and/or the data traffic load.

In general, there are many approaches that can be used in order to reduce the redundancy introduced by flooding and, thus, design energy-efficient broadcast mecha-

nisms. As detailed in Section 1.3, one interesting possibility is to adopt a probabilistic broadcasting scheme. According to the probabilistic forwarding strategy, denoted as Irresponsible Forwarding (IF), introduced in [29], a node computes the retransmission probability for each received packet taking into account the node spatial density, the transmission range (assumed fixed for all nodes), and the distance from the transmitter.

In this chapter, an innovative reactive routing protocol, denoted as irresponsible AODV (iAODV) is presented. The iAODV protocol is obtained from the AODV protocol by replacing, in its route discovery process, the flooding mechanism with IF. In order to exhaustively investigate the performance of the iAODV protocol, three different types of scenarios are considered: (i) pedestrian, (ii) pedestrian-vehicular, and (iii) vehicular.

This chapter is organized as follows. In Section 2.2, the IF technique is briefly recalled. In Section 2.3, the iAODV protocol is introduced: first the inclusion of IF within the route discovery process of AODV, leading to iAODV, is described; then, the capability of the iAODV protocol to mitigate the BSP is analytically evaluated. In Section 2.4, the three scenarios of interest (pedestrian, pedestrian-vehicular, and vehicular) are described. In Section 2.5, the performance of the iAODV protocol, directly compared with that of the AODV protocol, is investigated in the considered scenarios. Finally, Section 2.6 concludes the chapter.

2.2 Irresponsible Forwarding

IF is a probabilistic forwarding protocol according to which every node, upon reception of a packet to forward, computes (in a per-packet manner) its own retransmission probability. Since IF is based on the assumption of the knowledge of some topological network parameters, such as internode distance and node spatial density, each node is assumed to be equipped with a GPS transceiver.

In a single source scenario, the broadcast forwarding process of a single packet with IF is illustrated in Fig. 2.1 (a). The source node S transmits a packet inserting in the header: (i) its position, denoted as POS_{s}; (ii) its IP address denoted as $ADDR_{\mathrm{s}}$; and

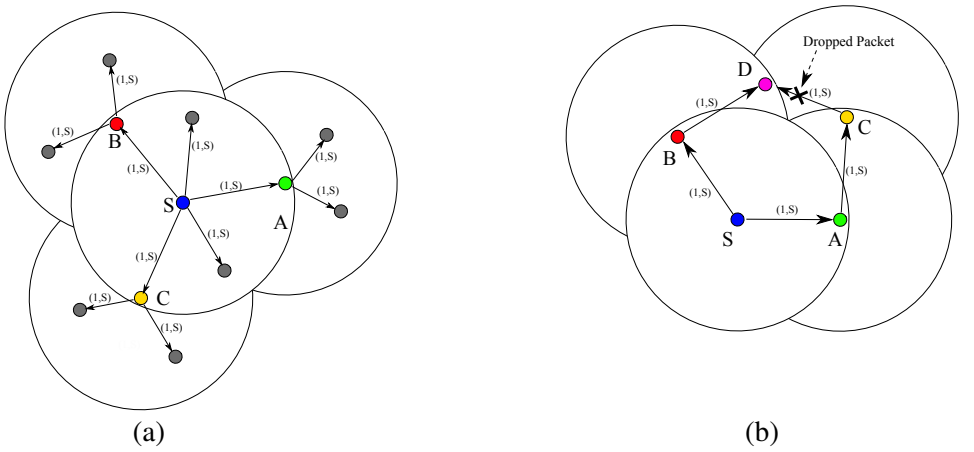


Figure 2.1: Representative examples of a single packet propagation with the IF technique.

(iii) a packet Sequence Number, denoted as *SN* (this is expedient for the transmission of streams of packets). Note that the combination $(SN, ADDR_S)$ allows to uniquely identify the packet within the network and, thus, is denoted as Unique ID (UID). In the example of Fig. 2.1 (a), node S transmits a packet with UID=(1,S). Each receiver of this packet is a neighbor of S and rebroadcasts independently with a probability computed according to a proper Probability Assignment Function (PAF). In particular, in the example of Fig. 2.1 (a), nodes A, B, and C retransmit the packet inserting, in the headers of the retransmitted packets, their positions (i.e., POS_A, POS_B, and POS_C) but keeping unaltered the UID (1,S). Each node which receives a packet from node A, checks the UID: the packet is dropped if already received earlier; otherwise, the node decides whether to retransmit or not according to the PAF previously mentioned. The process is similar for nodes B and C and repeats recursively during packet broadcasting. Fig. 2.1 (b) shows an example in which, due to multi path propagation, the packet with UID=(1,S) is transmitted to node D by both nodes B and C. In this case, since node B transmits before node C, the packet coming from node C is dropped by node D.

The choice of the PAF of IF is based on the intuitive observation that the farther the potential rebroadcaster is from the transmitter, the higher its associated rebroadcast probability should be, as this would yield the highest forward progress—this is reminiscent of the approach in [73]. Based on this idea, in [29] the PAF of IF is introduced for a monodimensional scenario (e.g., a narrow street). In a bidimensional scenario, the PAF proposed in [29] can be generalized as follows:

$$p = \exp\left\{-\frac{\sqrt{\rho}(z-d)}{c}\right\} \tag{2.1}$$

where: d (dimension: [m]) is the distance between a transmitting node and a potential rebroadcaster; z (dimension: [m]) is the transmission range; c is a shaping coefficient (adimensional), which can be used in order to tune the retransmission probability [36]; ρ (dimension: [nodes/km^2]) is the bidimensional node spatial density. Thanks to the POS parameter inserted in the packet header, a receiving node can directly compute its distance d from the transmitter. Moreover, each node can estimate its local node spatial density ρ by evaluating the distances from its (direct) neighbors (e.g., through the exchange of hello messages[1]). According to the PAF in (2.1), if the network is sparse, the overall retransmission probability is high in order to ensure complete connectivity. On the other hand, if the network is dense the overall retransmission probability is low in order to reduce useless redundant transmissions and, thus, collisions.

The idea behind the IF rebroadcast paradigm is that once a node receives a packet, it evaluates, in an average statistical sense, the presence of other nodes in its proximity. If the probability that another node can rebroadcast the packet is sufficiently high, then the node of interest "irresponsibly" chooses not to rebroadcast. In Fig. 2.2, the IF PAF (2.1) is shown, as a function of the internode distance d, for three different values of c. In all cases, $\rho = 900$ nodes/km^2 and $z = 100$ m. It can be observed that the shaping parameter c allows to "modulate" the behavior of the PAF.

We finally observe that assuming the presence of a GPS transceiver on each device may not be realistic in scenarios in which nodes are energy-constrained. How-

[1] A node simply estimate its local node spatial density as the ratio between number of nodes which reply to its hellos and πz^2, i.e., the area within its transmission range

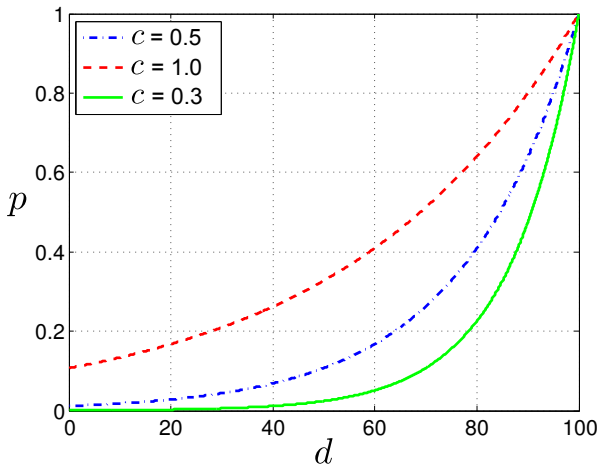


Figure 2.2: PAF (2.1) of IF, as a function of the internode distance, for various values of the shaping parameter c. In all cases, $\rho = 900$ nodes/km^2 and $z = 100$ m.

ever, in Section 6.3 it will be shown that the IF technique is robust to moderate errors in node's position estimation. This means that the GPS transceiver can be replaced with less accurate positioning estimation techniques—for example, the inter-node distance d could be estimated using the Received Signal Strength Indicator (RSSI) [81, 82].

2.3 Embedding IF into AODV

In this section, an overview on how the IF strategy can be embedded into the AODV protocol is first provided. Then, the ability of the iAODV protocol to mitigate the BSP is quantified analytically.

2.3.1 IF in the Route Discovery Process

In order to limit the number of broadcasted RREQ packets and, consequently, the BSP, the flooding mechanism, used in the route discovery process of AODV, has been replaced with IF. It can be first observed that, as explained in Section 1.3, with the

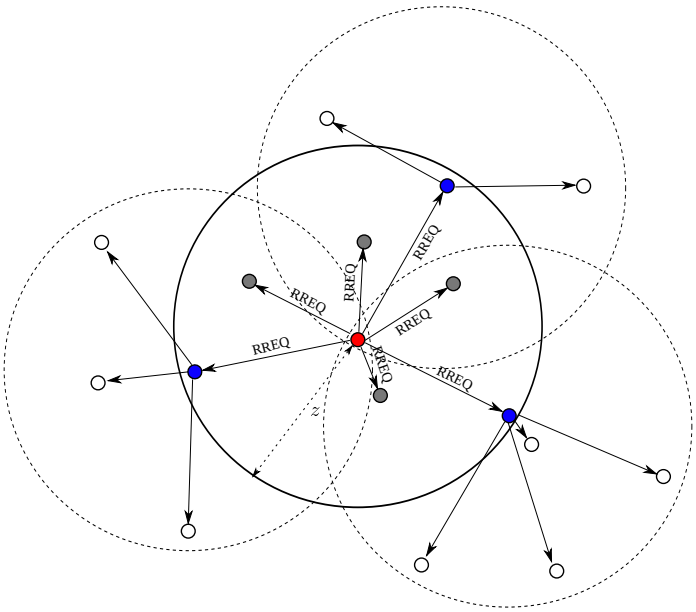


Figure 2.3: Illustrative example of iAODV route discovery process. Nodes which are far away from the source rebroadcast the RREQ with higher probability.

AODV protocol all RREQ packets are uniquely identified with a *broadcast id* which can be assimilated to the previously described UID used by the IF protocol. When a node receives a RREQ packet, it first checks the broadcast id: if it has already received another RREQ packet with the same broadcast id, the redundant RREQ packet is dropped; if this is not the case and if the node has no routing information to the destination, the RREQ packet is rebroadcasted with a probability given by (2.1). This new variant of the AODV protocol will be referred to as iAODV. In Figure 2.3, an illustrative example of the iAODV route discovery process is shown: according to the IF principle, the RREQs are propagated only (in a statistical sense) by the farthest nodes in the 1-st transmission domain. It is worth noting that the exchange of hello messages between nodes is a fundamental part of the AODV protocol. For this reason, each node is able to estimate the local node spatial density as the ratio between the number of nodes which reply to its hellos and πz^2, i.e., the area within

its transmission range.

2.3.2 BSP Mitigation: an Analytical Evaluation

In order to measure the capability of the iAODV protocol to mitigate the BSP, the number of saved redundant rebroadcasts brought by the use of IF in the route discovery phase of iAODV is analytically evaluated. While a similar analysis is carried out in [29] for a monodimensional scenario, in the following a more realistic bidimensional scenario is considered—this is more relevant for pedestrian, rather than vehicular, networks.

Let us consider the first rebroadcast round as shown in Figure 2.3, where the source node is placed in the center of its circular coverage area (with radius z) and transmits to all its neighbors the first RREQ packet. We denote the total number of rebroadcasts in this first rebroadcast round as N_{rtx}. The total number of nodes in the coverage area (i.e., the source neighbors) is, on average, the following:

$$\overline{N}_{\text{z}} = \rho \pi z^2. \tag{2.2}$$

According to the IF strategy, some of the neighbor nodes will rebroadcast the RREQ packet, while the others will be inhibited from doing it. Denoting with Γ the random variable "retransmission probability of a neighbor node" (obviously, $\Gamma \in [0,1]$), under the use of the IF protocol the average number of rebroadcasts in the first round can be written as follows:

$$\overline{N}_{\text{rtx}-\text{IF}} = \rho \pi z^2 \mathbb{E}\{\Gamma\} \tag{2.3}$$

where:

$$\mathbb{E}\{\Gamma\} = \int_{-\infty}^{\infty} \gamma f_{\Gamma}(\gamma)\, \mathrm{d}\gamma = \int_{0}^{1} \gamma f_{\Gamma}(\gamma)\, \mathrm{d}\gamma \tag{2.4}$$

and $f_{\Gamma}(\gamma)$ is the Probability Density Function (PDF) of Γ. Denoting as D the random variable representing the "distance between the source node and one of its neighbor," by applying the total probability theorem [38], $f_{\Gamma}(\gamma)$ can be rewritten as follows:

$$f_{\Gamma}(\gamma) = \int_{0}^{z} f_{\Gamma}(\gamma | D = \delta)\ \mathbb{P}\{D = \delta\}\, \mathrm{d}\delta = \int_{0}^{z} f_{\Gamma}(\gamma | D = \delta)\ f_{\text{D}}(\delta)\, \mathrm{d}\delta \tag{2.5}$$

where $f_{\mathrm{D}}(\delta)$ is the PDF of D. By replacing (2.5) into (2.4), one obtains:

$$\begin{aligned}\mathbb{E}\{\Gamma\} &= \int_0^1 \gamma \int_0^z f_\Gamma(\gamma|D=\delta)\, f_{\mathrm{D}}(\delta)\,\mathrm{d}\delta\,\mathrm{d}\gamma \\ &= \int_0^z \int_0^1 \gamma f_\Gamma(\gamma|D=\delta)\, f_{\mathrm{D}}(\delta)\,\mathrm{d}\gamma\,\mathrm{d}\delta\,. \end{aligned} \tag{2.6}$$

The PDF of Γ, conditioned to the fact that the neighbor node is at a distance δ from the source, is a Dirac delta function, i.e.,

$$f_\Gamma(\gamma|D=\delta) = \begin{cases} 1 & \text{if } \gamma = e^{-\frac{(z-\delta)}{c}\rho} \\ \\ 0 & \text{otherwise}\,. \end{cases} \tag{2.7}$$

Using (2.7) into (2.6) one obtains:

$$\mathbb{E}\{\Gamma\} = \int_0^z e^{-\frac{(z-\delta)}{c}\rho}\, f_{\mathrm{D}}(\delta)\,\mathrm{d}\delta \tag{2.8}$$

In order to find an expression for $f_{\mathrm{D}}(\delta)$, it is convenient to define a coordinate system where the source node is placed at the origin, so that $\mathscr{C} \triangleq \{(x,y) \in \mathbb{R}^2 : x^2+y^2 \leq z^2\}$ represents the set of coordinates within the coverage area of the source. Defining $\mathscr{E} \triangleq \{(x,y) \in \mathbb{R}^2 : x^2+y^2 \leq \delta^2\}$, the Cumulative Density Function (CDF) of D, denoted as $F_{\mathrm{D}}(\delta)$, can be expressed as the fraction between the area of the region identified by $\mathscr{E}$ and the coverage area, i.e.:

$$F_D(\delta) = \mathbb{P}\{D \leq \delta\} = \frac{\pi\delta^2}{\pi z^2} = \left(\frac{\delta}{z}\right)^2\,. \tag{2.9}$$

Therefore,

$$f_D(\delta) = \frac{\mathrm{d}F_D(\delta)}{\mathrm{d}\delta} = \frac{2\delta}{z^2}. \tag{2.10}$$

By replacing (4.7) into (2.8), the average retransmission probability with IF can be expressed as follows:

$$\mathbb{E}\{\Gamma\} = \frac{2}{z^2} e^{-\frac{z\rho}{c}} \int_0^z \delta e^{\frac{\delta\rho}{c}}\,\mathrm{d}\delta \;=\; \frac{2c}{z\rho}\left[1 + \frac{c}{z\rho}\left(e^{-\frac{z\rho}{c}} - 1\right)\right]. \tag{2.11}$$

Finally, by replacing (2.11) into (2.3) one obtains the following expression for the average number of rebroadcasters in the first rebroadcast domain:

$$\overline{N}_{\text{rtx}-\text{IF}} = 2\pi c \left[z + \frac{c}{\rho} \left(e^{-\frac{z\rho}{c}} - 1 \right) \right]. \tag{2.12}$$

Note that, as a consistency check, for $c \to \infty$ it follows that the right-hand side in (2.12) tends to the average number of transmissions in the first broadcast round with the flooding protocol, i.e.:

$$\lim_{c \to +\infty} \overline{N}_{\text{rtx}-\text{IF}} = \lim_{c \to +\infty} 2\pi c \left[z + \frac{c}{\rho} \left(e^{-\frac{z\rho}{c}} - 1 \right) \right] = \pi z^2 \rho \tag{2.13}$$

where the limit follows observing that the second-order Taylor series expansion of the term $\exp(-(z\rho)/c)$ is equal to $(1-(z\rho)/c+(z\rho)^2/2c^2$. We remark that, since with the flooding protocol each neighbor node rebroadcasts the RREQ packet, the average number of rebroadcasters is equal to $\overline{N}_{\text{z}}$ given by (2.2).

In Figure 2.4, $\overline{N}_{\text{rtx}}$ is shown, as a function of c, comparing the IF protocol with flooding. The node range z is set to 100 m while the node spatial density ρ is set to 2200 nodes/km^2. As expected, for very high values of c, $\overline{N}_{\text{rtx}}$ with IF converges to the value obtained by flooding. It can be observed that, by setting $c = 0.2$, the iAODV protocol can save roughly 20 rebroadcasts with respect to the AODV protocol, which corresponds to almost 30% of saved retransmissions.

We remark that the reduction of transmitted control messages predicted by the proposed analytical framework is related to a single source node in the first round of propagation: therefore, the total number of saved rebroadcasts in the entire network can be extremely larger, especially considering dense multi-source ad hoc networks. Regarding the second and following rebroadcast rounds, computing the number of rebroadcasts is much more complicated, since it depends not only by the number of rebroadcasters in the previous rebroadcast round, but also on their specific positions. This is a challenging problem and can be considered an interesting research extension. Finally, it is worth noting that the IF strategy mitigates the BSP non only by statistically reducing the number of rebroadcasted packets, as done, for example, in [31], but also by selecting, in an average statistical sense, the best rebroadcaster nodes and adapting itself to the network conditions.

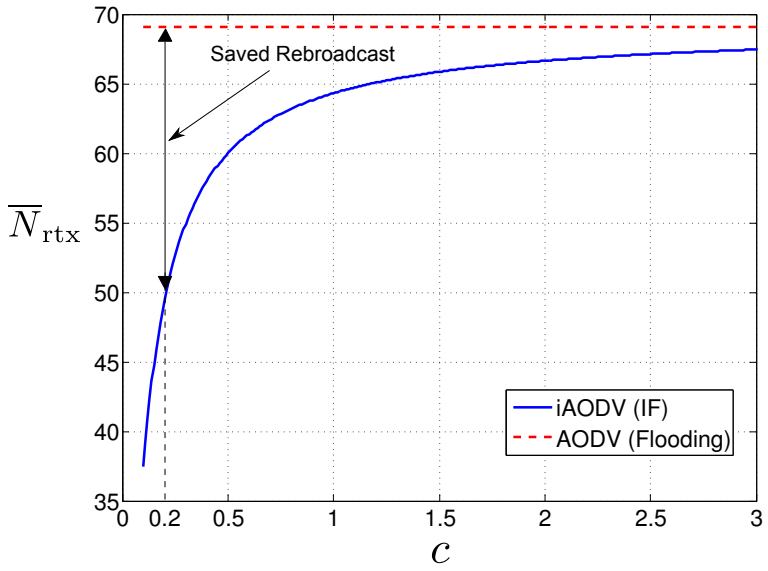


Figure 2.4: Average number of retransmissions, in the first rebroadcast round of the route discovery phase, as a function of c. The performance of iAODV (with IF in the route discovery phase) is directly compared with that of AODV (with flooding in the route discovery phase). In both cases, the node range z is set to 100 m while the node spatial density ρ is set to 2200 nodes/km^2.

2.4 Simulation Setup

In this section, the simulation set-up behind the performance analysis of the iAODV protocol is detailed. In particular, the three relevant and complementary networking scenarios of interest (pedestrian, pedestrian-vehicular, and vehicular) are described together with the selected performance metrics. All simulations are carried out with the ns-3 (ns-3.19) tool [39]. In all simulated scenarios: the number of nodes in network is denoted by N; each node has the same transmission range z (dimension: [m]); each source node generates packets of dimension P_{s} (dimension: [byte/pkt]) at a packet generation rate λ (dimension: [pkt/sec]). The packets are then transmitted with a fixed data rate (on the wireless channel) $R = 1$ Mbps. The number of source nodes is

denoted as N_{tx} and the corresponding destinations are randomly chosen among the set of all nodes. Since unicast transmissions are considered, the number of destinations is at most equal to N_{tx}. In particular, a single node may happen to be the destination for more than one source node.

At the network layer, the performance of iAODV is compared with those of AODV and AODV+G protocols [31]. As already mentioned in Section 1.3.5, the AODV+G protocol, embeds a static probabilistic broadcast in the route discovery process of the AODV protocol. In particular, the retransmission probability of the AODV+G protocol is set to $p = 0.65$ [31]. The characteristic control parameters of AODV, AODV+G, and iAODV protocols, outlined in Subsection 1.3.3, are set as indicated in [25], namely: $hello\ interval = 1$ sec and $allowed\ hello\ loss = 2$. For the lower layers, the wireless communication protocol stack defined by the ad-hoc IEEE 802.11b standard is used [8].

2.4.1 Pedestrian Scenario

This kind of scenario is representative for opportunistic ad-hoc networks of smartphones or tablets (namely, social proximity networks [19]). Since nodes correspond to pedestrians, this scenario is characterized by high node spatial density and low node speed. For example, one can imagine an application where a pedestrian may seldom send a very short information packet (e.g., containing his/her position) to intended destinations—for example this position information could be sent periodically, with a relatively long period (because of the low speed of the considered terminals). For the sake of performance analysis, nodes are assumed to be uniformly positioned over a square region with side L (dimension: [m]) given by:

$$L = \sqrt{\frac{N}{\rho}} \tag{2.14}$$

All nodes move according to the mobility model "RandomWayPointMobilityModel," available in ns-3, with speed s_{p} (dimension: [m/sec]).

An illustrative example of a pedestrian scenario with $N = 180$ nodes and $\rho = 1700$ nodes/km^2 is shown in Figure 2.5. The length of the side of the square region is

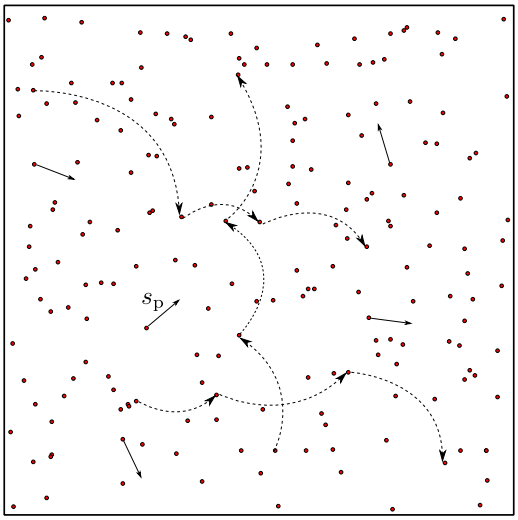


Figure 2.5: Illustrative example of pedestrian scenario. $N = 180$ nodes are deployed over a square region with a side $L \simeq 325$ m and a node spatial density $\rho = 1700$ nodes/km^2. For the sake of clarity, the speed vector (solid lines with arrows) are shown only for a few representative nodes. Multi-hop paths are represented through dashed lines.

$L \simeq 325$ m.

2.4.2 Pedestrian-Vehicular Scenario

In a pedestrian-vehicular scenario, both vehicles and pedestrians are present, so that a designer has to deal with heterogeneous devices, in terms of both speed and mobility patterns.

We define a scenario constituted by a single road, with two lanes, which surrounds a square region populated by pedestrians.The side of the square region is set as in the pedestrian scenario of Subsection 2.4.1, i.e., is given by (2.14). Nodes can be of two types: pedestrian or vehicular. In particular, N_{ped} pedestrian nodes are positioned randomly in the inner square region and move, without crossing the surrounding roads, with the same mobility model of the pedestrian scenario in Subsection 2.4.1 (i.e., ran-

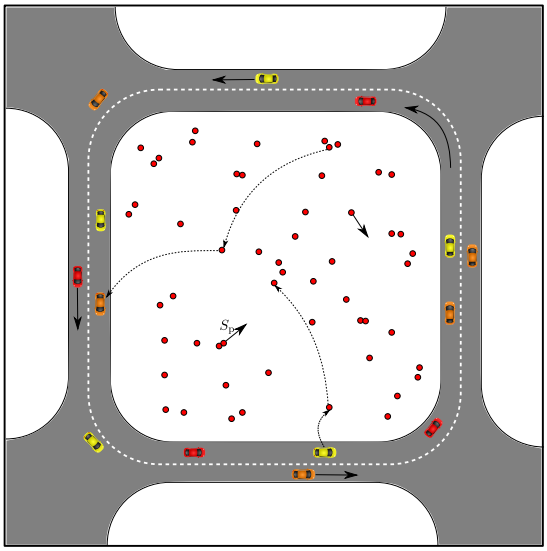


Figure 2.6: Illustrative example of *pedestrian-vehicular* scenario. $N_{\text{veh}} = 14$ vehicular nodes are positioned along the road while $N_{\text{ped}} = 56$ pedestrians are positioned into the inner square. For the sake of clarity, the speed vector (solid lines with arrows) is shown only for a few representative nodes. Multi-hop paths are represented through dashed lines

.

dom way point). The number of vehicles is denoted as N_{veh} and they are assumed to move along the road in a single driving direction. The ratio $N_{\text{ped}}/N_{\text{veh}}$ is fixed and set to 4—this is realistic for a "popular" square (e.g., a square with touristic attractions). The movement of the vehicles is generated with the SUMO open-source mobility simulator [40] integrated with the ns-3 simulator [41]. SUMO is a road traffic simulator that allows to create a vehicular scenario by using one of its external tools or by converting an existing map.

In Figure 2.6, an illustrative example of the scenario at hand, with $N_{\text{ped}} = 56$ and $N_{\text{veh}} = 14$, is shown.

2.4.3 Vehicular Scenario

The considered vehicular scenario is representative for the center of a large European city with many road intersections. In this kind of scenario, roads are typically narrow, with a single lane and a single driving direction. Moreover, the nodes' speeds are highly heterogeneous: in fact, although vehicles can move fast, they are constrained to abide by the traffic rules (priorities, traffic lights, etc.) forming queues and thus slowing down the overall vehicular traffic mobility. Since this highly dynamic mobility can strongly affect the performance of the used routing protocols, realistic VANET mobility models must be taken into account. In the last years, many approaches have been proposed in order to derive realistic mobility models for VANETs [40, 42]. In order to simulate realistic vehicular mobility, the Open Street Maps (OSM) tool [43] is exploited. OSM provides open and editable maps of the real world which can be exported into the SUMO format in order to obtain real-world vehicular mobility. Then, by integrating SUMO into ns-3, realistic VANET simulations can be run. As a representative vehicular scenario, a portion of the city center of Paris (namely the district between "Parc de la Plachette" and the "Montmartre cemetery"), shown in Figure 2.7, has been selected.

2.4.4 Performance Metrics

The simulation-based performance analysis is carried out investigating the following metrics: the throughput S (adimensional), the end-to-end delay D (dimension: [sec]), the total number of broadcasted packets N_{broad} (dimension: [pkt]), and the average communication distance d_{com} (dimension: [m]). We now shortly describe the considered metrics. The *throughput* is defined as the ratio between the number of packets that reach the intended destinations and all transmitted packets. The *end-to-end delay* is defined as the time during which a single packet stays in the network, from the generation instant (at its source) to the instant at which it reaches its destination. The end-to-end delay is obtained as the average of all per-packet delays. The *total number of broadcasted packets* is given by the sum of the number N_{hello} of hello messages and the number N_{RREQ} of RREQ packets transmitted during the entire simulation.

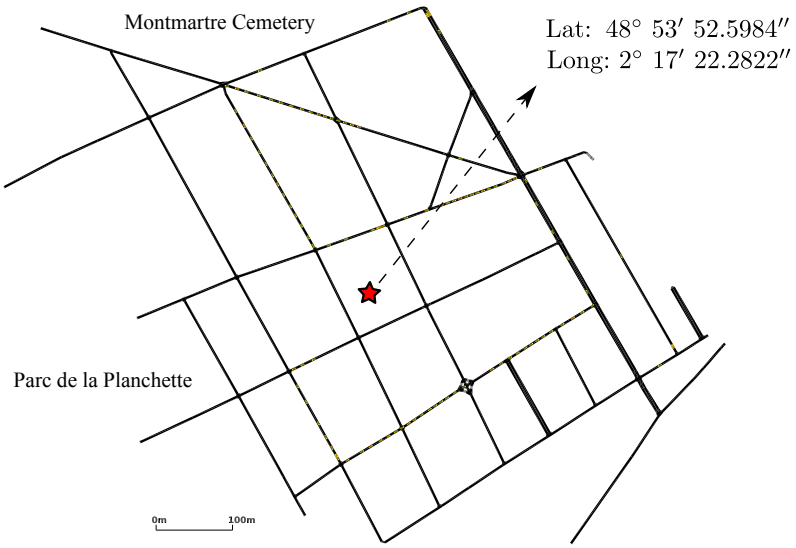


Figure 2.7: Illustrative example of VANET scenario: portion of the city center of Paris (namely, the district between "Parc de la Plachette" and the "Montmartre cemetery") imported into the SUMO mobility simulator.

The *average communication distance* is defined as the distance covered by a single packet, from the source to its final destination, averaged over all successfully delivered packets.

2.5 Performance Analysis

In this section, the simulation-based performance results obtained in the considered networking scenarios are explored.

2.5.1 Simulation Results in Pedestrian Scenarios

In Figure 2.8, the performances of AODV, iAODV, and AODV+G protocols, in terms of (a) throughput and (b) delay, as functions of the packet generation rate λ, are directly compared. The main network parameters are set as follows: $P_s = 40$ byte/pkt,

(a)

(b)

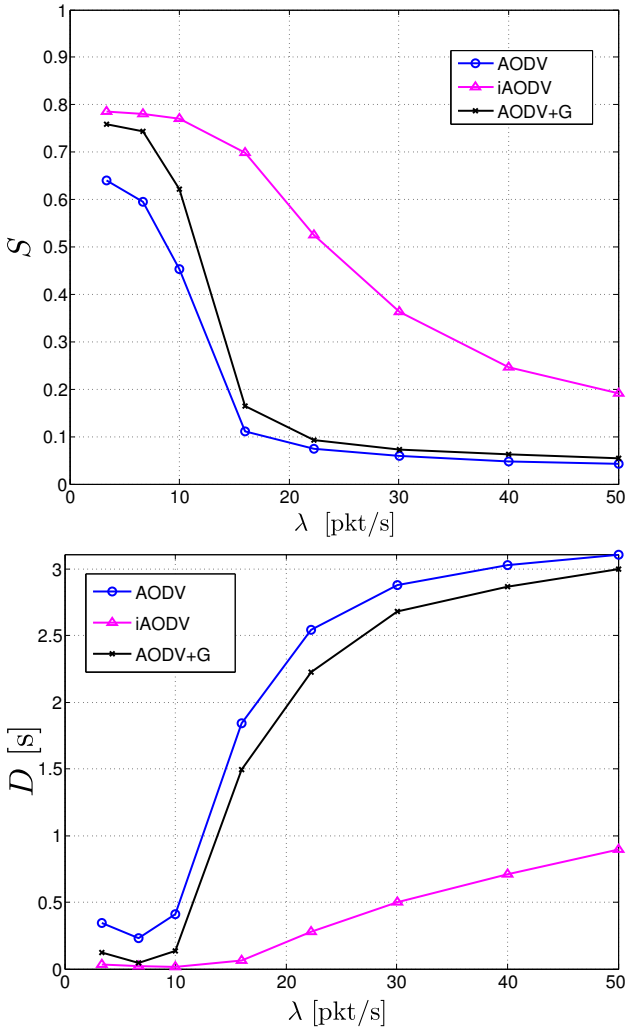


Figure 2.8: AODV, iAODV, and AODV+G protocols are directly compared, in terms of (a) throughput and (b) delay as functions of λ, in a *pedestrian scenario*. In all cases: $P_s = 40$ byte/pkt, $\rho = 1700$ node/km^2, $N = 180$ nodes, $N_{tx} = 40$ nodes, and $s_p = 1.5$ m/sec.

$\rho = 1700$ nodes/km^2, $N = 180$ nodes, $N_{\text{tx}} = 40$ nodes, and $s_{\text{p}} = 1.5$ m/sec. First of all, it can be observed that the iAODV protocol outperforms AODV and AODV+G protocols for all the considered values of packet generation rate. All the considered protocols reach the so-called saturation condition (in terms of throughput and delay) for large values of λ (namely, $\lambda \geq 40$ pkt/sec) [44]. More precisely, the network reaches a saturation regime when each source has always at least one packet in its transmission queue, so that increasing further λ does not change the network conditions and S and D remain approximately constant. The accumulation of packets in the transmission queues occurs for large values of λ also because, when the network load is high, the backoff mechanism of the lower layers (IEEE 802.11b) slows down the transmissions, in order to avoid collisions, as much as possible.

Focusing on Figure 2.8 (a), it can be observed that using the iAODV protocol, the network reaches the saturation conditions more slowly with respect to the AODV protocol. This is because, for a given value of λ, the iAODV protocol uses a smaller number of RREQ messages, with respect to the AODV protocol, leading to lower channel contention and faster transmissions, which, in turn, reduce the packets' queuing. These observations are confirmed by the delay results, which show that the delay of the iAODV protocol is significantly lower than that of the AODV protocol: this means that a single packet reaches its destination with a smaller number of backoffs, because of limited channel contentions, along the traversed hops.

Note that, in Figure 2.8, the performance of the AODV+G protocol is slightly better than, but trend-wise very similar to, that of the AODV protocol.It is worth noting that the AODV+G protocol, although designed to reduce the BSP, is outperformed by the iAODV protocol for all the considered values of λ. This is mainly due to the AODV+G protocol's inability to effectively select the rebroadcaster nodes. By selecting the rebroadcasters in a random manner, the AODV+G protocol leads to the creation of multi-hop routes with a larger number of hops with respect to the iAODV protocol. This, in turn, increases the collision probability and the number of experienced backoffs.

In order to explore the amount of traffic overhead generated by the considered routing protocols in the pedestrian scenario, in Figure 2.9 N_{broad} is shown, as a func-

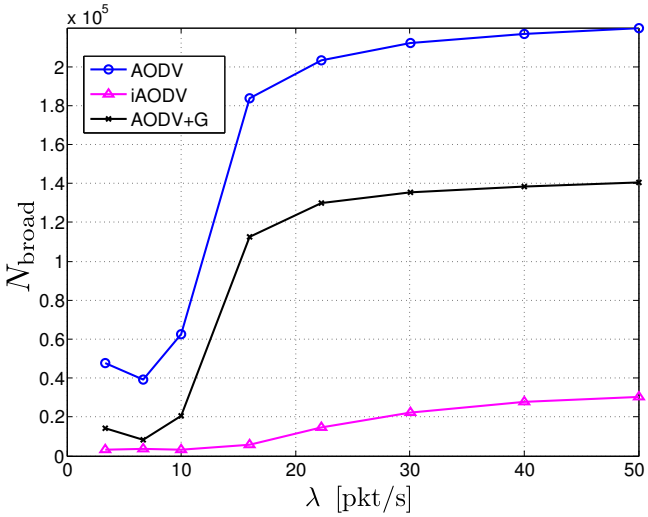


Figure 2.9: Total number of broadcasted packets, as functions of λ, in the *pedestrian scenario*. The AODV, iAODV and AODV+G protocols are compared. In all cases: $P_s = 40$ byte/pkt, $\rho = 1700$ node/km^2, $N = 180$ nodes, $N_{tx} = 40$ nodes, and $s_p = 1.5$ m/sec.

tion of λ, in the same conditions of Figure 2.8. It can be observed that the use of iAODV significantly limits the number of broadcasted messages, especially in the saturation regime (high traffic load). This leads to a better occupation of the radio channel, thus justifying the global performance improvement observed in Figure 2.8. Note that the modifications made in the route discovery process (from AODV to iAODV) do not affect the broadcast of hello messages. However, in this kind of scenario the total number N_{hello} of hello messages is negligible with respect the total number N_{RREQ} of generated RREQ packets.

Focusing on the performance of the AODV+G protocol in Figure 2.9, it can be observed that the number of RREQ packets reduces as well, with respect to the AODV protocol, limiting the BSP. However, the iAODV protocol uses a significantly smaller number of RREQ packets than the AODV+G protocol, while still guaranteeing better performance. As already observed, this is due to the fact that the iAODV protocol

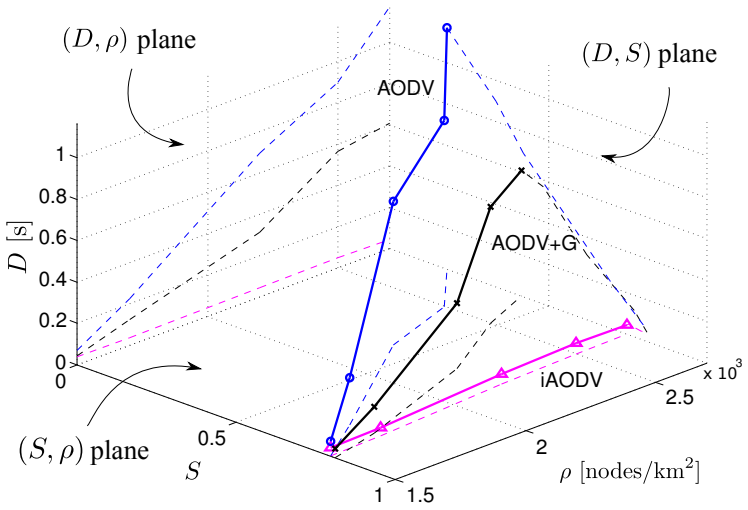


Figure 2.10: Three-dimensional characterization of delay and throughput, as functions of ρ, in the *pedestrian scenario*: AODV, iAODV and AODV+G are compared. In all cases: $P_s = 40$ byte/pkt, $\lambda = 3.33$ pkt/sec, $N_{tx} = 40$ nodes, and $s_p = 1.5$ m/sec.

selects more efficiently the rebroadcaster nodes. For example, for $\lambda = 16$ pkt/sec, the AODV+G protocol saves roughly 39% of redundant RREQ packets with $S \simeq 0.17$ (low throughput), while the iAODV protocol can save 90% of rebroadcasts while still guaranteeing $S \simeq 0.7$.

As already said, the IF strategy takes into account the node spatial density. In this way, the overall retransmission probability adapts itself to the network conditions. In order to get more insights about this feature of IF embedded into iAODV, the impact of the node spatial density on the performance of the considered protocols is investigated. The node spatial density is changed by varying the number of nodes and keeping the simulation area (i.e., the side L) fixed. In Figure 2.10, a comparative analysis of iAODV, AODV, and AODV+G protocols is carried out in a three-dimensional space, jointly considering node spatial density, throughput, and delay. For each value of ρ, the corresponding values of S and D are computed and the point (ρ, S, D) is drawn for all the considered protocols. The packet generation rate λ is fixed to 3.33 pkt/sec and the remaining simulation parameters are set as in Figure 2.8. The

projections of all curves on all possible planes (namely: (S,ρ),(D,ρ),(D,S)) are also shown.

- The projections on the plane (D, ρ) show that the node spatial density has a negative impact on the delay performance of the AODV protocol. Increasing ρ leads to an increase of the channel contention level, as there are more and more nodes within the transmission range of each other, thus increasing the collision probability. However, it can be observed that the delay of the iAODV protocol is constant with respect to the node spatial density and this proves that IF can adapt effectively its behaviour to the network conditions. The performance of the AODV+G protocol lies between those of AODV and iAODV protocols, with a trend similar to that of the AODV protocol.
- Considering the projections on the plane (S, ρ), it can be concluded that the same insights drawn for D are valid also for S. In particular, this performance metric becomes approximately independent of the node spatial density with the use of the iAODV protocol.
- Considering the projections on the plane (D,S), the overall independence of the performance of the iAODV protocol from the node spatial density becomes evident. Conversely, the AODV+G protocol, which limits the BSP in a static manner, cannot adapt itself to the network conditions and its performance degrades for high values of ρ.

In Figure 2.11, the average communication distance d_{com} is shown as a function of ρ, in the same conditions of Figure 2.10. The more efficient channel utilization brought by IF allows to support longer communication distances. This means that a single packet can be transmitted across multiple hops without being affected by collisions. At the opposite, the route discovery process of the AODV protocol floods the network with RREQ messages, increasing the probability of collisions and reducing the number of successful hops, thus making only destinations close to the source reachable. This aspect becomes more evident for higher node spatial densities (and, correspondingly, channel contention levels). On the contrary, with the iAODV protocol d_{com} seems to be independent of the node spatial density, as already observed

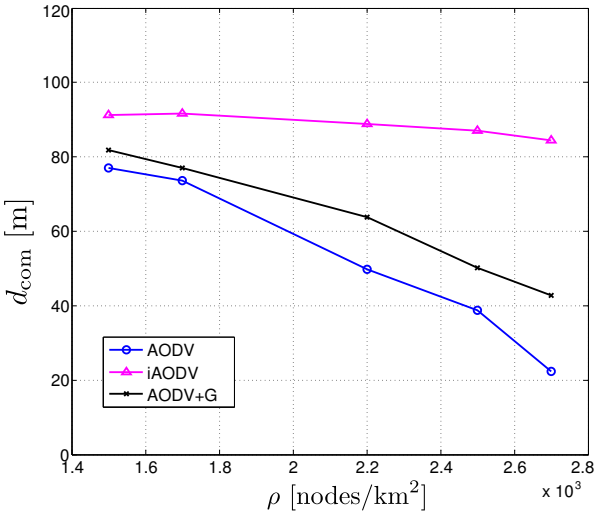


Figure 2.11: Average communication distance, as a function of ρ, in the *pedestrian scenario*: AODV iAODV and AODV+G are compared. In all cases: $P_{\mathrm{s}} = 40$ byte/pkt, $\lambda = 3.33$ pkt/sec, $N_{\mathrm{tx}} = 40$ nodes, and $s_{\mathrm{p}} = 1.5$ m/sec.

in Figure 2.10 for S and D. This confirms once more the adaptivity of IF and the corresponding benefits brought by its use in the route discovery phase of iAODV. As already observed before, the performance of the AODV+G protocol is trend-wise similar to that of the AODV protocol, with a performance improvement significantly smaller than that guaranteed by the iAODV protocol

As mentioned in Subsection 2.4.1, the pedestrian scenario is relevant to an application where pedestrians may send very short data and, in this case, the value of the packet generation rate could be relatively small. In this type of scenario, the obtained results show that the iAODV protocol outperforms the AODV and AODV+G protocols in all considered network conditions, even for medium-low values of λ. For example, with reference to Figure 2.8 (a) it can be observed that, for $\lambda = 10$ pkt/sec, using iAODV leads to a throughput increase, with respect to AODV/AODV+G, of about 73%/26%.

2.5.2 Simulation Results in Pedestrian-Vehicular Scenarios

As anticipated in Subsection 2.4.2, the pedestrian-vehicular scenario is characterized by the presence of both pedestrians and vehicles. With respect to the pedestrian scenario, smaller values of the number of devices and of the node spatial density have been considered—this is expedient to evaluate the efficiency of the iAODV protocol even in sparse networks scenarios. Since, in this scenario, the amount of information to be sent is not necessarily limited, accurate modeling calls for higher values of P_s and λ.

In Figure 2.12, (a) the throughput S and (b) the delay D are shown as functions of λ, comparing directly iAODV, AODV, and AODV+G protocols. The main system parameters are set as follows: $N = 160$ nodes, $\rho = 900$ nodes/km^2, $N_{tx} = 30$ nodes, $P_s = 128$ byte/pkt, $s_p = 1.5$ m/sec, and $c = 0.3$. Note that the packet generation rate may reach values which are twice the maximum value considered in the pedestrian case. Focusing on Figure 2.12 (a), it can be observed that, for low values of λ, all protocols have roughly the same performance. This is because, with respect to the pedestrian scenario, the values of N and N_{tx} are smaller, and the node spatial density is almost halved. In these conditions, even considering a higher value of P_s, the channel contention is strongly reduced, so that even the flooding strategy, used by AODV in the route discovery process, can perform well. However, for medium-high values of the packet generation rate, the inefficient use of the channel brought by flooding leads the AODV protocol to a fast performance degradation. Focusing on the AODV+G protocol, it can be observed that, even though it outperforms the AODV protocol, it incurs a significant performance degradation for medium-high values of the network load. Conversely, with the iAODV protocol the network does not reach the saturation conditions even for the highest considered values of λ. The delay performance, shown in Figure 2.12 (b), confirms these conclusions: for low values of λ, all protocols have good performance; for higher values of λ, the delays of the AODV and AODV+G protocols increase, while the delay of the iAODV protocol remains very low.

In addition to the packet generation rate, another parameter that strongly affects the total amount of information sent is the packet size P_s. In Figure 2.13, a three

(a)

(b)

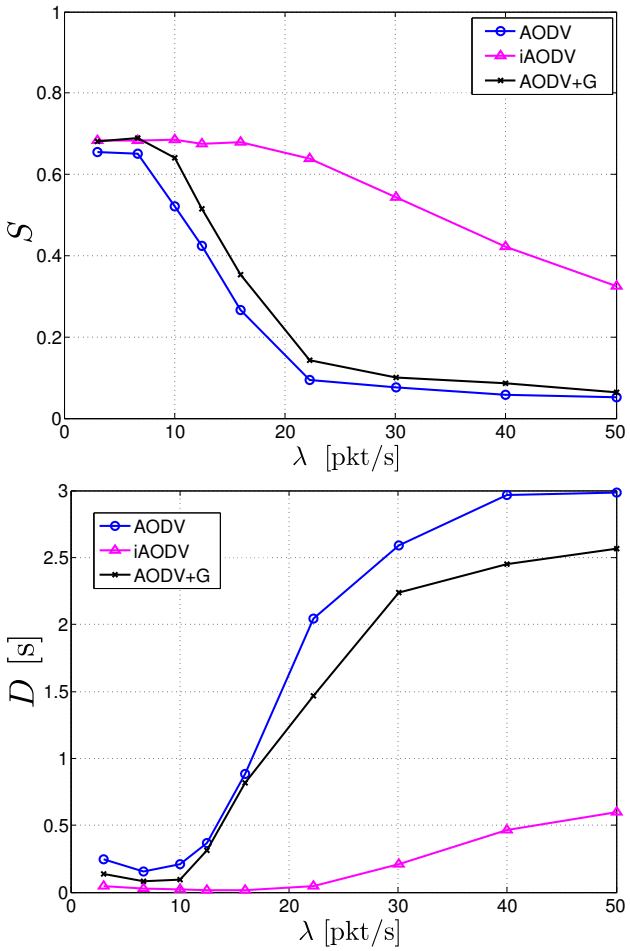


Figure 2.12: (a) Throughput and (b) delay, as functions of λ, in the *pedestrian-vehicular scenario*: the AODV, AODV+G and iAODV protocols are compared. In all cases: $N = 160$ nodes, $\rho = 900$ nodes/km^2, $N_{\text{tx}} = 30$ nodes, $P_{\text{s}} = 128$ byte/pkt, $s_{\text{p}} = 1.5$ m/sec, and $c = 0.3$.

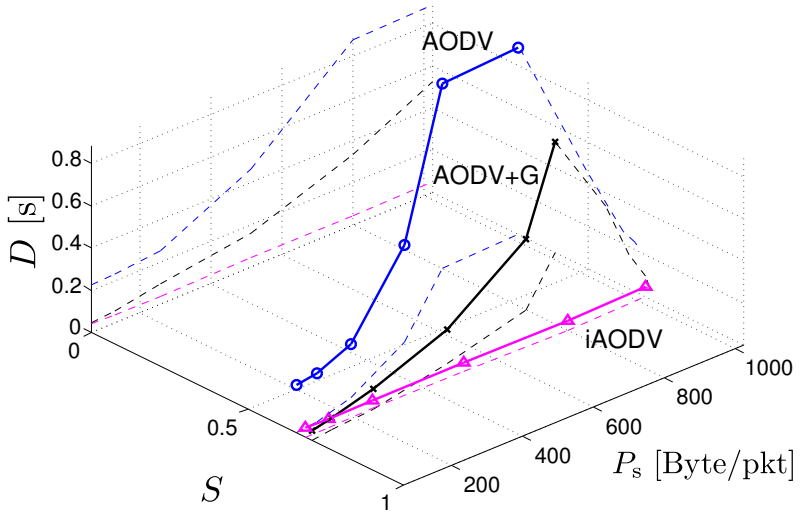


Figure 2.13: Three-dimensional (delay, throughput, and P_s) characterization, in the *pedestrian-vehicular scenario*: the AODV, AODV+G, and iAODV protocols are compared. In all cases: $N = 160$ nodes, $\rho = 900$ nodes/km^2, $N_{tx} = 30$ nodes, $\lambda = 4$ pkt/sec, $s_p = 1.5$ m/sec, and $c = 0.3$.

dimensional performance analysis is carried out jointly considering D, S, and P_s. The main system parameters are set as in Figure 2.12: in particular, the value of λ is set to 4 pkt/sec. It can be observed that the iAODV protocol outperforms the AODV protocol for all considered values of P_s. As usual, by focusing on the projections of the curves on the three "side planes," the following insights can be derived.

- From the projection on the plane (D, P_s), it can be observed that increasing the value of the packet size the delay entailed by the AODV protocol explodes. This is due to the fact that with longer packets, a transmission act takes longer, i.e., the channel is "captured" for a longer time. This results in a higher collision probability and leads to network saturation even for low values of λ. However, using the iAODV protocol, the delay does not explode: this is because limiting the number of transmitted RREQ packets saves bandwidth and the channel can thus be used for longer transmissions without the need of a frequent use of the

backoff mechanisms.

- From the projection on the plane (S, P_s), it can be observed that an increase of P_s leads, again, to a performance degradation of the AODV protocol. For large values of P_s, the iAODV protocol prevents the network from entering into the saturation regime, thus guaranteeing a better utilization of the radio channel. This makes the throughput of the iAODV protocol to remain almost constant for all the considered values of P_s.

- From the projection on the plane (S, D), it can be observed that the performance of the iAODV protocol is basically independent from P_s. As already observed before, even though the AODV+G protocol outperforms the AODV protocol, its performance does not remain acceptable for increasing values of P_s.

2.5.3 Simulation Results in Vehicular Scenarios

As anticipated in Subsection 2.4.3, a vehicular scenario representative of a big city center is considered. In particular, vehicular traffic has been simulated in a real portion of the city center of Paris. As mentioned, in such scenario nodes' mobility can strongly affect the performance of a routing protocol. Therefore, accurate modeling is crucial for the evaluation of different vehicular traffic conditions.

In Figure 2.14, the performances of the considered routing protocols, in terms of S and D as functions of λ, are shown through a three-dimensional representation. The main system parameters are set as follows: $N = 100$ nodes, $P_s = 128$ byte/pkt, $N_{tx} = 40$ nodes, and $c = 0.3$. It can be observed that the results are quite similar to those obtained in the pedestrian-vehicular scenario (Figure 2.12). In particular, while the performances of all protocols are very similar for small values of λ, the iAODV protocol outperforms the AODV and AODV+G protocols for medium-high values of the network load. The reason of this behaviour is manly due to the fact that, in the simulated settings, the vehicular scenario includes both vehicles which are free to move and queued vehicles. Therefore, this scenario can be interpreted as a pedestrian-vehicular scenario in which pedestrians correspond to the queued vehicles (almost static).

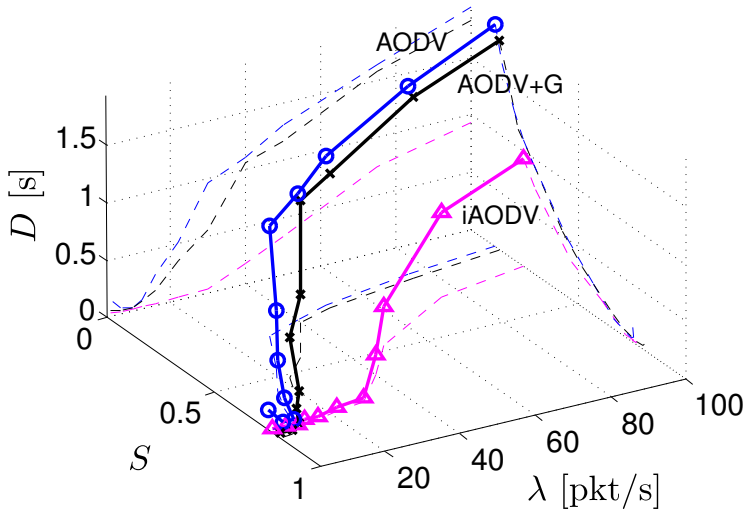


Figure 2.14: Three-dimensional characterization of delay and throughput, as functions of λ, in the *vehicular scenario*: the AODV, AODV+G, and iAODV protocols are compared. In all cases: $N = 100$ nodes, $P_s = 128$ byte/pkt, $N_{tx} = 40$ nodes, and $c = 0.3$.

In order to obtain an exhaustive analysis considering different vehicular traffic conditions, the impact of the number N of vehicles is analyzed. For small values of N, the light traffic conditions allow vehicles to move fast. Conversely, for large values of N the road traffic is congested and the creation of long queues of vehicles slows down the overall vehicular mobility. In Figure 2.15, the performances of the considered routing protocols, in terms of (a) throughput and (b) delay, as functions of N, are shown. The main network parameters are set as follows: $\lambda = 12$ pkt/sec, $P_s = 128$ byte/pkt, and $c = 0.3$. Regarding N_{tx}, the ratio N_{tx}/N is fixed and equal to $2/3$. It can be observed that, for light road traffic conditions, all routing protocols have a similar performance in terms of both S and D. When the vehicular traffic is light, there are no queued vehicles, and the node spatial density reduces: the AODV protocol can thus perform well since the channel contention level is quite low. However, for increasing values of N, vehicular traffic congestion corresponds to a significant increase of the node spatial density and the available bandwidth thus reduces: the

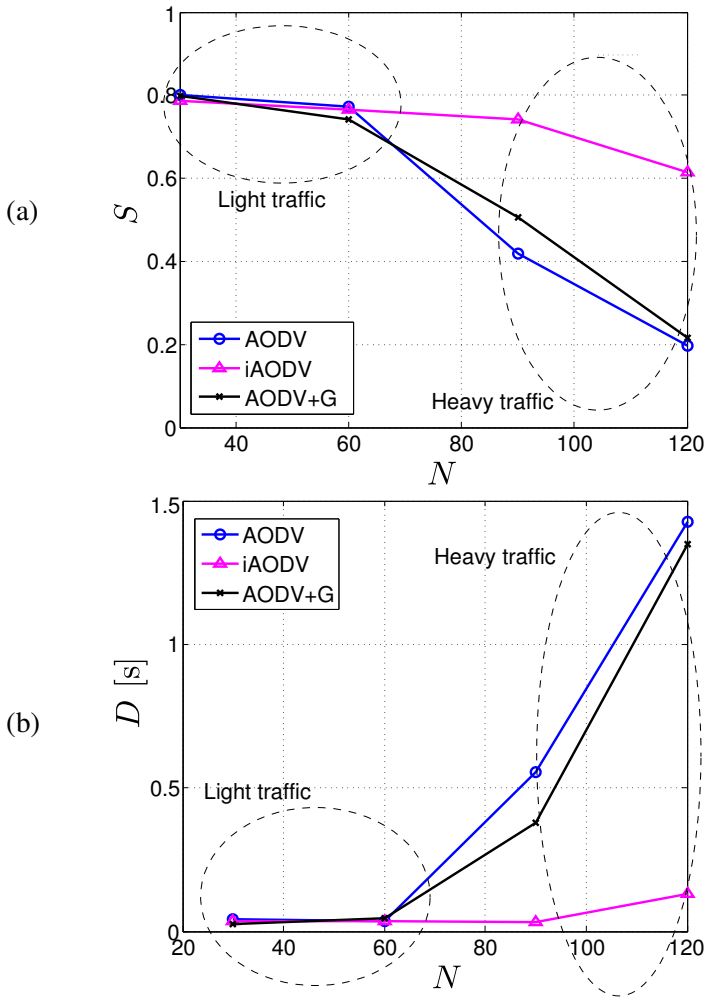


Figure 2.15: (a) Throughput and (b) delay, as functions of N, in the *vehicular scenario*. The AODV, AODV+G, iAODV protocols are directly compared. In all cases: $\lambda = 12$ pkt/sec, $P_s = 128$ byte/pkt, $N_{tx} = 2N/3$, and $c = 0.3$.

wasteful use, caused by flooding, of the radio channel degrades the performance of the AODV protocol. As already observed in the previously analyzed scenarios, the AODV+G protocol cannot guarantee a high performance for increasing values of the node spatial density. Conversely, the iAODV protocol, by effectively adapting itself to the network conditions, guarantees a good performance, in terms of S and D, even in heavy road traffic conditions.

2.6 Conclusions

In this chapter, a novel reactive routing protocol, denoted as iAODV has been proposed. This protocol is derived from the AODV protocol by replacing the flooding mechanism used in its route discovery phase with the probabilistic forwarding mechanism denoted as IF. Three scenarios have been considered: the first one is representative of a pedestrian ad hoc network; the second is representative of a pedestrian-vehicular scenario which involves both vehicles and pedestrian nodes; the third is a vehicular scenario corresponding to a real portion of the city center of Paris. In all cases, and for almost all the considered values of the network parameters, the iAODV protocol outperforms the AODV protocol and the AODV+G protocol. This is mainly due to the fact that the number of control messages is effectively reduced by the use of IF, thus reducing the contention level and making the channel utilization more efficient.

We remark that the proposed IF-based route discovery process can be applied to any reactive routing protocol which shares the same route discovery phase of the AODV protocol, e.g., the DSR protocol. This represents an interesting research extension.

Chapter 3

Decentralized Detection in clustered VSNs

3.1 Introduction

In the last decade, commercial vehicles have witnessed an exponential growth of their sensing, computational, and communication capabilities. This huge improvement is enabling the implementation of a large number of innovative services and applications, including: safety, traffic management, smart navigation, environmental monitoring, etc. By exploiting their sensing and communication capabilities, the vehicles can cooperate to create so-called Vehicular Sensor Networks (VSNs) [45]. VSNs have peculiar characteristics at various levels, from communication, networking, and data processing. From a communication perspective, the vehicles continuously gather, process, and share location-relevant sensor data (e.g., road conditions, pollution, etc.). Informations collection and dissemination can be performed using inter-vehicular communications [46] and/or relying on the presence of roadside infrastructure [47]. Moreover, each vehicle is likely to contain at least a smartphone, which is itself a powerful sensing platform. In this context, cluster-based networking is an attractive solution to reduce network congestion and to simplify routing and data aggregation/dissemination [48].

The goal of this chapter is to present a decentralized detection scheme for data acquisition in clustered VSNs, which fits well with the requirements of on-demand detection applications. In particular, the proposed scheme might be used to determine, in a timely manner, if in a given city area (e.g., several blocks) there has been a critical situation (e.g., road congestion). A possible application of interest is the dissemination of this information to prevent other vehicles from running into this congested area (e.g., adaptive cruise control for congestion avoidance [53, 54]). The proposed sensing and detection scheme foresees a two-phase communication mechanism. First of all, a *downlink phase* is triggered by a remote sink, with data collection duties, in order to form a clustered topology, constituted by ephemeral clusters (i.e., with limited lifetime) with associated Cluster Heads (CHs). The downlink phase is carried out through an innovative protocol, denoted as Cluster-head Election Irresponsible Forwarding (CEIF), which significantly improves the multihop probabilistic broadcast protocol, denoted as CIF, originally proposed in [55]. The so-formed clustered VSN is then used, during the (second) *uplink phase*, for data aggregation and/or local per-cluster fusion carried out at the CHs.

The performance of CEIF is analyzed by considering realistic dynamic ("on the move") conditions, in both highway-like and urban-like mobility scenarios. In particular, the performance of the proposed VSN clustered decentralized detection scheme is investigated considering mostly IEEE 802.11b communications between smartphones, as well as IEEE 802.11p between vehicles. A reclustering procedure, to be activated after ephemerals clusters break, is also presented.

This chapter is structured as follows. In Section 3.2 and Section 3.3, preliminaries on the system and communication models, respectively, are provided. In Section 3.4, the decentralized detection mechanism is described. In Section 3.5, the performance of the proposed scheme is analyzed in a static scenario, i.e., under average steady-state conditions. In Section 3.6, the impact of mobility on the system performance in highway-like scenarios is analyzed, from both clustering and sensing points of view. In Section 3.7, the system performance is investigated in a realistic urban-like scenario. Finally, concluding remarks are given in Section 3.8.

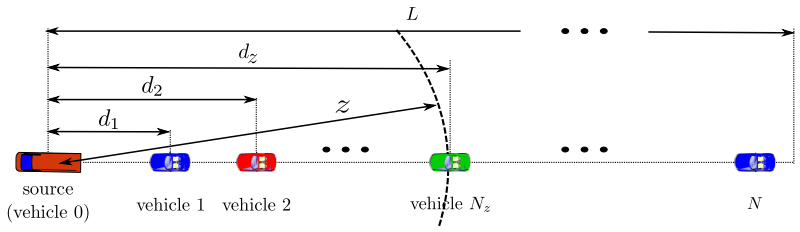


Figure 3.1: Pictorial description of a linear VSN.

3.2 System Model

Figure 3.1 shows the linear network topology of reference for a VSN: N nodes are placed in a one-dimensional scenario. This is representative of a highway-like scenario—in Section 3.7, an urban-like scenario will be considered. Each node is uniquely identified by an index $i \in \{1,2,\ldots,N\}$. The source node, denoted as node 0, is placed at the left end of the network. In order to derive the proposed clustering protocol (i.e., CEIF), steady-state conditions are first considered, i.e., a static network where nodes are positioned according to a one-dimensional Poisson point process with parameter ρ_s, where ρ_s is the linear vehicle spatial density (dimension: [veh/m])—the validity of this assumption is confirmed by empirical traffic data [57]. In Section 3.6, this assumption will be relaxed by analyzing more realistic VSNs with mobile nodes.

Each vehicle has a fixed transmission range, denoted as z (dimension: [m]), which depends on the transmit power and on the propagation model. In particular, the latter is assumed to be deterministic and the following models will be considered: Friis and Two Ray Ground [58]. Each vehicle is equipped with a GPS receiver—namely, each on-board smartphone. As a consequence, each vehicle knows its own position at any given time—this is realistic in most vehicular conditions (but galleries). The maximum network length of the linear VSN is denoted as L (dimension: [m]), so that the number N of vehicles in $[0,L]$ can be modeled as a Poisson random variable with parameter $\rho_s L$.

All vehicles observe a spatially constant phenomenon, i.e., a phenomenon whose

status does not change from vehicle to vehicle along the road. For example, vehicles could monitor if the average vehicle spatial density on the road overcomes a critical threshold (i.e., there is traffic congestion): the VNS would declare that it does if it is declared by most of the vehicles on the road. This phenomenon is typical in applications such as adaptive cruise control, where the information on the local road conditions may be used to improve the behavior of vehicles approaching that road [53, 54]. The observed binary phenomenon can be generally modeled as follows:

$$H = \begin{cases} H_0 & \text{with probability } p_0 \\ H_1 & \text{with probability } 1 - p_0 \end{cases}$$

where $p_0 \triangleq \mathbb{P}\{H = H_0\}$, being $\mathbb{P}\{\mathscr{A}\}$ the probability that the event $\mathscr{A}$ happens. The value H_0 can be interpreted as representative of a situation where the underlying physical phenomenon is, on average (along the road), below a given threshold. At the opposite, the value H_1 corresponds to the fact that the underlying physical phenomenon is, on average (along the road), above a given threshold. In the following, for notational simplicity, it will be considered that $H_0 = 0$ and $H_1 = 1$.

3.3 Inter-vehicle Communications and Clustered VANET Creation

In this section, the communication model behind the proposed VSN-based distributed detection scheme is introduced. First, a *downlink phase* is envisioned, where the sink broadcasts a query to all vehicles in the network, in order to obtain information about the phenomenon of interest. During this phase, the CEIF protocol, besides guaranteeing fast information dissemination, automatically creates a clustered VNS topology: each cluster has a single CH and all vehicles in the cluster communicate directly to it. After a clustered network topology has been generated, during the *uplink phase* in each cluster the data sensed by the vehicles (namely, by on-board smartphones) are sent to their corresponding CH where local (per cluster) fusion is performed. Then, the fused data are transferred from CHs to the sink, which takes the final decision

on the status of the observed phenomenon. The uplink collection phase has to be completed before the clusters break down because of nodes' mobility [59].

3.3.1 Downlink Phase

Before illustrating the basic operational principle of the CEIF protocol, it may be helpful to define the basic concept of Transmission Domain (TD). In general, in multi-hop broadcasting scenarios a source node starts transmitting a packet which is directed to all other nodes in the network. This initial transmission is denoted as the 0-th hop transmission, while the source itself identifies the so-called 0-th TD. The packet transmitted by the source is then received by its neighbors, which rebroadcast the packet in order to propagate it to all nodes which are beyond the source node range—these neighbors constitute the 1-st TD. In general, the group of nodes that rebroadcast in j-th transmission hop is defined as TD_j. In the "basic" linear scenario described in Section 3.2, the number of TDs required to cover all the network is a random variable, denoted as N_{TD}, taking values in the set $\{\lceil L/z \rceil, \lceil L/z \rceil + 1, ..., N\}$.

The CEIF protocol has been designed in order to choose a single CH in every TD, thus creating a unique set of communicating CHs able to cover the entire area of interest. The nodes which are not designated as CHs become children of a CH, leading to the formation of clusters with similar dimensions. The CEIF protocol relies on 3 types of control packets: (i) Cluster Initialization Packet (CIP); (ii) Probe Packet (PP); (iii) Cluster Confirmation Packet (CCP). In Figure 3.2, for illustrative purposes, the message exchange is shown in TD_1. More generally, CEIF involves two main operational subphases, which can be described as follows.

The first subphase requires the execution of three steps in every TD of the network (sequentially starting from TD_1) and terminates when the three basic steps have been completed in all TDs. We now describe the basic three steps with reference to the j-th TD ($j = 1, \ldots, N_{\mathrm{TD}}$).

I.A At a generic instant t_1, a node of TD_{j-1} (the remote sink in the case of $j = 1$) sends a CIP with a transmit power $P_{\mathrm{t}}^{\mathrm{CIP}}$, containing its own position and a unique IDentification (ID). Upon the reception of a CIP from a node of TD_{j-1},

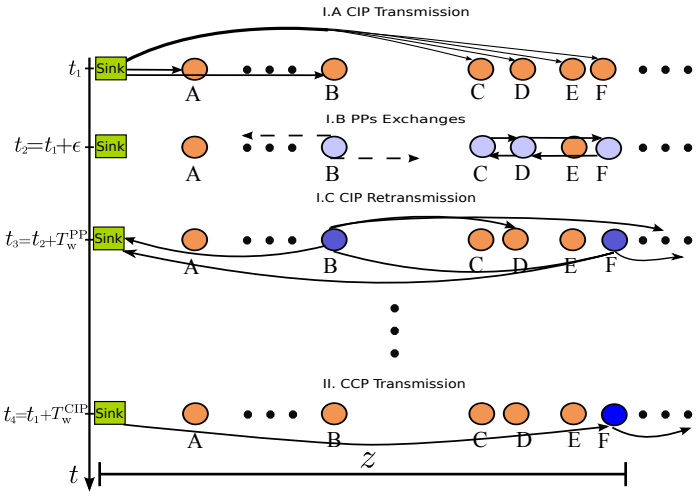


Figure 3.2: Sequence of operation carried out by the CEIF protocol in TD_1.

the receiver automatically becomes a member of TD_j.

I.B At $t_2 = t_1 + \varepsilon$, where ε accounts for on-board processing time, the nodes of TD_j start competing to designate a CH. In particular, every node in TD_j elects itself as "candidate CH" according to the PAF (2.1) given by IF. At this point, each candidate CH schedules the retransmission of a very short PP bearing: (i) the distance d from the sender of the CIP, and (ii) the ID of the CIP. Nodes which do not elect itself candidate CH simply discard the CIP. The PPs are not forwarded and are transmitted with a high priority, in order to reduce the overall latency, and with a low power which has been heuristically set to $P_t^{PP} = P_t^{CIP}/4$, in order to reduce network congestion and channel interference. Every candidate CH node sends a PP and then waits for a short interval, denoted as T_w^{PP} (dimension: [ms]): if, within this interval, it receives at least a PP containing a value of distance larger than its own, it stops forwarding and discards the packet; otherwise, it jumps to step I.C.

I.C At $t_3 = t_2 + T_w^{PP}$, the CIP is finally forwarded by the designated forwarding nodes of TD_j. The first subphase can be considered completed when the nodes

of the last TD send their CIPs. Note that, since PPs are transmitted with a lower power, with respect to CIP, more than one candidate CH node may broadcast the CIP. For example, in Figure 3.2 both nodes B and F rebroadcast the CIP.

The goal of the second subphase is to determine the actual CHs. It begins at epoch $t_4 = t_1 + T_{\mathrm{w}}^{\mathrm{CIP}}$, where $T_{\mathrm{w}}^{\mathrm{CIP}} \gg \overline{N}_{\mathrm{TD}}(\varepsilon + T_{\mathrm{w}}^{\mathrm{PP}})$ in order to guarantee that the second subphase starts after the average time required by the first subphase to complete. Initially, on the basis of its own information, every node shall elect its own CH. Note that a candidate CH can become aware of being the farthest candidate CH of its TD by simply listening to the CIP transmitted by the other candidate CHs of the same TD in the first subphase. If this is the case, the candidate CH elects itself as CH for its TD. At the same time (i.e., at t_4), the sink sends a CCP that shall be retransmitted only by the CHs, till the end of the network. By simply listening to the CCP, the remaining nodes can become aware of the identity of the true CHs.

According to the subphases of the downlink phase summarized above, the CEIF protocol can efficiently build, in a decentralized manner, a clustered topology, where each node elects its own CH without pursuing a common global consensus.

3.3.2 Uplink Phase

The uplink phase exploits the clustered topology created during the downlink phase. More precisely, during the uplink phase, the data acquired by the N vehicles of the VSN are transmitted to the final remote sink in order to estimate the phenomenon status. Note that, unlike a static wireless sensor network, the created VSN can be used as long as its structure does not break down because of vehicle mobility. In other words, there is a maximum amount of data which can be collected. The impact of mobility will be investigated in detail in Section 3.6.

The observed signal at the i-th vehicle can be expressed as

$$r_i = H \cdot s + w_i \qquad i = 1, \ldots, N \tag{3.1}$$

where $\{w_i\}$ are additive noise samples and s, which models the sensing quality, is considered as a deterministic quantity (the same for all vehicles). In particular, the

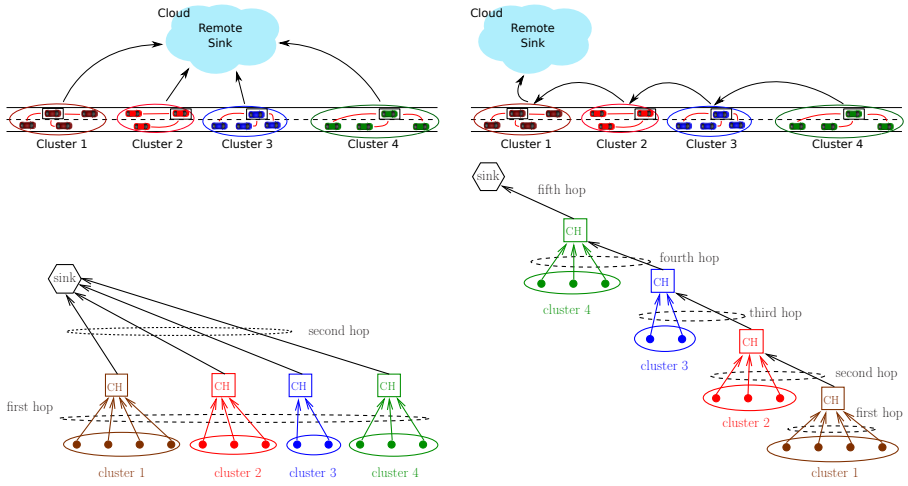
parameter s is related to the sensor sensitivity. Assuming that the noise samples $\{w_i\}$ are independent random variables with the same Gaussian distribution $\mathcal{N}(0,\sigma^2)$, the common observation signal-to-noise ratio (SNR) at the vehicles is $\mathrm{SNR}_{\mathrm{obs}} \triangleq s^2/\sigma^2$ [52]. Each vehicle makes a decision comparing its observation r_i with a threshold value $\tau = s/2$ and computes a local decision $u_i = U(r_i - \tau)$, where $U(\cdot)$ is the unit step function. Note that a vehicle could transmit one single decision per packet or, by collecting consecutive phenomenon observations, it could transmit packets with a larger number of decisions. The selected strategy depends on the desired trade-off between data and overhead per transmitted packet. However, investigating this aspect goes beyond the scope of this analysis.

Suppose that during the downlink phase the CEIF protocol has led to the creation of $N_c < N$ clusters. Each vehicle can communicate only with its local CH. Possible clustered topologies are shown in Figure 3.3, according to the particular communication strategy towards the remote sink: (a) direct communications between the CHs and the sink (e.g., through an infrastructure-based network) and (b) multi-hop communications from the CHs to the sink.

IT is worth noting that in Figure 3.3, the sink is able to communicate with all the CHs, during the VSN lifetime, while they are moving. This may be a crucial issue in the presence of WiFi communications between the CHs and the sink. However, this limitation can be overcome assuming that the sink is a "logical" sink (e.g., a server in the cloud) and CHs can communicate with it through cellular communications (3G/4G). This, for instance, is the cross-network approach proposed in [60, 61]. In the remainder of the chapter, this aspect is not further discussed, assuming that a CH, which has to communicate with the remote sink, can actually do it.

3.4 Fusion Rule and Probability of Error

According to the theoretical framework presented in [52], in the presence of a spatially constant phenomenon a key performance indicator is the probability of decision error on the final phenomenon status estimate at the remote sink. This probability can



(a) (b)

Figure 3.3: Network topologies (upper part) and their logical representations (lower part): (a) direct communications between CHs and remote sink (b) multi-hop communications between CHs and remote sink.

be expressed as:

$$\overline{P}_{\mathrm{e}} = \mathbb{P}\{\widehat{H} = H_1 | H_0\}\mathbb{P}\{H_0\} + \mathbb{P}\{\widehat{H} = H_0 | H_1\}\mathbb{P}\{H_1\}$$

where $\widehat{H}$ is the phenomenon estimate and the probabilities $\{\mathbb{P}\{\widehat{H} = H_\ell | H_m\}\}_{\ell,m=0}^{1}$ depend on the particular network topology (number of clusters and sensors per cluster) and the considered fusion rule. The fusion rule to be considered, either at a CH or at the remote sink, can be given by the following general expression:

$$\Phi(x_1,\ldots,x_M,k) \triangleq \begin{cases} 0 & \text{if } \sum_{m=1}^{M} x_m < k \\ 1 & \text{if } \sum_{m=1}^{M} x_m \geq k \end{cases} \tag{3.2}$$

where $x_1,\ldots,x_M$ are M binary data ($x_m \in \{0,1\}$) to be fused together and k is the decision threshold. For even values of M, $k = M/2$; for odd values of M, $k = \lfloor M/2 \rfloor + 1$. Note that the binary data $\{x_m\}$ may be the data observed at the vehicles and to

be fused together at the CH or the data, generated by the CHs, to be fused at the remote sink. The fusion rule in (3.2) is a majority fusion rule which guarantees a good performance in the presence of a phenomenon with equally likely statuses.

For a fixed value of the number of clusters, denoted as N_c, a clustered network topology can be described by the vector $\mathscr{D} \triangleq (\mathscr{D}_c^{(1)}, \mathscr{D}_c^{(2)}, \ldots, \mathscr{D}_c^{(N_c)})$, where $\mathscr{D}_c^{(j)}$ is the number of nodes in the generic j-th cluster ($j = 1, \ldots, N_c$). As detailed in [52], if the network topology $\mathscr{D}$ varies, the amount of information fused at each CH changes and, in turn, the probability of error changes as well. Therefore, the probability of error is a function of $\boldsymbol{\mathscr{D}}$ and it is meaningful to compute the average, with respect to the Probability Mass Function (PMF) of $\mathscr{D}$. In order to do this, the PMFs of $\mathscr{D}_c^{(j)}$, $j = 1, 2, \ldots, N_c$, and N_c are needed in (3.3). For the ease of simplicity, the value of the number N_c of clusters is set to its average $\overline{N}_c$. Under these assumptions, the average probability of decision error, with respect to the clustering configuration, can be computed as follows:

$$\overline{P}_e(\mathrm{SNR}_{obs}) = \mathbb{E}_{\mathscr{D}|N_c=\overline{N}_c}\left[\overline{P}_e\left(\mathrm{SNR}_{obs}|\mathscr{D}, N_c = \overline{N}_c\right)\right]. \tag{3.3}$$

3.5 Performance Analysis in Steady-State (Static) Scenarios

In this section, the performance of the proposed CEIF protocol is first analyzed in a static scenario, i.e., without considering vehicles' mobility. This kind of scenario is representative of a steady-state mobile scenario in which the CH election procedure has been performed and clusters are static—this is a limiting case, in the presence of mobility, where all vehicles move at exactly the same speed. In such settings, the obtained performance is a benchmark, due to the fact that mobility breaks the clusters and, consequently, degrades the performance, as will be shown in Section 3.6.

3.5.1 Set-up

Simulations are carried out using the ns-3 simulator [39], with CEIF on top of an IEEE 802.11b wireless communication stack. The main parameters of the overall

Packet Size	100 bytes
$CW_{\min}$	31
Carrier Frequency	2.4 GHz
$T_{\mathrm{w}}^{\mathrm{PP}}$	10 ms
Data Rate	1 Mbps
$T_{\mathrm{w}}^{\mathrm{CIP}}$	200 ms

Table 3.1: Main network simulation parameters for CEIF on top of IEEE 802.11b.

protocol (CEIF on top of IEEE 802.11b) are summarized in Table 3.5.1. We consider a linear network with length $L = 8z$ and $\rho_{\mathrm{s}} = 0.02$ veh/m. The following values of $\rho_{\mathrm{s}} z$ are considered: 5 veh ($z = 250$ m), 10 veh ($z = 500$ m), 15 veh ($z = 750$ m), 20 veh ($z = 1000$ m). Note that values of $\rho_{\mathrm{s}} z$ smaller than 10 veh are representative of disconnected VSNs, whereas values larger than 10 veh are typical of (highly) connected networks.

We point out that all simulations have been carried out also considering IEEE 802.11p, which is specifically designed for VANETs, as MAC protocol [62]. However, the traffic load of the considered VSN is relatively small and, therefore, the MAC protocol has a negligible impact. This is in agreement with the results in [63], where a comparison between IEEE 802.11b-based and IEEE 802.11p-based VANETS, in terms of throughput, has been performed for small levels of traffic load (i.e., under the same conditions of this work). Therefore, the IEEE 802.11b is chosen as a MAC protocol, which is more representative of VSNs based on on-board smartphones.

3.5.2 Results

The first parameter of interest for the analysis is the PMF of the number $\mathscr{D}_{\mathrm{c}}^{(j)}$ of nodes in the generic j-th cluster ($j = 1, \ldots, N_{\mathrm{c}}$), which is needed to obtain the average probability of decision error (3.3). The final PMF is obtained by averaging over 500 simulation runs: for each of them a different network topology, corresponding to a specific configuration $\mathscr{D}$, is generated. In Figure 3.4, the PMF is shown, consid-

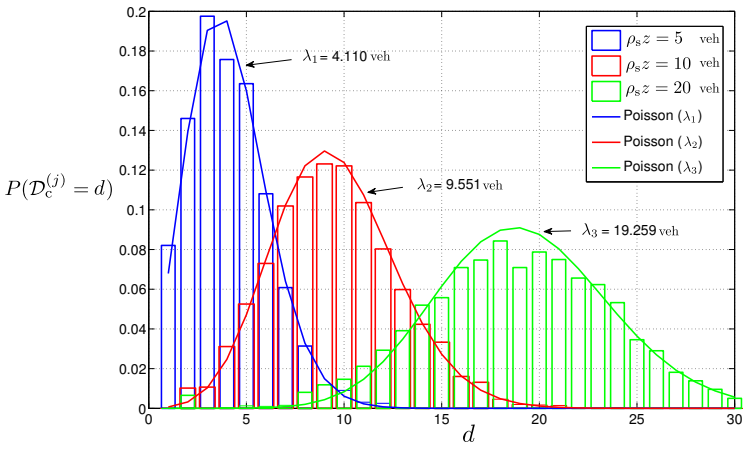


Figure 3.4: PMF of the number of nodes per cluster. Various values of $\rho_s z$ are considered.

ering various values of $\rho_s z$. As one can observe, the shape of the PMFs is the same, regardless of the value of $\rho_s z$. Despite a rigorous theoretical proof of this result is an open research problem, it can be observed that the curves have a Poisson-like shape. Using classical fitting tools, one obtains that the average value of the approximating Poisson distribution is 4.110 veh, 9.551 veh, and 19.259 veh for $\rho_s z$ equal to 5, 10, and 20 vehicles, respectively. As expected, for increasing values of $\rho_s z$, it is more likely to obtain larger cluster sizes, since there is a larger number of nodes in the transmission range of each other. In order to evaluate (3.3), the distribution of the number of clusters N_c is also needed. For the ease of simplicity, as mentioned earlier, this value is set to the average value $\overline{N}_c$ determined through simulations. More precisely, it follows that $\overline{N}_c = 7$ for $\rho_s z = 5$ veh and $\overline{N}_c = 8$ for $\rho_s z$ equal to 10 veh, 15 veh, and 20 veh.

In Figure 3.5, the probability of decision error is shown, as a function of the vehicle observation SNR, in a scenario with CHs directly connected with the sink (scenario in Figure 3.3 (a)). Different values of $\rho_s z$ are considered. As one can see, the larger the value of $\rho_s z$, the better the performance. This has to be expected, since the simulations confirm the intuitive fact that a larger value of $\rho_s z$ corresponds to a larger

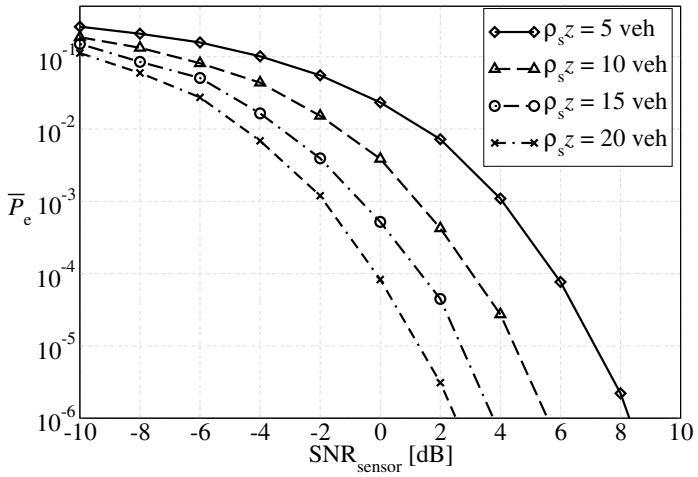


Figure 3.5: Average probability of decision error at the remote sink, as a function of the vehicle observation SNR, in a scenario with CHs directly connected with the sink (scenario of Figure 3.3 (a)). Various values of $\rho_s z$ are considered.

average number of vehicles per cluster. In fact, it is well known that the performance of decentralized detection schemes, considering majority-like fusion rules, improves by using a larger amount of sensors [52].

Considering the inter-cluster multi-hop scenario in Figure 3.3 (b), majority-like fusion (according to (3.2)) can be carried out at each intermediate CH. In this case, however, the delay is much higher than in the scenario of Figure 3.3 (a), since intermediate communication and processing is needed before obtaining the final phenomenon status estimate at the remote sink. In general, since the signal processing time is negligible, with respect to the communication time, the delay can be approximated as $D \simeq n_{\mathrm{h}} T_{\mathrm{h}}$ (dimension: [s]), where T_{h} (dimension: [s]) is the time necessary to transfer a packet from one CH to the next one or to the sink (at each hierarchical level the packet size does not increase) and n_{h} is the maximum number of hops between the vehicles and the remote sink. For instance, in the scenario in Figure 3.3 (a) $n_{\mathrm{h}} = 2$, whereas in the scenario in Figure 3.3 (b) $n_{\mathrm{h}} = 4$ (more generally, $n_{\mathrm{h}} = N_{\mathrm{c}} + 1$).

From the results it turns out that the performance of the scheme in Figure 3.3 (b),

with majority fusion at intermediate CHs, is poor, since only two binary decisions are fused at each CH. With $M = 2$, the decision threshold is $k = M/2 = 1$ and, according to the fusion rule in (3.2), a decision in favor of H_1 is taken if at least one of the two CHs are in favor of H_1. However, this decision rule is thus biased towards H_1 and this is detrimental in scenarios where the presence of the phenomenon (e.g., traffic congestion) is rare.

In order to evaluate the system performance when the phenomenon statuses are not equally likely, the case of a rare event is considered (i.e., $p_0 > p_1$). For this reason, a majority-like fusion rule biased in favor of H_0 considered , i.e., a decision in favor of H_1, at each intermediate CH in Figure 3.3 (b), is taken only if both CHs' decisions are in favor of H_1. This corresponds to setting $k = M = 2$ in (3.2), thus leading to the following fusion rule:

$$\Phi(x_1, x_2) \triangleq \begin{cases} 1 & \text{if } x_1 = x_2 = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{3.4}$$

Therefore, each CH locally decides applying the majority rule in (3.2) to the binary data coming from the vehicles in its own cluster. Then, the decision taken by the CH is fused with that coming from the CH in the previous hop using the modified rule in (3.4).

In Figure 3.6, the probability of decision error is shown, as a function of the vehicle observation SNR, in the scenario of Figure 3.3 (b) considering the decision rule (3.4) at the CHs. Different values of ρ_{sz} are considered. In a scenario with equal a priori probabilities of the phenomenon (with $p_0 = 0.5$), the performance drastically worsens with respect to the case with direct communications between the CHs and the sink (Figure 3.5). This is due to the fact that, when the number of hops increases, the number of information fusions also increases and, therefore, the amount of information transferred across the network reduces, as already observed in [52]. Since, however, the fusion rule is biased in favor of H_0, the performance improves when the observed phenomenon is rare, e.g., in the case with $p_0 = 0.9$ and $p_1 = 1 - p_0 = 0.1$.

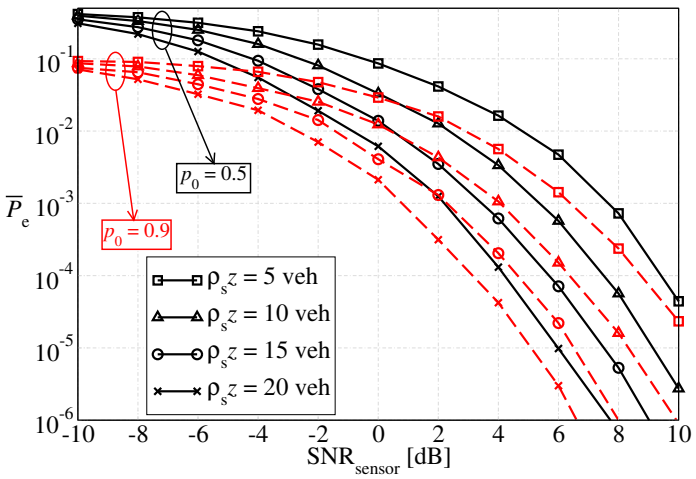


Figure 3.6: Average probability of decision error at the remote sink, as a function of the vehicle observation SNR, in a scenario with CHs connected with the sink through multi-hop communications (scenario of Figure 3.3 (b)). Various values of $\rho_s z$ and of p_0 are considered. The fusion rule (3.4) is used at the CHs.

3.5.3 Soft Fusion

In order to improve the performance in VSNs with multi-hop topologies (Figure 3.3 (b)), a soft output-based fusion rule is also considered. Denote as $n_{\mathrm{d},1}^{(j)}$ the exact number of decisions in favor of H_1 at the j-th cluster ($j = 1,\ldots,N_{\mathrm{c}}$). Similarly, $n_{\mathrm{d},0}^{(j)} = d_{\mathrm{c}}^{(j)} - n_{\mathrm{d},1}^{(j)}$ is the number of decisions in favor of H_0 in the j-th cluster. At this point, the following Logarithmic Likelihood Ratio (LLR) can be introduced:

$$\mathscr{L}_j \triangleq \ln \frac{n_{\mathrm{d},1}^{(j)}/d_{\mathrm{c}}^{(j)}}{n_{\mathrm{d},0}^{(j)}/d_{\mathrm{c}}^{(j)}} = \ln \frac{n_{\mathrm{d},1}^{(j)}}{n_{\mathrm{d},0}^{(j)}}$$

which corresponds to the logarithm of the ratio between the probability that the decision of a CH is in favor of H_1 and the probability that the decision of the CH is in favor of H_0. Obviously, $\mathscr{L}_j > 0$ if H_1 is more likely and $\mathscr{L}_j < 0$ if H_0 is more likely. Since $\forall j > 1$ each CH receives the LLR from the $(j-1)$-th cluster, denoted as $\mathscr{L}_{j-1}^{\mathrm{UP}}$, the LLR generated at the j-th cluster, to be passed to the elect CHs, can be expressed

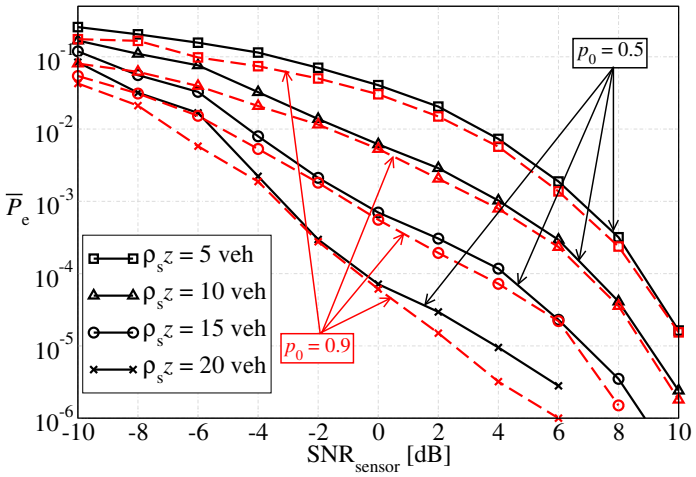


Figure 3.7: Average probability of decision error at the remote sink, as a function of the vehicle observation SNR, for the scenario in Figure 3.3 (b). Various values of $\rho_s z$ and p_0 are considered. In all cases, the LLR-based fusion rule in (3.5) is used at each intermediate CH.

as follows:

$$\mathscr{L}_j^{\mathrm{UP}} = \mathscr{L}_{j-1}^{\mathrm{UP}} + \mathscr{L}_j.$$

Finally, the sink decides with the following rule:[1]

$$\widehat{H} = \begin{cases} H_0 & \text{if } \mathscr{L}_{N_c}^{\mathrm{UP}} < 0 \\ H_1 & \text{if } \mathscr{L}_{N_c}^{\mathrm{UP}} \geq 0. \end{cases} \tag{3.5}$$

In Figure 3.7, the probability of decision error is shown, as a function of the vehicle observation SNR, in the scenario in Figure 3.3 (b), using the LLR-based fusion rule in (3.5). Comparing the results in Figure 3.7 with those in Figure 3.6, it can be concluded that the LLR-based fusion rule outperforms the majority fusion rule for both considered values of p_0 (0.5 and 0.9). This should be expected, since more information is transferred across the network—in fact, each CH transmits a LLR instead

[1]Note that in the scheme in Figure 3.3 (b) the decision taken by the sink coincides with that taken by the last (hierarchically higher) CH.

of a single bit. It is worth noting that this improved performance comes at the price of a higher energy consumption, due to the fact that the transmission of an LLR (even if quantized) requires a larger number of bits and, therefore, the energy consumption with soft fusion is higher than with majority fusion. However, depending on the chosen communication protocol, this energy penalty may be negligible. This is the case, for instance, with IEEE 802.11b communications.

3.5.4 Approximate Performance Analysis

A natural question arising at this point is the following. All simulation results in the figures presented in the previous subsections are obtained by averaging the probability of decision error with respect to the PMF of the number of nodes in each cluster. However, from Figure 3.4 one can compute the average number of nodes per cluster, denoted as $\overline{\mathscr{D}}_{\mathrm{c}}$, and assume that all clusters are composed by exactly $\overline{\mathscr{D}}_{\mathrm{c}}$ vehicles. Therefore, it is interesting to understand the relation between $\overline{P}_{\mathrm{e}}$ (as computed before) and the probability of error assuming that all clusters have the same number of nodes $\overline{\mathscr{D}}_{\mathrm{c}}$. The latter can be computed according to the analytical framework proposed in [52] for the scenario of Figure 3.3 (a). In Figure 3.8, the average error probability is shown, as a function of the observation SNR, considering the fusion rule in (3.2), for various values of $\rho_{\mathrm{s}z}$. The performance of $\overline{P}_{\mathrm{e}}$ (solid lines) is compared with the probability of error with clusters of the same size $\overline{\mathscr{D}}_{\mathrm{c}}$ (dashed lines). One can observe that, for each value of $\rho_{\mathrm{s}z}$, the gap between the two curves is limited. Moreover, the performances are trend-wise very close and, therefore, the "exact" average error probability can be accurately approximated, assuming that all clusters have the same size $\overline{\mathscr{D}}_{\mathrm{c}}$. This result will be used in the next sections, where the performance in the presence of mobility will be analyzed. In particular, it will be assumed that, on average, all clusters will be composed by the same number of vehicles—this is meaningful in the presence of relatively uniform traffic conditions.

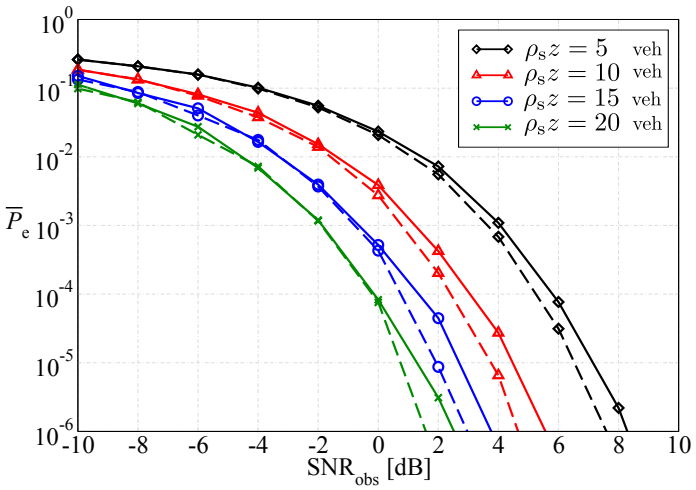


Figure 3.8: Average error probability, as a function of the SNR, for various values of $\rho_s z$. The average error probability according to (3.3) (solid lines) is compared with the average probability of error when all clusters have the same size $\overline{\mathscr{D}}_c$ (dashed lines).

3.6 Performance Analysis: Mobile Scenario

In this section, the impact of vehicle mobility on the performance of the proposed vehicular decentralized detection scheme is investigated. In particular, all results refer to the 2-hop scenario in Figure 3.3 (a) with fusion rule given by (3.2). As anticipated in Section 3.3, this scenario is realistic since that next generation intelligent transportation systems are expected to be equipped with a heterogeneous mix of hardware and communication technologies [60, 61].

3.6.1 Set-Up

In order to analyze the impact of mobility, the Simulation of Urban MObility (SUMO) traffic mobility suite [40] has been integrated into the ns-3 simulator. SUMO implements a sophisticated car-following model which is an extension of those developed by Stefan Krauß [64, 65].

Using SUMO, a highway-style scenario have been analyzed where the speed of

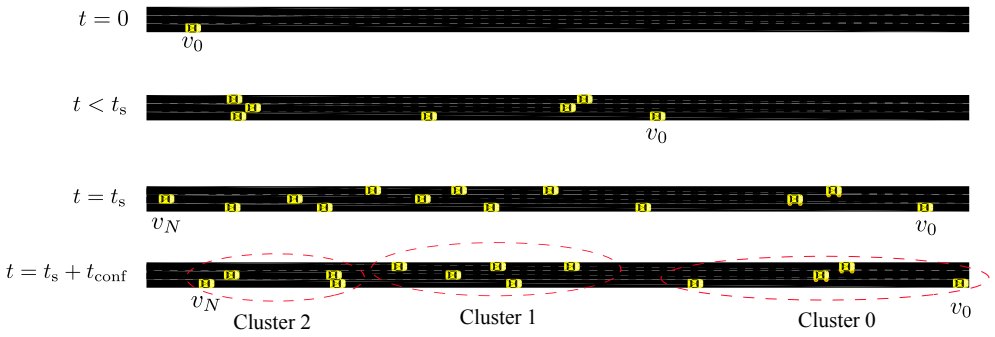


Figure 3.9: Representative example of network evolution and clustering configuration in a mobile scenario.

each vehicle is a normally distributed random variable with mean μ_s (dimension: [m/s]) and deviation σ_s (dimension: [m/s]). According to the legal limits in Italian highways (namely, 35 m/s) μ_s and σ_s have been set to 35 m/s and 4 m/s. This means that half of the vehicles drive below the legal speed limit, but the other half of them exceed it—a realistic assumption in Italy. A road with length $L = 40$ km is considered with three lanes and an approximately constant vehicle spatial density ρ_s equal to 0.015 veh/m. The node range z is set so that the following values of $\rho_s z$ are obtained: 7.5 veh, 11.25 veh, and 15 veh.

3.6.2 Cluster Formation

The network evolution, with an illustrative representation shown in Figure 3.9, can be described as follows. At time instant $t = 0$, the first vehicle, denoted as v_0, enters the highway. Then, other vehicles (namely $v_i,\ i = 1,\ldots,N$) follow, one at a time, v_0 with their own speeds and abiding by the car-following model used by SUMO. At time instant t_s, the remote sink initializes the cluster formation and the last vehicle v_N starts the configuration process. From a simulation point of view, the starting time t_s is selected in order to make sure that all vehicles have entered the road. In particular, $N = 200$ vehicles are considered and the starting time is heuristically set to $t_s = 340$ s. The clusters are created at time $t_s + t_{conf}$, where t_{conf} is the time needed to create

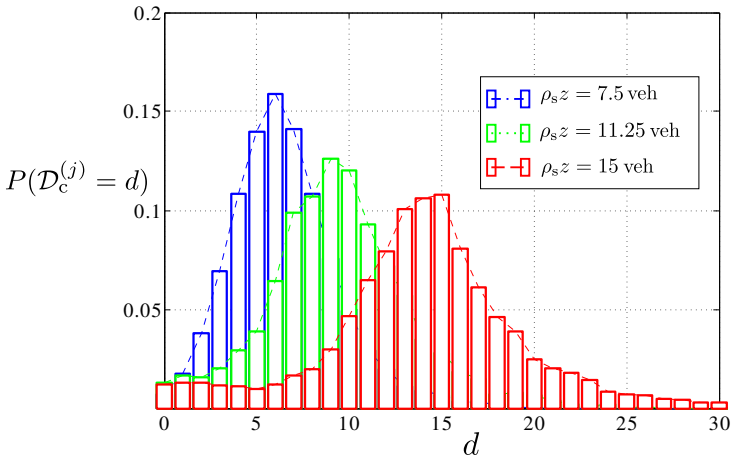


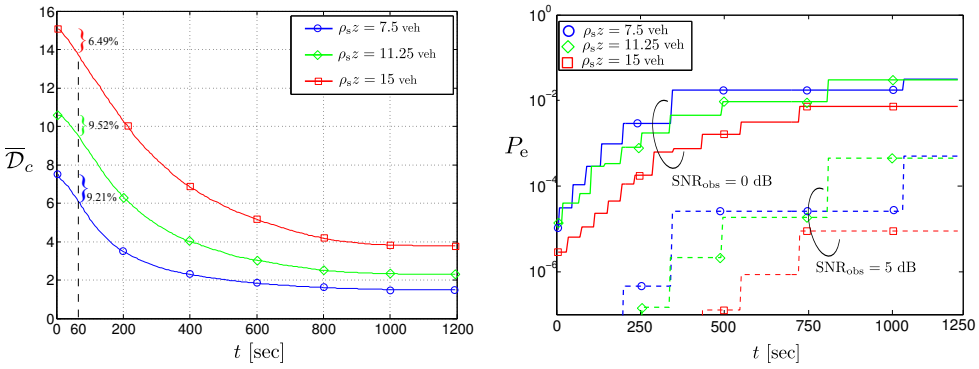
Figure 3.10: PMF of $\mathscr{D}_c^{(j)}$ in the mobile scenario. Various values of $\rho_s z$ are considered.

the clusters. In this scenario, nodes move overtaking each other and the clustering configuration changes as time goes by.

As a consistency check, in Figure 3.10 the PMF of $\mathscr{D}_c^{(i)}$, $i = 1,2,\ldots,\overline{N}_c$ is shown, in the mobile scenario introduced in the previous paragraph, at the end of the clustering procedure. The PMF is obtained by averaging over 500 trials where, in each simulation run, a different mobility configuration is generated by SUMO. Note that the shape of the PMF is approximately the same of that in Figure 3.4 for the static scenario. In particular, the cluster size is an increasing function of $\rho_s z$, since the number of nodes within the transmission range of each other increases. This result justifies the steady-state Poisson node distribution in the static scenario of Section 3.5.

3.6.3 Cluster Evolution and Network Lifetime

As mentioned earlier in Section 3.5, the number of clusters N_c and the number of nodes in each cluster (i.e., the overall network topological structure $\mathscr{D} = \{\mathscr{D}_c^{(1)},\ldots,\mathscr{D}_c^{(N_c)}\}$ have a direct impact on the average probability of decision error. In particular, mobility has a strong influence on the network topology and, as time goes by, the per-



(a) (b)

Figure 3.11: Performance evolution, as a function of time, in terms of: (a) average number of nodes in a cluster $\overline{\mathcal{D}}_c$ and (b) average probability of error for two values of SNR_{obs}: 0 dB (solid lines) and 5 dB (dashed lines).

formance can heavily degrade, as clusters tend to "stretch," i.e., vehicles disconnect from their CHs.

In Figure 3.11, the performance evolution is shown as a function of time in terms of: (a) average number of nodes in a cluster $\overline{\mathcal{D}}_c$ and (b) average probability of decision error. In case (b), two values of the observation SNR are considered: $\text{SNR}_{\text{obs}} = 0$ dB (solid lines) or 5 dB (dashed lines). In case (a), one can observe that, as time evolves, the average number of nodes in a cluster decreases. This is due to the fact that vehicles have different speeds and, in a given cluster, a vehicle may overtake its CH, finally exiting from its transmission range. However, this process is relatively slow: in fact, after one minute, the average reduction of the number of vehicles in a cluster is of 9.213%, 9.524%, and 6.487% when $\rho_s z$ is equal to 7.5 veh, 11.25 veh, and 15 veh, respectively. For $t \to +\infty$, $\overline{\mathcal{D}}_c$ is likely to tend to 1 (i.e., the CH remains alone), since all nodes may go outside the node range of the corresponding CH and it is very unlikely that two nodes have the same speed. This limiting behavior is not observed in Figure 3.11, due to the fact that the road has a finite length. In order to determine the probability of error in case (b), the average number of nodes in a cluster is obtained by quantizing the curves of case (a). Moreover, the simulations results show that the

average numbers of clusters $\overline{N}_{\mathrm{c}}$ is 22, 15, and 13 for values of $\rho_{\mathrm{s}}z$ equal to 7.5 veh, 11.25 veh, and 15 veh, respectively. It can be observed that the probability of error is a monotonically increasing function of the time and the case with $\rho_{\mathrm{s}}z = 15$ veh guarantees the lowest value of $\overline{P}_{\mathrm{e}}$ at any time. This is due to the fact that the larger the number of nodes per cluster, the lower the probability of decision error. Note that, for $t \simeq 750$ sec, the case with $\rho_{\mathrm{s}}z = 11.25$ veh corresponds to a higher value of $\overline{P}_{\mathrm{e}}$ with respect to the case with $\rho_{\mathrm{s}}z = 7.5$ veh. This is due to the fact that, even if the number of vehicles per cluster is quite similar in both cases, with $\rho_{\mathrm{s}}z = 7.5$ veh the number of clusters is larger.

Since, as shown in Figure 3.11 (a), the average number of nodes per cluster decreases over time, it follows that clusters "break." Therefore, when $\overline{\mathscr{D}}_c$ reduces below a given threshold, the cluster creation procedure should be restarted to avoid that the probability of decision error becomes too high. Based on this motivation, the *network lifetime* is crucial and has a significant impact on the quantity of information that can be collected before the clustered network topology breaks down. In order to get more insights on the network lifetime, it is first necessary to clearly define this metric.

Various definitions of network lifetime have been proposed in the literature, depending on the application of interest. For instance, one can define the network lifetime as (i) the time interval (after clustered VANET creation) until the first cluster breaks or (ii) the time interval until a given Quality of Service (QoS), e.g., a given probability of error at the remote sink, is guaranteed. If the latter definition of network lifetime is considered, one can determine its value from the results in Figure 3.11 (b), by observing the time instant at which a target maximum probability of error is achieved. As an example, if the target probability of error (i.e., the chosen QoS indicator) is 10^{-4}, $\rho_{\mathrm{s}}z = 15$ veh, and $\mathrm{SNR}_{\mathrm{obs}} = 0$ dB, then the desired QoS is guaranteed until $t \simeq 250$ s. If, instead, the former definition of network lifetime is considered, then the *cluster lifetime* first needs to be defined. For example, one can assume that a cluster breaks when a given percentage of its vehicles disappear, i.e., they are no longer connected to the CH (and, in turn, to the remote sink). We denote as $\Gamma_{\ell}^{(i)}$ the random variable representing the duration of the link between a CH and the i-th of its children vehicles—for the j-th cluster ($j = 1, \ldots, N_{\mathrm{c}}$), $i \in \{1, \ldots, \mathscr{D}_{\mathrm{c}}^{(j)}\}$.

In [1], the authors prove that the PDF of $\Gamma_\ell^{(i)}$ has the following log-normal distribution:

$$f_{\Gamma_\ell^{(i)}}(\gamma) = \frac{1}{\gamma\sqrt{2\pi\sigma_\Gamma}} \exp\left[-\frac{(\ln\gamma - \mu_\Gamma)^2}{2\sigma_\Gamma^2}\right]$$

with μ_Γ (dimensions: [s]) and σ_Γ (dimensions: [s]) proper parameters which can be derived from experimental measurement campaigns.

Let assume that a cluster breaks when the first of its vehicles disconnects. According to this assumption, the cluster duration can be given by the following expression:

$$\Gamma_c = \min\left\{\Gamma_\ell^{(1)}, \ldots, \Gamma_\ell^{(n)}\right\} \tag{3.6}$$

where n is the number of vehicles in a given cluster[2] (without considering the CH). Since each variable $\Gamma_\ell^{(i)}$ follows a log-normal distribution, it can be given the following expression:

$$\Gamma_\ell^{(i)} = e^{\mu_\Gamma + \sigma_\Gamma Z_i} \tag{3.7}$$

where $Z_i \sim \mathcal{N}(0,1)$. By replacing (3.7) into (3.6), one obtains:

$$\Gamma_c = \min\left\{\Gamma_\ell^{(1)}, \ldots, \Gamma_\ell^{(n)}\right\} = e^{\mu_\Gamma + \sigma_\Gamma \min\{Z_1,\ldots,Z_n\}} = e^{\mu_\Gamma + \sigma_\Gamma Z} \tag{3.8}$$

where $Z \triangleq \min\{Z_1, \ldots, Z_n\} \sim \mathcal{N}(0, 1/n)$. This proves that, identifying the cluster death in correspondence to the instant of first vehicle disconnection from the CH, the PDF of the cluster lifetime is log-normal as well.

More generally, one can assume that a cluster lives until a given number v of links breaks (with $v > 1$). By denoting as $\Gamma_c^{(v)}$ ($v = 1, \ldots, n$) the random variable

[2]Note that, in the previous sections, the number of vehicles in the j-th cluster has been denoted as $\mathscr{D}_c^{(j)}$. However, for the ease of simplicity, in the following the number of vehicles in the generic cluster will be denoted as n.

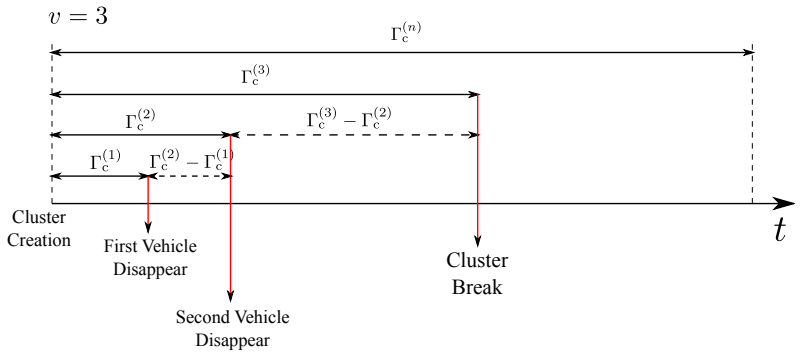


Figure 3.12: Representative example of the cluster evolution process. The cluster lives until a number $v = 3$ of links breaks.

representing the time until v vehicles disconnect from the cluster, one can write:

$$\begin{aligned}
\Gamma_c^{(1)} &= \min\{\Gamma_\ell^{(1)},\ldots,\Gamma_\ell^{(n)}\} = \Gamma_c \\
\Gamma_c^{(2)} &= \min\{\ \{\Gamma_\ell^{(1)},\ldots,\Gamma_\ell^{(n)}\} \setminus \{\Gamma_c^{(1)}\}\ \} \\
&\vdots \\
\Gamma_c^{(v)} &= \min\{\ \{\Gamma_\ell^{(1)},\ldots,\Gamma_\ell^{(n)}\} \setminus \{\Gamma_c^{(1)},\ldots,\Gamma_c^{(v-1)}\}\ \} \\
&\vdots \\
\Gamma_c^{(n)} &= \min\{\ \{\Gamma_\ell^{(1)},\ldots,\Gamma_\ell^{(n)}\} \setminus \{\Gamma_c^{(1)},\ldots,\Gamma_c^{(n-1)}\}\ \}.
\end{aligned} \tag{3.9}$$

Referring to Figure 3.12, one can express $\Gamma_c^{(v)}$ as follows:

$$\Gamma_c^{(v)} = \Gamma_c^{(1)} + (\Gamma_c^{(2)} - \Gamma_c^{(1)}) + (\Gamma_c^{(3)} - \Gamma_c^{(2)}) + \ldots + (\Gamma_c^{(v)} - \Gamma_c^{(v-1)}). \tag{3.10}$$

Observing that $\Gamma_c^{(1)}$ is a log-normal random variable and $\{\Gamma_c^{(i)} - \Gamma_c^{(i-1)}\}_{i=2}^{v}$ are approximately log-normal [66], it follows that $\Gamma_c^{(v)}$ is approximately log-normal. The value of the parameters of the final log-normal distribution can be found in [67]. This means that even assuming that more than one vehicle has to disconnect to make the cluster die, the cluster duration still follows a log-normal distribution.

In Figure 3.13, the empirical PDF of the cluster lifetime, obtained from 500 simulation runs, with $\rho_s z = 7.5$ veh, is compared with a log-normal distribution with optimized parameters. Two definitions of cluster lifetime are considered: $v = 1$ and

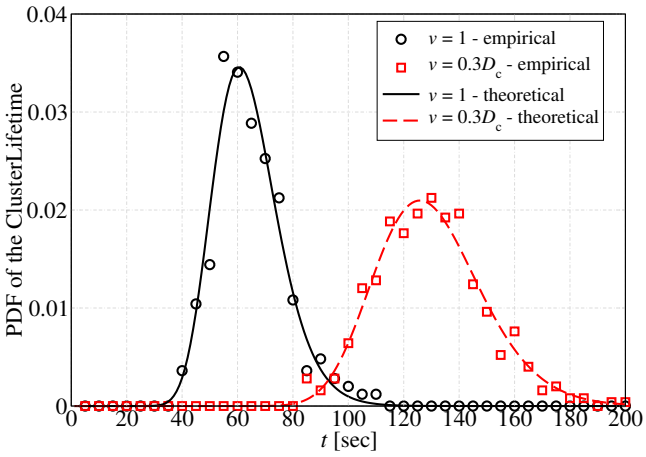


Figure 3.13: Comparison of the empirical PDF of the cluster lifetime with the log-normal theoretical distribution of [1] for $\rho_s z = 7.5$ veh. Two definitions of cluster lifetime are considered: $v = 1$ and v corresponding to 30% of the nodes in the cluster.

v corresponding to 30% of the nodes in the cluster—note that similar results can be obtained for other values of $\rho_s z$. One can observe that the simulation results confirm the theoretical log-normal distribution. In particular, by minimizing the mean square error between the simulation-based results and the log-normal PDF one finds that the parameters of the log-normal distribution are: $\mu_\Gamma = 4.14$ s and $\sigma_\Gamma = 0.19$ s for $v = 1$; and $\mu_\Gamma = 4.86$ s and $\sigma_\Gamma = 0.15$ s for v corresponding to 30% of the number of nodes in the cluster.

Focusing on the results of Figure 3.13, it is worth noting that the cluster lifetime can be quite long. In fact, recalling that

$$\mathbb{E}\{\Gamma_c\} = \exp\left\{\mu_\Gamma + \frac{\sigma_\Gamma^2}{2}\right\} \text{ [s]}$$

the cluster lifetime is approximately 64 s in the case of cluster death after the first link disconnection and 130 s in the case of cluster death after 30% of the links break.

Knowing the statistical distribution of the cluster lifetime, one can compute the total transferable amount of information before the cluster breaks, depending on the

underlying communications/networking protocols. This goes behind the scope of this work and represents an interesting research extension.

3.6.4 Cluster Maintenance and Reclustering

As outlined in Subsection 3.6.3, the cluster duration has a significant impact on the performance of the considered VSN. In order to prolong the overall network lifetime, different mechanisms can be envisioned to maintain the formed clusters and, eventually, to perform reclustering in the presence of vehicles leaving their own clusters. A possible solution is that all CHs periodically transmit, with period T_{ccp} (dimension: [s]), CCPs. A non-CH vehicle receiving a CCP checks if the CCP generator is at a distance shorter than $z/2$: if this is the case, the vehicle joins the new CH and, therefore, nodes exiting a cluster may have the possibility of joining a new one. The choice of the value of T_{ccp} is a tradeoff between a finer reclustering (and, therefore, a smaller information loss from vehicles) and a higher network overhead. One should note that, in the presence of mobility, some vehicles may leave the monitored area and portions of this area may even be uncovered. Since unconnected vehicles cannot participate in the data collection in the VSN, it is necessary to repeat, with period much longer than T_{ccp}, the entire CEIF clustering procedure described in Subsection 3.3.1. This aspect, however, has been neglected in the simulations, since the information loss caused by connectionless regions is negligible during the VSN operational lifetime.

In Figure 3.14, $\overline{\mathscr{D}}_c$ is shown, as a function of time, in the presence of reclustering with $T_{\text{ccp}} = 100$ s. As already observed in Figure 3.11 (a), as time passes $\overline{\mathscr{D}}_c$ decreases due to mobility. However, when CCPs are generated, nodes outside their CHs' transmission ranges can now join a new cluster and the value $\overline{\mathscr{D}}_c$ thus suddenly increases. However, it is worth noting that, especially for the larger values of $\rho_{\text{s}} z$, the recovery procedure becomes less effective when time passes, since the value of $\overline{\mathscr{D}}_c$ immediately after each CCP generation is smaller and smaller. This is due to the fact that a larger value of $\rho_{\text{s}} z$ implies a larger cluster size and, therefore, the probability that a vehicle does not receive any CCP during reclustering increases.

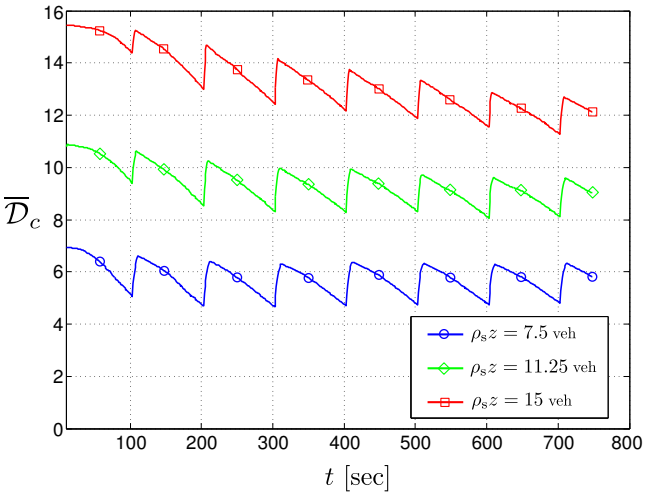


Figure 3.14: $\overline{\mathcal{D}}_c$, as a function of time, in the presence of reclustering with $T_{\text{ccp}} = 100$ s.

3.7 Performance Analysis in Urban Scenarios

In order to further investigate the impact of mobility on the performance of the proposed VSN, in this section a urban-like vehicular scenario is considered. One should note that the CEIF protocol and the corresponding clustering procedure have been designed for linear networks (e.g., highway-like) and are not optimized for bidimensional urban-like scenarios. In this section, the focus is on a simulation-based investigation of the performance of CEIF in urban scenarios: the design of proper communication/networking protocols for such scenarios goes beyond the scope of this work and may be the subject of future research.

The considered scenario is representative of the center of a large city with many road intersections, narrow roads, and single driving directions. Moreover, the nodes' speeds are highly heterogeneous: in fact, vehicles can move fast but they are constrained to abide by the traffic rules (priorities, traffic lights, etc.) forming queues and thus slowing down the overall vehicular traffic mobility. Since this is a highly dynamic scenario, more complex realistic mobility models are needed. To this end, the

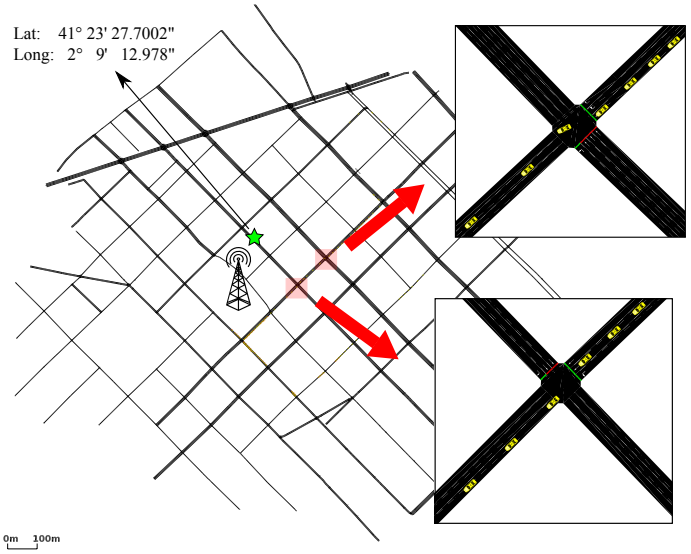


Figure 3.15: Illustrative example of urban VNS scenario: portion of the city center of Barcelona imported into the SUMO mobility simulator

.

Open Street Maps (OSM) tool has been exploited, which provides open and editable maps of real cities to be imported into SUMO [43]. As a representative vehicular scenario, a portion of the city center of Barcelona, shown in Figure 3.15 has been selected. Vehicles move with a speed lower than or equal to the legal limits of the roads, stopping when needed, i.e., in correspondence to traffic lights and priorities. The node range z is set to 250 m and the number of considered vehicles is $N = 100$. At time instant $t_s = 550$ s the cluster formation procedure is triggered by the remote sink which, in this case, is represented by a base station placed in the center of the simulated area. As in the highway scenario considered in Section 3.6, the selected value of t_s guarantees that all vehicles enter the simulated area.

In Figure 3.16, $\overline{\mathscr{D}}_c$ is shown, as a function of time, in the considered urban-like scenario depicted in Figure 3.15. First, it can be noted that, because of mobility, $\overline{\mathscr{D}}_c$ is a decreasing function of time. However, the decaying is faster than in the highway-like scenario shown in Figure 3.9: for instance, after 1 min the average reduction is of

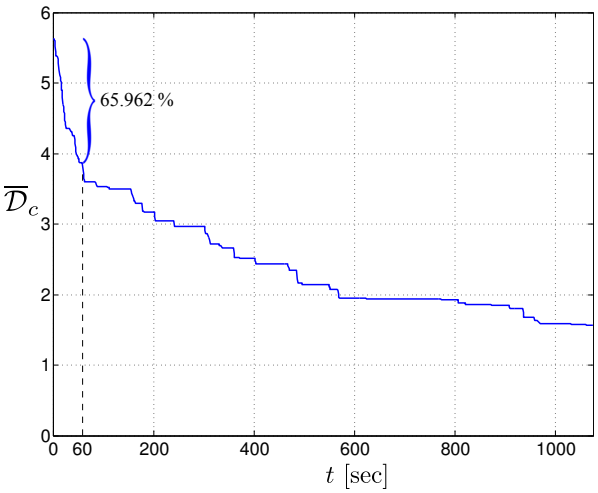


Figure 3.16: $\overline{\mathscr{D}}_c$, as a function of time, in the considered urban-like scenario.

65.962%. This is due to the fact that urban-like mobility is much more dynamic than highway-like mobility. As an example, traffic lights and priorities can be very detrimental for the cluster lifetime, as groups of vehicles may remain stopped in a queue thus leaving their own clusters. This also causes, unlike what happens in highway-like scenarios, the piecewise-constant trend of $\overline{\mathscr{D}}_c$. We can thus conclude that in the presence of urban-like mobility reclustering should be triggered more frequently.

3.8 Concluding Remarks

In this chapter, the performance of decentralized detection schemes for clustered VSNs have been analyzed. Two phases have been envisioned: a downlink phase, during which a novel clustering broadcast protocol, denoted as CEIF, is employed; and an uplink phase, during which the vehicles perform, through the clustered topology, decentralized detection of a spatially constant phenomenon of interest. Different clustered topologies and fusion rules have been considered. The performance of the proposed VSN-based distributed detection schemes has been analyzed in terms of probability of error on the phenomenon estimate and network lifetime, considering

different mobility models (namely, highway-like and urban-like). Unlike clustered sensor networks, where the clustering structure is a design aspect, the proposed vehicular distributed detection schemes exploit the natural formation of ephemeral vehicle clusters. Our results clearly show that the maximum amount of data collectible during the clustered VSN lifetime is more affected by the node mobility level than by the specific clustering structure. This suggest that "bursty" data collection strategies should be considered, together with proper (local) reclustering strategies. This is especially relevant in urban-like scenarios.

Chapter 4

Multi-hop Broadcast Communications in Pedestrian Ad-hoc Networks

4.1 Introduction

As already anticipated in Section 2.4.1, in Pedestrian Ad-hoc NETworks (PANETs) nodes are represented by devices carried by pedestrians (e.g., hand-held smartphones or wearable devices). The main characteristics of this kind of ad-hoc network are: (i) presence of many information sources; (ii) high node spatial density; and (iii) low node speed (almost static). The main focus of this chapter is related to PANETs applications where nodes send very small amounts of information data (e.g., geographical coordinates or alert messages) to all other nodes via multi-hop transmissions. This is meaningful, for example, for proximity-based social networking applications.

Also in this case, the probabilistic broadcasting with IF represents a possible solution. In particular, as seen in Chapter 2, IF is able to improve the efficiency of multi-hop broadcasting with respect to flooding by saving retransmissions effectively selecting the best rebroadcasters. However, IF has been originally designed to work into monodimensional networks (highway-like VANETs) and suffers from intrinsic

inefficiencies when exploited in scenarios where nodes have a bidimensional distribution.

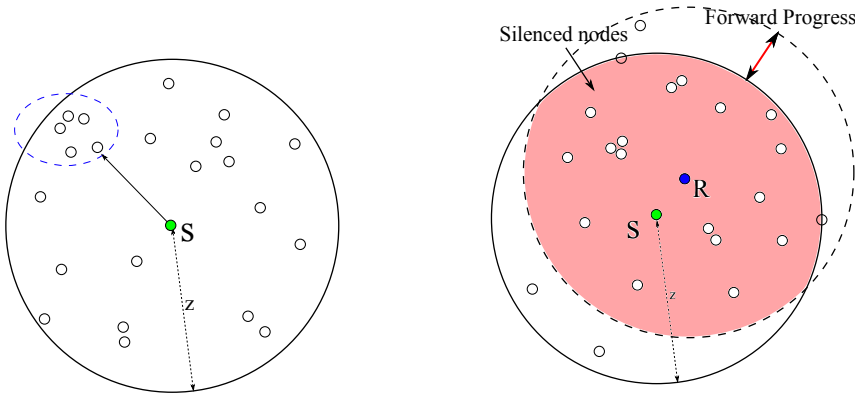
In this chapter, a novel probabilistic broadcasting strategy, denoted as Distance-based Silencing Irresponsible Forwarding (DiSIF) is proposed. This strategy "stems" from IF and tries to address its inefficiencies through a novel silencing mechanism which effectively selects rebroadcasters by introducing an initial contention phase. In order to have a comparison benchmark for the performance of the considered protocols, a lower bound for the average number of hops performed in a single multi hop communication route is derived.

The remainder of this chapter is organized as follows. In Section 4.2, a first simple extension of the IF protocol, denoted as silencing IF (SIF) is briefly introduced. In Section 4.3 the DiSIF protocol, with its novel embedded silencing mechanism, is presented. In Section 4.4, a lower bound on the average number of hops along a single communication route, together with a simpler approximation, is provided. In Section 4.5, the system performance is investigated. Finally, Section 4.6 concludes the chapter.

4.2 Silencing Irresponsible Forwarding (SIF)

SIF is an extension of IF which introduces the concept of *silencing*. According to the silencing mechanism, when a rebrocaster node (say node r) receives a packet with a certain UID (say u) it first checks its transmission queue: if a packet with UID u is found (i.e., it has already been received but has not been retransmitted, yet), it is removed from the queue, as another neighbor node has already transmitted the same packet. In this way, node r is silenced for the transmission of the packet with UID u. The use of silencing corresponds to the fact that the “fastest” retransmitter (among the set of those which have decided to retransmit) silences the others. In other words, even if a node decides to retransmit a packet, it may refrain from doing so if the same packet has already been retransmitted by another node.

The rationale behind the "fast" silencing technique used by SIF is to limit the potentially large number of rebroadcasts brought by IF. As will be show in the next



(a) (b)

Figure 4.1: Representative examples of inefficiencies of (a) IF and (b) SIF.

section, this problem is more critical in a dense, multi-source, bidimensional network.

4.3 Distance-based Silencing Irresponsible Forwarding (DiSIF)

4.3.1 The Dark Force: Shortcomings of IF and SIF

Although IF and SIF may perform well also in bidimensional PANET scenarios—as shown in [69]—they suffer from intrinsic inefficiencies.

With the IF PAF given in (2.1), two nodes at the same distance from the source have the same retransmission probability. Therefore, it is possible that a group of neighboring nodes, located at the boundary of the coverage area of the source, simultaneously rebroadcast the packet, possibly colliding. A representative example of this situation, which is likely to appear in dense PANET scenarios, is shown in Fig. 4.1 (a).

SIF does not incur IF's collision risk described in the previous paragraph, as the silencing technique guarantees that only the fastest rebroadcaster retransmits the packet. However, the fastest rebroadcaster may not be the "best" rebroadcaster: since SIF is a probabilistic forwarding protocol, a node close to the source may choose

to retransmit the packet silencing many potential rebroadcasters farther from the source. This, in turn, results in a low forward progress and may prevent the originally transmitted packet from propagating in some directions. A representative example, in which almost all potential rebroadcasters are silenced by a node close to the source, is shown in Fig. 4.1(b), where S is the source and R is the rebroadcaster.

4.3.2 A New Hope: DiSIF

The DiSIF protocol implements, through an initial contention phase, a novel silencing technique which is more efficient, in bidimensional multi-source networks, than the silencing technique embedded into SIF. More specifically, this new silencing mechanism guarantees that the “farthest” (instead of the “fastest”) rebroadcaster silences the others.

Two types of packets are defined in the DiSIF protocol: (i) DATA packets, which contain the information to propagate; and (ii) Probe Packets (PPs), which are short control packets used by the DiSIF silencing mechanism. In a single source scenario, the forwarding process of DiSIF can be summarized as follows. At a generic instant t_0 , the source node s transmits a DATA packet. As already seen for the IF protocol, node s puts in the packet header its own position POS_{s} and the UID pair given by $(SN, ADDR_{\mathrm{s}})$. The packet is then received by the source neighbors at a time $t_1 = t_0 + \varepsilon$, where ε is the (average) propagation time.[1] Upon reception of the DATA packet, a source neighbor, after checking the UID of the packet, has two options: the packet is dropped if already received earlier; otherwise, the neighbor starts "competing" with the other neighbors to designate a set of rebroadcasters. In particular, each neighbor node elects itself as a *candidate rebroadcaster* with a probability given by (2.1). We denote as n_{cand} the number of candidate rebroadcasters. A candidate rebroadcaster node k ($k \in \{1, \ldots, n_{\mathrm{cand}}\}$) schedules a transmission of a PP bearing its distance, denoted as d_{ks}, from the source together with the same UID $(SN, ADDR_{\mathrm{s}})$ of the received DATA packet). The transmission of the PP is scheduled by each candidate

[1] Note that ε may vary from neighbor to neighbor but, for the sake of simplicity, it is assumed that is equal for all neighbors. This corresponds to considering an equivalent average propagation time and is reasonable in dense PANETs (e.g., set of hand-held smartphones in a crowd).

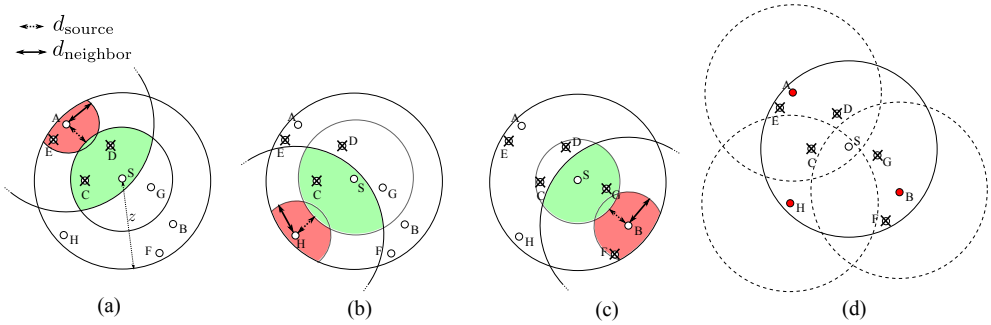


Figure 4.2: Representative example of the situation, in the 1-st hop, due to the DiSIF propagation process.

rebroadcaster at a time randomly distributed between t_1 and $t_1 + t_{\text{wait}}/2$, where t_{wait} is a DiSIF parameter to be properly optimized. Another candidate rebroadcaster, say node j ($j \in \{1,\ldots,n_{\text{cand}}\} \setminus \{k\}$), which receives the PP sent from node k, silences itself if at least one of the two following conditions apply:

$$d_{js} + d_{\text{source}} < d_{ks} \tag{4.1a}$$

$$d_{kj} < d_{\text{neighbor}} \tag{4.1b}$$

where d_{source} (dimension: [m]) and d_{neighbor} (dimension: [m]) are parameters of the DiSIF protocol. The rationale behind condition (4.1a) is that if a candidate rebroadcaster is sufficiently closer, with respect to another candidate rebroadcaster, to the source it should silence itself because its retransmission will be redundant. Condition (4.1b) guarantees that two candidate rebroadcasters, which are sufficiently close to each other, do not simultaneously rebroadcast the packet, thus decreasing the collision probability.

When a candidate rebroadcaster silences itself, it removes the transmission of its PP (if still to be sent) and ignores all other future received PPs with the same UID. At time $t_1 + t_{\text{wait}}$, each unsilenced candidate rebroadcaster retransmits the DATA packet and the process repeats recursively.

In Fig. 4.2, an illustrative example of the 1-st hop in the DiSIF propagation pro-

cess with single source is given. All nodes are assumed to have the same transmission range z. Even though the DATA packet transmitted by the source, denoted as S, is received by all neighbors, for the sake of graphical clarity only the candidate rebroadcasters are shown. Moreover, since this is a single-source, single-packet propagation example, the UID of DATA and PP packets is omitted, as it is always equal to $(1,\mathrm{S})$. In Fig. 4.2 (a), node A is the first candidate rebroadcaster which transmits a PP—note that condition (4.1a) leads to silencing all candidate rebroadcasters in the green area, while condition (4.1b) corresponds to silencing all candidate rebroadcasters in the red area. Node A is far from the source and silences nodes C and D as $d_{\mathrm{CS}} + d_{\mathrm{source}} < d_{\mathrm{AS}}$ and $d_{\mathrm{DS}} + d_{\mathrm{source}} < d_{\mathrm{AS}}$. Node E is also silenced because $d_{\mathrm{AE}} < d_{\mathrm{neighbor}}$. In Fig. 4.2 (b), since node H transmits its PP, node C would be silenced once more, as $d_{\mathrm{CS}} + d_{\mathrm{source}} < d_{\mathrm{HS}}$: however, since node C had already been silenced by node A, it drops the PP received from node H. Finally, in Fig. 4.2 (c) node B transmits its PP and silences nodes G and F. As shown in Fig. 4.2 (d), at the end of the contention phase only nodes A, B, and H are still unsilenced, so that they proceed to rebroadcast the DATA packet. The process then repeats recursively. In this example, for the sake of simplicity, the farthest nodes from the source (i.e., A, H, and B) first transmit the PP silencing the nearest nodes (i.e., C, D, G, and E). However, even if one of the nearest (to the source) nodes, decides to first transmit the PP, the farthest nodes would not be silenced because condition (4.1a) would not be fulfilled and, even in this case, the nearest node will be silenced by a subsequent PP transmission by one of the farthest node.

It is worth noting that, considering a more general multi-source scenario, since both DATA and PP packets can be uniquely identified, each packet dissemination process is independent of the others. For example, in Fig. 4.3 two source nodes are present, denoted as S_1 and S_2, which both transmit DATA packets to the same node A: the transmitted DATA packets have UIDs $(1,\mathrm{S}_1)$ and $(1,\mathrm{S}_2)$, respectively. When node A receives the packet from node S_1, it becomes a candidate rebroacaster for the packet with UID $(1,\mathrm{S}_1)$ and enters in the contention phase, scheduling the transmission of a PP bearing its distance from S_1 and with UID equal to $(1,\mathrm{S}_1)$. When node A receives the packet coming from node S_2, it becomes a candidate rebroadcaster also for the

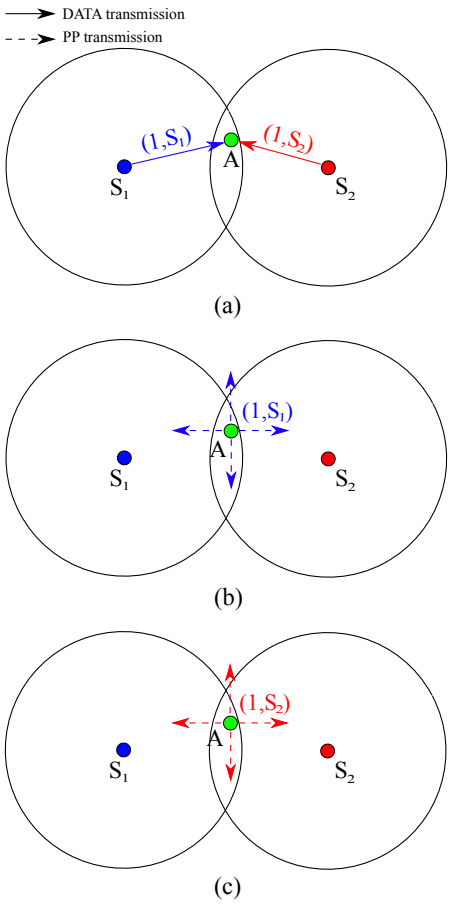


Figure 4.3: Representative example of DiSIF propagation with two sources of information: (a) nodes S_1 and S_2 send DATA packet to A; (b) node A first rebroadcasts the PP with UID (1,S_1); (c) node A rebroadcasts the PP with UID (1,S_2).

packet with UID $(1, S_2)$ and puts in its transmission queue a PP bearing its distance from S_2 and with UID equal to $(1, S_2)$. After transmitting both PPs, if A has not been silenced by other nodes, it rebroadcasts both the DATA packet with UID $(1, S_1)$ and the DATA packet with UID $(1, S_2)$.

The main goal of DiSIF's forwarding strategy is to reduce the number of performed hops during broadcast propagation. This, in turn, increases the propagation efficiency by decreasing the channel contention level and, therefore, collisions. This policy is used in other existing multi-hop forwarding strategies: for example, while DiSIF tries to maximize the distance between transmitting node and rebroadcaster, in [71] the GeRaF protocol tries to minimize the distance between rebroadcaster and final destination. In order to quantify the propagation efficiency of DiSIF, in the following Section 4.4 analytical lower bounds on the average number of hops in a point-to-point (i.e., unicast) communication route are derived.

4.4 Lower Bounding the Average Number of Hops in a Unicast Communication Route

4.4.1 An Exact Lower Bound

In a unicast multi-hop communication, the number of hops is given by the number of relay nodes between source and destination plus one. In the PANET scenario of interest, denoting with n the number (fixed) of nodes in the network, a single broadcast from one of the n nodes can be modeled as $n-1$ multi-hop unicast transmissions. Defining as N_{hop} (random variable)[2] the number of hops of one of these unicast multi-hop communication routes, its expected value $\mathbb{E}\{N_{\text{hop}}\}$ is a relevant metric to evaluate the propagation efficiency of a multi-hop broadcasting protocol.

In order to derive a lower bound on $\mathbb{E}\{N_{\text{hop}}\}$, an "ideal" unicast multi-hop communication route is considered in which: (i) relay nodes lay on the straight line between source and destination; and (ii) a relay node i is on the boundary of the node

[2] In the following, all random variables are denoted with uppercase letters while other variables are denoted with lowercase letters.

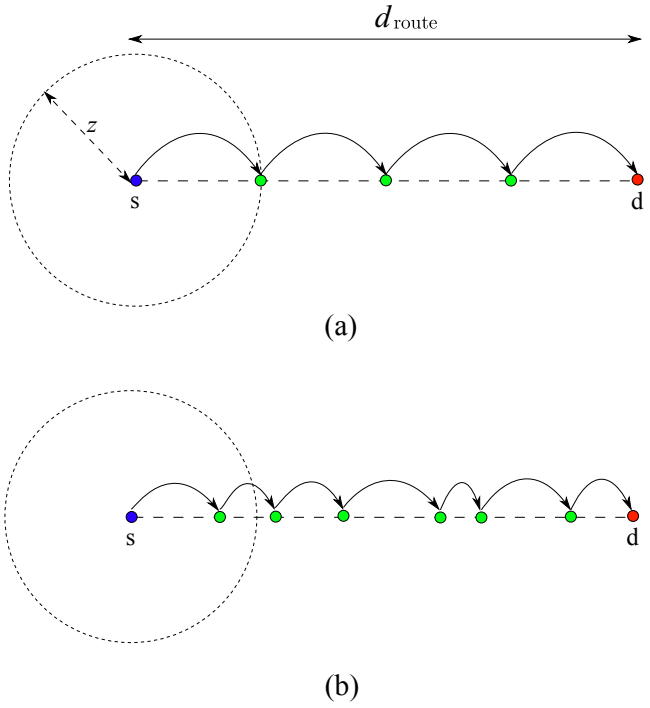


Figure 4.4: Representative example of ideal multi hop communications: (a) relay nodes on the boundary of the node range of the previous relay node; (b) relay nodes randomly deployed between source and destination.

range of the previous relay node $i-1$ ($i \in \{2,\ldots,N_{\text{hop}}\}$) or the source ($i=1$). In this scenario, which is shown in Fig. 4.4 (a), the number of hops is minimized and can be expressed as d_{route}/z, where d_{route} (dimension: [m]) is the distance between source and destination and z (dimension: [m]) is the already introduced *fixed* transmission range. Therefore, one can write:

$$\mathbb{E}\{N_{\text{hop}}\} \geq \frac{\mathbb{E}\{D\}}{z} \tag{4.2}$$

where D is the distance (random variable) between node pairs (i.e., the source-destination pairs) in the PANET—in particular, d_{route} in Fig. 4.4 (a) corresponds to a realization of D.

The lower bound at the right-hand side of (4.2) depends on the statistical distribution of D. Assuming that nodes are randomly deployed over a square region with side ℓ, the corresponding Cumulative Distribution Function (CDF) of D is derived in [83], from which the following Probability Density Function (PDF) follows:

$$f_D(\delta) = \begin{cases} 0 & \delta < 0 \\ \frac{2\pi\delta}{\ell^2} - \frac{8\delta^2}{\ell^3} + \frac{2\delta^3}{\ell^4} & 0 \le \delta < \ell \\ \frac{\delta}{\ell^4}\Big[4\ell^2\sqrt{b-1} + \frac{2\ell^2}{\sqrt{b-1}} + & \\ \quad + \frac{2\delta^2}{\sqrt{b-1}} + \frac{2(l^2-\delta^2)^2}{\ell^2\sqrt{(b-1)^3}} + & \\ \quad -4\ell^2 \arcsin(\frac{b-2}{b}) - 4\ell^2 + & \\ \quad - \frac{4\ell^4}{\sqrt{\ell^2(\delta^2-\ell^2)}} - 2\delta^2\Big] & \ell \le \delta < \sqrt{2}\ell \\ 0 & \delta \ge \sqrt{2}\ell \end{cases} \tag{4.3}$$

where $b \triangleq \delta^2/\ell^2$. In Fig. 4.5(a), the PDF $f_{\mathrm{D}}(\delta)$ is shown for two different values of ℓ: 100 m and 200 m. In both cases, the analytical PDF is compared with the PDF obtained through Matlab simulations. As can be observed, analytical and simulated PDFs almost overlap, thus validating the expression (4.3) for $f_D(\delta)$. The maximum value of the internode distance is obviously $\sqrt{2}\ell$, which corresponds to the length of the diagonal of the square region. After reaching a maximum value, the PDF rapidly decreases and the probability that two nodes lie more than ℓ meters apart is very low (lower than 0.029 and 0.024 when ℓ is equal to 100 m and 200 m respectively).

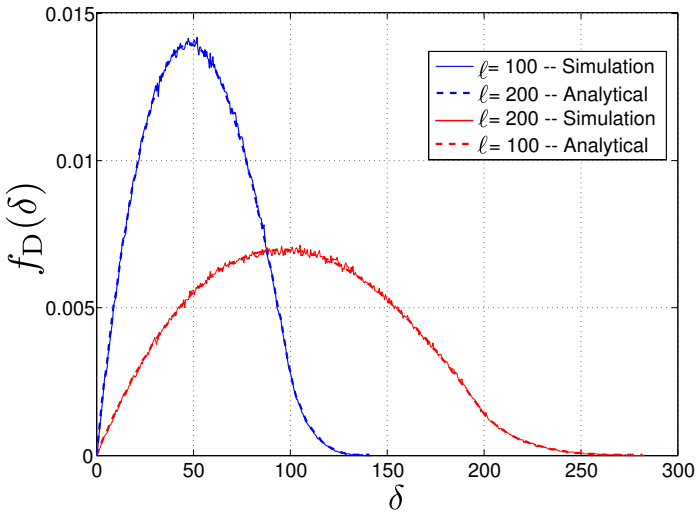


Figure 4.5: (a) PDF of D for $\ell = 100$ m and $\ell = 200$ m. Analytical and simulation-based curves are compared.

The average value of D follows directly from (4.3):

$$\begin{aligned}\mathbb{E}\{D\} &= \int_0^{\ell\sqrt{2}} \delta f_D(\delta)\,\mathrm{d}\delta \\ &= \frac{4+2\sqrt{2}+10\ln(1+\sqrt{2})}{30}\,\ell. \end{aligned} \tag{4.4}$$

By replacing (4.4) at the right-hand side of (4.2), the following lower bound for $\mathbb{E}\{N_{\text{hop}}\}$ is obtained:

$$\begin{aligned} LB_{\text{nhop}} \triangleq \frac{\mathbb{E}\{D\}}{z} &= \frac{4+2\sqrt{2}+10\ln(1+\sqrt{2})}{30}\,\frac{\ell}{z} \\ &\simeq \frac{\ell}{2z}. \end{aligned} \tag{4.5}$$

In Fig. 4.6, the lower bound LB_{nhop} is shown, as a function of ℓ, for $z = 83$ m. Analytical and simulation results are compared. As predicted by (4.5), LB_{nhop} is a linearly increasing function of ℓ.

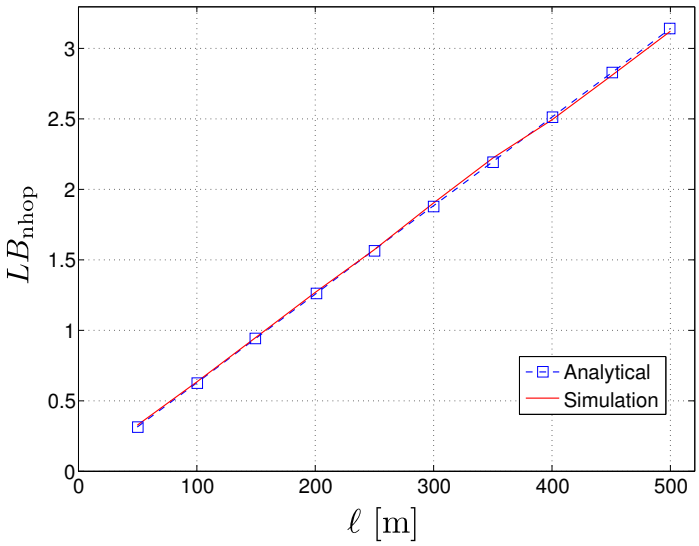


Figure 4.6: LB_{nhop}, as a function of ℓ, with $z = 83$ m.

4.4.2 An Approximate Lower Bound

One of the assumptions behind the ideal multi-hop communication route considered in Subsection 4.4.1 is that two consecutive relay nodes are z meters apart from each other, i.e., at the maximum possible distance. This is a strong assumption for probabilistic forwarding protocols, as even a node which is close to the source may decide to rebroadcast the packet as shown in Fig. 4.4 (b). Therefore, this assumption is now relaxed. When a rebroadcaster, say the i-th with coordinates (x_i, y_i), transmits a packet, all nodes in the circle centered in (x_i, y_i) with radius z receive the packet: each of these nodes is a potential rebroadcaster for the next hop. In order to realistically evaluate the distance between relay nodes, the expected value of the random variable given by the distance between i and one of its neighbor, denoted as R, can be derived. Then, by replacing z with $\mathbb{E}\{R\}$ in (4.2), an approximate lower bound for $\mathbb{E}\{N_{\text{hop}}\}$ can be derived—this approximation is no longer an "exact" lower bound. For analytical tractability, it is convenient to define a coordinate system with its origin at (x_i, y_i) (which thus become $(0,0)$) so that the set of coordinates

$\mathscr{C} = \{(x,y) : x^2 + y^2 \leq z^2\}$ represents the coverage area of node i. The CDF of R, denoted as $F_R(r) = P\{R \leq r\}$, can thus be computed as the probability that a node lies in the area $\mathscr{E} = \{(x,y) : x^2 + y^2 \leq r^2\}$, so that one can write:[3]

$$F_R(r) = \frac{\text{Area}(\mathscr{E})}{\text{Area}(\mathscr{C})} = \frac{\pi r^2}{\pi z^2} = \left(\frac{r}{z}\right)^2 \tag{4.6}$$

and, consequently:

$$f_R(r) = \frac{\mathrm{d}F_R(r)}{\mathrm{d}r} = \frac{2r}{z^2}. \tag{4.7}$$

The expected value of R is:

$$\begin{aligned} \mathbb{E}\{R\} &= \int_{-\infty}^{\infty} r f_R(r)\,\mathrm{d}r \\ &= \int_0^z r f_R(r)\,\mathrm{d}r = \frac{2}{3}z. \end{aligned} \tag{4.8}$$

Finally, by replacing z (maximum hop length) with $\mathbb{E}\{R\}$ in (4.2), the following approximate bound is obtained:

$$\mathbb{E}\{N_{\text{hop}}\} \gtrapprox \frac{\mathbb{E}\{D\}}{\mathbb{E}\{R\}} = \frac{3\mathbb{E}\{D\}}{2z} \triangleq LB_{\text{approx}} \tag{4.9}$$

where $\gtrapprox$ stands for "on the order of 1 or greater than" and its use can be motivated as follows. It cannot be claimed that (4.9) is an exact lower bound, as $\mathbb{E}\{D\}/\mathbb{E}\{R\}$ is not necessarily smaller than the average value of the number of hops. Intuitively, however, the larger is the average number of hops in a multi-hop route, the more accurate is the use of $\mathbb{E}\{R\}$ to estimate the average length of every hop (statistical regularity). This motivates one to consider the approximation sign "$\approx$" in (4.9). Moreover, the fact that in all scenarios considered in Section 4.5, (4.5) will be a loose lower bound (it is very idealistic), whereas (4.9) will be almost always a tighter lower bound, motivates the use of the strict inequality sign "$>$" in (4.9). Overall, the choice of the approximate inequality notation $\gtrapprox$ in (4.9) seems the most appropriate.

[3]Expression (4.6) underlies the implicit assumption that within the coverage area of a node there is at least one other node, i.e., there is no disconnected node.

Owing to (4.5), the approximate lower bound in (4.9) can be further expressed as follows:

$$\begin{aligned} LB_{\text{approx}} &= \frac{3}{2} LB_{\text{nhop}} \\ &= \frac{12+6\sqrt{2}+30\ln(1+\sqrt{2})}{60} \frac{\ell}{z} \\ &\simeq \frac{3\ell}{4z}. \end{aligned} \tag{4.10}$$

In order to keep the terminology simple, in the next section (4.10) will be referred as approximate lower bound.

4.5 Performance Analysis

4.5.1 Simulation Setup

All simulations are carried out with the well known discrete-event network simulator ns-3 (ns-3.19) [39]. In all simulated scenarios, the following assumptions hold. As already anticipated in Subsection 4.4.1, nodes are assumed to be randomly deployed over a square region of side $\ell = \sqrt{n/\rho}$.

All nodes move according to the random way point mobility model, available in the ns-3 simulator, with average speed s_{p} (dimension: [m/s]). Each node has the same transmission range z. Regarding the DiSIF protocol parameters, in all simulations: $t_{\text{wait}} = 5$ ms, $d_{\text{neighbor}} = z/3$, $d_{\text{source}} = z/10$, and $c = 1$. The number of source nodes in the network is n_{tx} and each of them transmits a burst of n_{p} DATA packets of fixed size p_{s} (dimension: [byte/pkt]). Packets are generated at a fixed rate λ (dimension: [pkt/sec]) and transmitted at a fixed data rate (on the wireless channel) equal to 1 Mbps. The performance of the proposed broadcasting protocol DiSIF is compared with those of flooding [28], GOSSIP [31], IF [29], and SIF [68]. For the MAC and physical layers, the wireless communication protocol stack defined by the ad-hoc IEEE 802.11 b standard is used [8].

The considered performance metrics are listed in Table 4.1 and are shortly described in the following. The Packet Delivery Ratio (*PDR*) is defined as the global

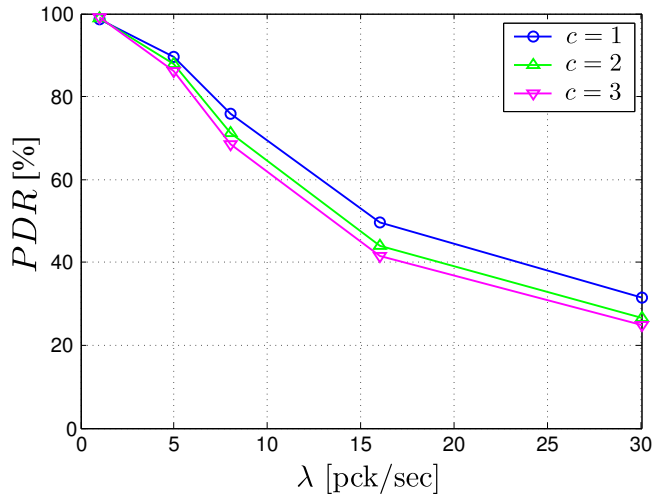
Metric	Symbol	Dimension
Packet Delivery Ratio	PDR	[adimensional]
Average number of hops	$\mathbb{E}\{N_{\text{hop}}\}$	[hop]
Average number of retransmissions per generated packet	$n_{\text{broad/pck}}$	[pck]
Average end-to-end delay	DEL	[sec]

Table 4.1: Considered performance metrics.

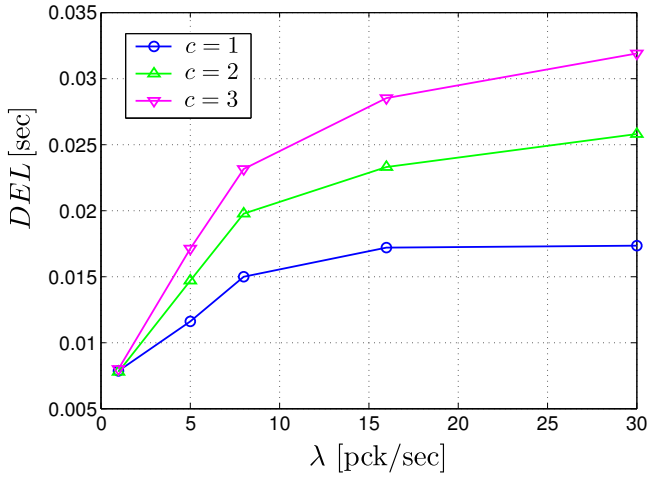
percentage (with respect to the total number of transmitted packets) of correctly received DATA packets. The average number of hops $\mathbb{E}\{N_{\text{hop}}\}$ in a communication route, already introduced analytically in Section 4.4, is obtained by averaging over all communication routes in network. The average number $n_{\text{broad/pck}}$ of retransmissions triggered by a single packet generation (at its source) is given by the ratio between the total number of broadcast transmission acts in the network and the total number of generated packets $n_{\text{p}}n_{\text{tx}}$. The *per-packet* end-to-end delay is defined as the time during which a single DATA packet stays in the network, from the generation instant at its source till the reception instant at its destination. The *average* end-to-end delay (DEL) is obtained as the arithmetic average of the end-to-end delays of all correctly received DATA packets.

4.5.2 Simulation Results

Before starting with a comparative performance evaluation of the DiSIF protocol with other existing broadcast protocols, the impact of the parameter c on its performance is first analyzed. In Fig. 4.7, the performance of DiSIF, in terms of (a) PDR and (b) DEL, is shown as a function of the packet generation rate λ, considering various values of the parameter c. The main network parameters are set as follows: $p_{\text{s}} =$



(a)



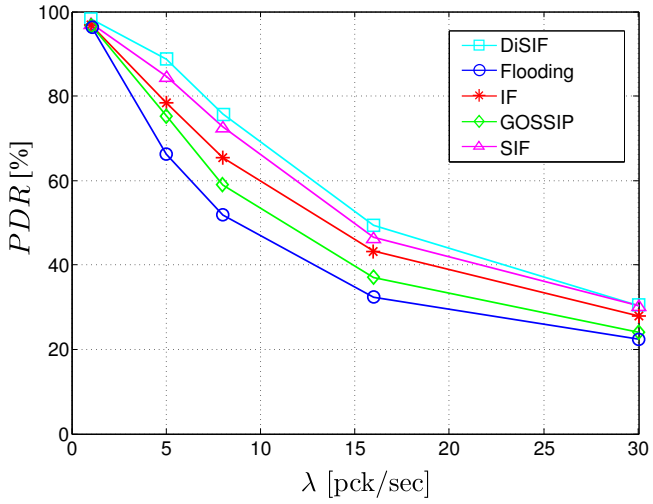
(b)

Figure 4.7: (a) *PDR* and (b) *DEL* as functions of the packet generation rate. The performances of the DiSIF protocol with various values of the parameter c are considered. The network parameters are set as follows: $p_{\mathrm{s}} = 128$ byte/pkt, $\rho = 2000$ nodes/km^2, $n = 150$ nodes, $n_{\mathrm{tx}} = 30$ nodes, and $s_{\mathrm{p}} = 1.5$ m/s.

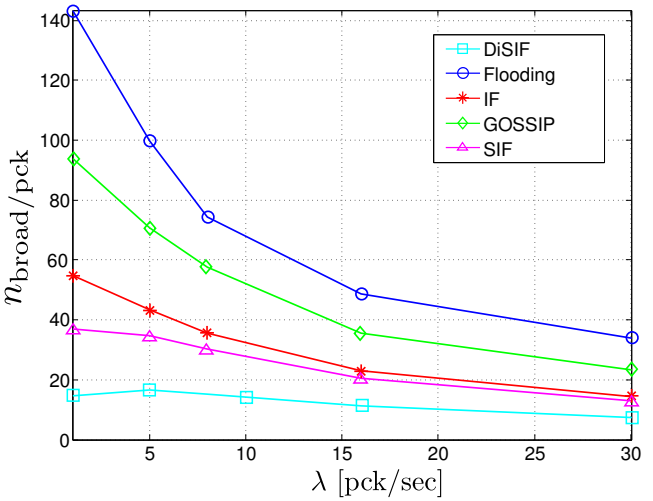
128 byte/pkt, $\rho = 2000$ nodes/km^2, $n = 150$ nodes, $n_{\text{tx}} = 30$ nodes, and $s_{\text{p}} = 1.5$ m/s. Focusing on Fig. 4.7 (a), it can be observed that, for low values of λ, DiSIF has a high *PDR* and low *DEL*. Increasing the packet generation rate leads to an increase of the overall channel contention level, which, in turn, decreases the *PDR*. In Fig. 4.7 (b), a performance degradation, in terms of *DEL*, can be observed for increasing values of λ. In particular, after a "threshold" value of λ (approximately, 15 pck/sec), the network enters into a "saturation" regime: this happens when the transmission queue of each node in the network is full. In these settings, increasing further λ leads the nodes to drop generated packets and does not change the protocol performance in terms of *DEL* (it is worth noting that only the correctly received DATA packets are considered in order to compute *DEL*).

Since the results in Fig. 4.7 (a) and Fig. 4.7 (b) show that the best performance is obtained for $c = 1$, this value will be used in the following simulations. Note that values of c lower than 1 have not been considered because, in this case, even if the protocol obtains a slight performance improvement for large values of λ, the small number of rebroadcasters prevents the packets to reach all nodes even for low values of the packet generation rate.

In Fig. 4.8, (a) *PDR* and (b) $n_{\text{broad/pck}}$ are shown as functions of the packet generation rate λ. The performances of all considered protocols are directly compared and the main network parameters are set as in Fig. 4.7. First of all, it can be observed that DiSIF performs better than all other considered broadcasting protocols for all the values of the packet generation rate. Similarly to what has been observed in Fig. 4.7 (a), the results in Fig. 4.8 (a) show that for very small values of λ the *PDR* is almost 100% for all simulated protocols. In these settings, where nodes can transmit one at a time and there are no concurrent transmissions, even flooding performs well. However, for increasing values of λ, the inefficient use, by flooding, of the radio channel rapidly degrades the performance because of highly redundant transmissions, which lead to collisions. Conversely, the DiSIF protocol guarantees a *PDR* over 80% for $\lambda \leq 7$ pkt/sec, owing to its limited redundancy, which corresponds to a better occupation of the radio channel. For high values of λ, transmissions are highly overlapped and interference and collisions become critical for all protocols.



(a)



(b)

Figure 4.8: (a) *PDR* and (b) $n_{\text{broad/pck}}$ as functions of the packet generation rate. The performances of various broadcasting protocols are directly compared. The network parameters are set as follows: $p_{\text{s}} = 128$ byte/pkt, $\rho = 2000$ nodes/km^2, $n = 150$ nodes, $n_{\text{tx}} = 30$ nodes, and $s_{\text{p}} = 1.5$ m/s.

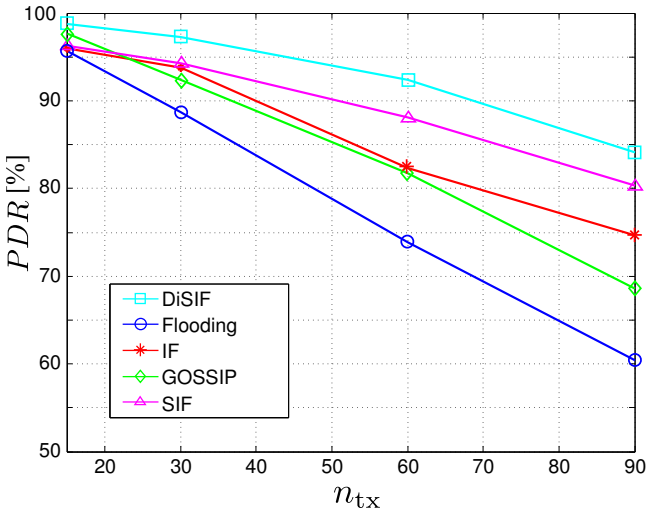


Figure 4.9: *PDR* as a function of n_{tx}. Network parameters are set as follows: $p_{\mathrm{s}} = 128$ byte/pkt, $\rho = 2000$ nodes/km^2, $n = 150$ nodes, $\lambda = 0.5$ pkt/s, $s_{\mathrm{p}} = 1.5$ m/s, and $c = 1$. Various broadcasting protocols are directly compared.

The results in Fig. 4.8 (b) show that, increasing the overall network load, each packet generation (at its source) corresponds to a progressively smaller number of rebroadcasts. This is due to the fact that, because of collisions, generated packets are not delivered to all nodes in the network, so that the nodes act as rebroadcasters less frequently. In other words, the multi-hop propagation is limited by collisions and this is confirmed by the *PDR* results in Fig. 4.8 (a). By comparing Fig. 4.8 (a) with Fig. 4.8 (b), it can be observed that for medium-low values of λ, i.e., in the operative conditions of PANETs, the DiSIF protocol guarantees higher values of *PDR* by performing a consistently smaller number of rebroadcasts. This confirms the effectiveness of the new silencing mechanism, which selects the best rebroadcasters among all neighbors, reducing the channel contention level. It is worth noting that DiSIF, by reducing the number of rebroadcasts, also reduces the overall energy consumption, which is an important issue, in particular when nodes are energy-constrained [84, 85].

The presence of many information sources is a crucial aspect in PANETs, because

large values of n_{tx} can drastically increase the overall network load. In order to study the impact of the number of sources on the system performance, in Fig. 4.9 the *PDR* is shown as a function of n_{tx}. The main network parameters are set as in Fig. 4.7, in particular, the value of the packet generation rate is set to $\lambda = 0.5$ pkt/s. It can be observed that, for all the considered values of n_{tx}, DiSIF outperforms all other protocols. Note that the performance of flooding degrades for large values of n_{tx}, regardless of the relatively small value of λ ($\lambda = 0.5$ pkt/s). This underlines that a large number of source nodes (typical of PANETs) may lead to a high network load even for low values of the per-node packet generation rate. Therefore, efficient (non-redundant) management of the radio channel is crucial in PANETs.

As mentioned in Section **??**, the IF strategy takes into account the node spatial density ρ. This feature allows IF to adapt itself to the network conditions and is inherited by DiSIF. In order to get more insights about this adaptivity, in Fig. 4.10 the *PDR* is shown as a function of the node spatial density. The main network parameters are set as follows: $p_{\text{s}} = 128$ byte/pkt, $n_{\text{tx}} = 80$ nodes, $\lambda = 0.5$ pkt/s, and $s_{\text{p}} = 1.5$ m/s. The node spatial density ρ is varied by keeping constant the side length ℓ of the square network region and varying the number n of nodes in the network. Increasing ρ increases the channel contention level, as more and more nodes are within the transmission range of each other. This directly results in a higher collision probability with flooding and the performance degrades rapidly. Conversely, it can be observed that DiSIF can adapt effectively its behavior to the network conditions.

In addition to packet generation rate and node spatial density, another important network parameter is the packet size p_{s}. This parameter is critical because longer packets correspond to longer transmission times, which increase the collision probability. In order to study the impact of increasing values of p_{s}, in Fig. 4.11 (a) *PDR* and (b) *DEL* are shown as functions of the packet size. The main network parameters are set as follows: $\rho = 2000$ nodes/km^2, $n = 200$ nodes, $\lambda = 1$ pkt/s, $s_{\text{p}} = 1.5$ m/s, $n_{\text{tx}} = 80$ nodes, and $c = 1$. In Fig. 4.11 (a), it can be observed that increasing values of p_{s} lead to a rapid performance degradation of IF, SIF, flooding, and GOSSIP. This is due to the fact that, with longer packets, a transmission act "captures" the channel for a longer time, resulting in a higher collision probability. Conversely, the perfor-

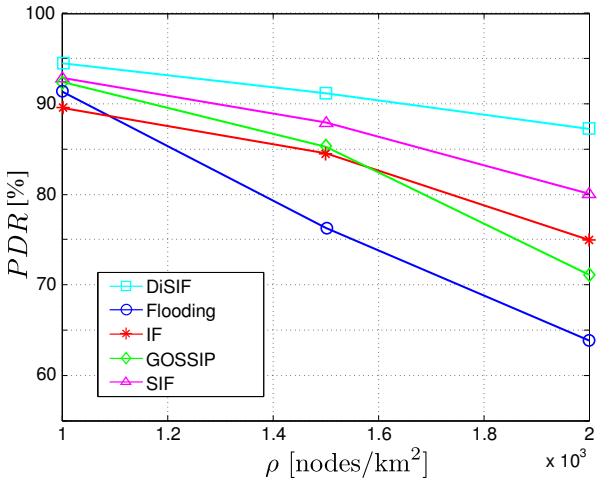
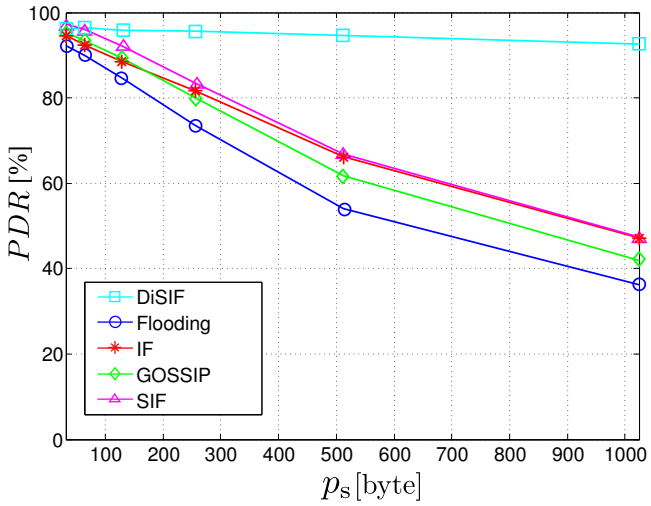
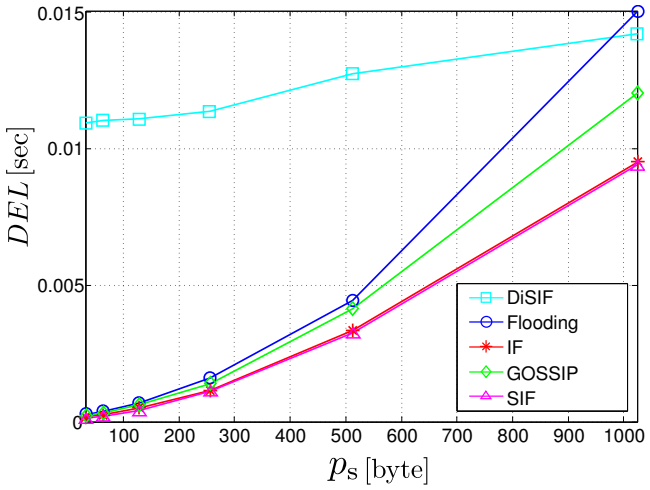


Figure 4.10: *PDR* as a function of the nodes spatial density. Network parameters are set as follows: $p_s = 128$ byte/pkt, $n_{tx} = 80$ nodes, $n = 200$ nodes, $\lambda = 0.5$ pkt/s, $s_p = 1.5$ m/s, and $c = 1$. Various broadcasting protocols are directly compared.

mance of DiSIF is not affected by the increase of the packet dimension. This result is mostly related to the silencing mechanism of DiSIF which, in its first (contention) phase, allows the transmission of only PPs in order to select the actual rebroadcasters. These packets are very short, so that they "capture" the channel for a short time and collisions among them are unlikely. Then, after t_{wait}, only the actual rebroadcasters transmit (longer) DATA packets. These nodes, owing to the silencing mechanism of DiSIF, are typically far away from each other (see Fig. 4.2), so even long DATA packets do not generate interference and the performance of DiSIF remains roughly the same for increasing values of p_s. It is worth noting that, since the contention phase of DiSIF slows down rebroadcasting of DATA packets, the drawback of this strategy is an overall increase of the end-to-end delay *DEL*. This aspect is confirmed by the results in Fig. 4.11 (b), which show that the delay of DiSIF is higher that those of the other protocols. However, at high values of p_s, DiSIF outperforms flooding. In general, the delay of DiSIF is acceptable (e.g., below 5 ms) for PANET-based "social" applications.



(a)



(b)

Figure 4.11: (a) *PDR* and (b) *DEL* as a function of p_s. The main network parameters are set as follows: $\rho = 2000$ nodes/km^2, $n = 200$ nodes, $\lambda = 1$ pkt/s, $s_p = 1.5$ m/s, $n_{tx} = 80$ nodes, and $c = 1$. Various broadcasting protocols are directly compared.

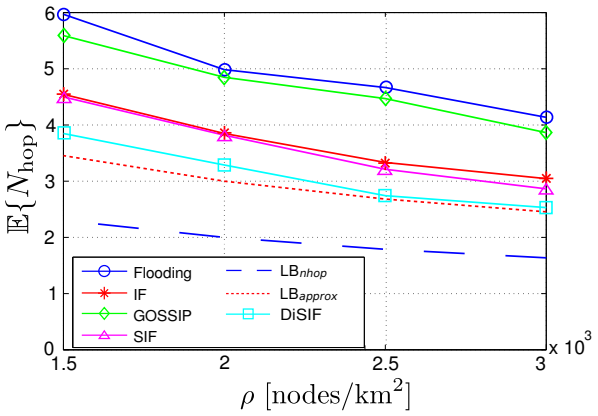


Figure 4.12: $\mathbb{E}\{N_{\text{hop}}\}$ as a function of ρ. The performances of the considered broadcasting protocols are directly compared with the bounds LB_{nhop} (eq.(4.5)) and LB_{approx} (eq.(4.10)).

In Fig. 4.12, the average number of hops is shown as a function of the node spatial density ρ. The simulation-based performances of the considered protocols are directly compared with LB_{nhop} and LB_{approx}. The main network parameters are set as in Fig. 4.10, but for the node spatial density ρ, which is varied by varying the side length ℓ of the square network region and keeping the number n of network nodes fixed to 200. In these settings, increasing values of ρ reduce the average number of hops since nodes get closer to each other. From the results in Fig. 4.12, it can be observed that flooding and GOSSIP require, on average, large numbers of hops. This is due to the fact that the rebroadcaster selection strategy of these protocols does not take into account the internode distance, so that even a node close to the source may rebroadcast a packet, resulting in a low forward progress and increasing the collision probability. Conversely, the IF and SIF rebroadcaster selection strategies favor nodes which are far away from the source and this results in a smaller average number of hops and a lower collision probability. DiSIF guarantees the smallest average number of hops among all simulated protocols for all the considered values of ρ. These results underline, once more, the effectiveness of the new silencing strategy embedded in DiSIF. By comparing the protocols' performances with the proposed bounds, it can

be observed that the performance of DiSIF is very close to LB_{approx}.

4.5.3 Impact of Positioning Error

As described in Section 2.2, the probabilistic forwarding approach represented by IF needs requires the knowledge of some topological network parameters, such as internode distance and node spatial density. This is true also for DiSIF since it exploits IF in order to select the candidate rebroadcasters.

For this reason, each node has been assumed to be equipped with a GPS transceiver. The GPS, deployed in diverse networking settings and increasingly common (e.g., in the majority of smartphones), is often exploited in many existing broadcasting techniques [73, 87]. However, a GPS system can be affected by an error in many ways: propagation errors, signal multipath, receiver clock errors, GPS satellite orbit errors, and others—the interested reader is referred to [88] for a more accurate description. The GPS positioning error, which is rarely taken into account in the communication protocols literature, can severely damage the performance of a topology-based broadcast forwarding technique. For this reason, an investigation on the impact of the positioning error on the performance of DiSIF is now performed.

Denoting as $(x_t; y_t)$ the true coordinates of a node, the GPS positioning error can be modeled as follows:

$$\begin{cases} x = x_t + n_x \\ y = y_t + n_y \end{cases}$$

where n_x and n_y are independent zero-mean Gaussian random variables with standard deviation σ_n (dimension: [m]). In Fig. 4.13, the performance of DiSIF, in the presence of GPS positioning error, is investigated in terms of *PDR* as a function of the packet generation rate λ. Three different values of σ_n are considered: 0 m (corresponding to perfect localization), 31 m, and 100 m. Moreover, in order to have a performance reference benchmark, the *PDR* of GOSSIP is also shown. The main network parameters are set as in Fig. 4.7. It can be observed that the performances with $\sigma = 31$ m and $\sigma = 0$ m are quite similar: this means that DiSIF is robust to moderate localization errors. The reason behind this robustness is the probabilistic forwarding strategy of DiSIF inherited from IF. In the first contention phase, according to the IF strategy,

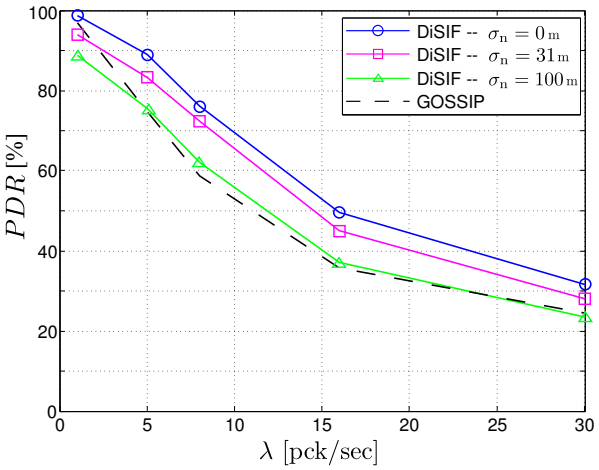


Figure 4.13: DiSIF performance, considering the GPS positioning error, in term of *PDR* as function of λ. Three different values of σ_n are considered: 0 m, 31 m, and 100 m. For comparison purposes the performance of GOSSIP is also shown. The main network parameters are set as in Fig. 4.8.

each node uses the node spatial density and the distance from the source in order to evaluate the presence of other nodes in its proximity and to elect candidate rebroadcasters. However, this is done only in an average statistical sense, so that a moderate position estimation inaccuracy does not affect the performance. Conversely, in the actual rebroadcasters' selection phase, localization errors can lead to silencing nodes which should retransmit the packets and, thus, some inefficiencies may arise. However, the observed results show that these inefficiencies induce a limited performance loss. Focusing on the case with $\sigma_n = 100$ m, it can be noted that the performance of DiSIF converges to that of GOSSIP. This is due to the fact that, in our simulations, the node range z is set to 83 m, so that an error of ± 100 m is equivalent to a random selection of the rebroadcasters. The main implication of the robustness of DiSIF to the localization error is that the GPS system can be replaced with less accurate positioning estimation techniques. For example, internode distance estimation could be based on the RSSI [81, 82]. This aspect is attractive in energy-constrained scenarios

such as WSNs or LPLNs, where the use of GPS may not be a viable option.

4.6 Conclusions

In this chapter, a novel broadcast forwarding strategy, denoted as DiSIF, has been proposed. This technique improves the previously proposed IF and SIF protocols. For almost all the considered values of the network parameters and for all the considered performance metrics, results show that DiSIF outperforms, besides IF and SIF, other existing static probabilistic forwarding protocols such as GOSSIP and flooding. This is mainly due to the new silencing mechanism of the DiSIF protocol, which limits the number of retransmissions and effectively selects the best rebroadcasters, making the channel utilization more efficient. By comparing simulation results with theoretical findings, it turns out that DiSIF is close to be the optimal rebroadcast strategy in terms of minimization of the number of hops per communication route. Finally, the performance in the presence of a GPS positioning error shows that DiSIF is robust to moderate GPS error inaccuracy. This suggests that less accurate positioning estimation techniques could be successfully combined with DiSIF.

Conclusions

This thesis has been focused on the efficient multi-hop data dissemination in Smart Cities and IoT environments. This kind of communication paradigm represents a powerful tool which can be used in many scenarios, ranging from the delivery of safety/alert messages to all vehicles of a VANETs to the propagation of control packets in routing protocol for MANETs. In IoT and Smart Cities, heterogeneous communication technologies together with the high level of node spatial density create an high level of interference. The adoption of multi-hop broadcast techniques like flooding, generates a large level of redundancy which decreases the efficiency of the multi-hop propagation due to collisions. In these cases, make a good use of the available bandwidth by saving retransmissions and selecting the best rebroadcasters assumes a fundamental importance and allows to efficiently cope with the so-called Broadcast Storm Problem. In this thesis, some innovative multi-hop dissemination techniques have been presented, mostly based on probabilistic forwarding.

First, a novel reactive routing protocol, denoted as iAODV has been proposed. This protocol is derived from the AODV protocol by replacing the flooding mechanism used in its route discovery phase with the probabilistic forwarding mechanism denoted as IF. A simulative analysis of the obtained performance of iAODV has been performed in different scenarios such as pedestrian network and VANETs. Different simulation tools, namely ns-3, SUMO and OSM, have been exploited in order to obtain realistic simulations in terms of both node mobility and communication protocols. An analytical evaluation, able to show the amount of saved retransmission brought by the iAODV protocol has been also performed. In all cases, and for

almost all the considered values of the network parameters, the iAODV protocol outperforms the AODV protocol and another version of AODV obtained by replacing flooding with static probabilistic forwarding and denoted as AODV+G protocol. This is mainly due to the fact that the number of control messages is effectively reduced by the use of IF, thus reducing the contention level and making the channel utilization more efficient. Moreover, the thanks to the retransmission probability of IF, the best rebroadcasters are chosen (i.e., the farther from the transmitter) and this allow to efficiently mitigate the broadcast storm problem.

Later, data dissemination with probabilistic forwarding has been exploited in VANETs scenarios in order to create the so-called Vehicular Sensor Networks (VSNs). With VSN, vehicles continuously gather, process, and share location-relevant sensor data (e.g., road conditions, pollution, etc.). Information collection and dissemination is performed using inter-vehicular communications and/or relying on the presence of roadside infrastructure. In this context, two phases have been envisioned: a downlink phase, during which a novel clustering broadcast protocol, denoted as CEIF, is employed to create clusters of vehicles; and an uplink phase, during which the clusters are exploited in order to perform decentralized detection of a spatially constant phenomenon of interest. Different clustered topologies and fusion rules have been considered. The performance of the proposed VSN-based distributed detection schemes has been analyzed in terms of probability of error on the phenomenon estimate and network lifetime, considering different mobility models (namely, highway-like and urban-like). Unlike clustered sensor networks, where the clustering structure is a design aspect, the proposed vehicular distributed detection schemes exploit the natural formation of ephemeral vehicle clusters. Results clearly show that the maximum amount of data collectible during the clustered VSN lifetime is more affected by the node mobility level than by the specific clustering structure. This suggest that "bursty" data collection strategies should be considered, together with proper (local) reclustering strategies. This is especially relevant in urban-like scenarios.

Finally, a novel broadcast forwarding strategy, denoted as DiSIF, has been proposed. This technique improves the previously proposed IF protocol and is more suitable to be used in scenarios where nodes have a bidimensional space distribu-

tion. In particular, DiSIF extends IF by introducing a novel contention phase able to effectively select the best rebroadcasters taking into account the distance from the transmitting node. For almost all the considered values of the network parameters and for all the considered performance metrics, the results show that DiSIF outperforms, besides IF, other existing static probabilistic forwarding protocols such as GOSSIP and flooding. This is mainly due to the new silencing mechanism of the DiSIF protocol, which limits the number of retransmissions and effectively selects the best rebroadcasters, making the channel utilization more efficient. This performance improvement comes at the cost of an overall increase of the end-to-end delay which is introduced by the contention mechanism. By comparing simulation results with theoretical findings, it turns out that DiSIF is close to be the optimal rebroadcast strategy in terms of minimization of the number of hops per communication route. The performance in the presence of a GPS positioning error has been also investigated. Results shows that DiSIF is robust to moderate GPS error inaccuracy and this suggests that less accurate positioning estimation techniques could be successfully combined with DiSIF.

List of Publications

International Journals

- A. Gorrieri and G. Ferrari, "Irresponsible DSR", *submitted to IEEE Communications Letters*, January 2016.
- A. Gorrieri and G. Ferrari, "DiSIF: A distance-based silencing technique for multi-hop broadcast communications in pedestrian ad-hoc networks", *IEEE Tansactions on Mobile Computing*, accepted in December 2015.
- A. Gorrieri, M. Martalò, S. Busanelli, and G. Ferrari, "Clustering and sensing with decentralized detection in vehicular ad hoc networks," *Ad Hoc Networks*, Special Issue on *Vehicular Networking for Mobile Crowd Sensing*, vol. 36, Part 2, January 2016, pp. 450-464.
  DOI:10.1016/j.adhoc.2015.05.019.
- L. Belli, S. Cirani, L. Davoli, A. Gorrieri, M. Mancin, M. Picone, and G. Ferrari, "Design and deployment of an IoT application-oriented testbed", *IEEE Computer*, Special Issue *Activating the Internet of things*, vol. 48, no. 8, September 2015, pp. 32-40. DOI:10.1109/MC.2015.253.
- A. Gorrieri and G. Ferrari "Irresponsible AODV routing", *Vehicular Communications*, vol. 2, no. 1, pp. 47-57, January 2015. Doi: 10.1016/j.vehcom.2015.01.002.

International Conferences

- L. Belli, S. Cirani, A. Gorrieri, M. Picone "A novel smart object-driven UI generation approach for mobile devices in the Internet of Things", *Proc. 1st International Workshop on Experiences with the Design and Implementation of Smart Objects* (SmartObjects'15), Paris, France, September 2015, pp. 1-6. DOI: 10.1145/2797044.2797046

- A. Gorrieri "Efficient multi-hop broadcast data dissemination for IoT and Smart Cities applications", in *Proc. 2015 MobiSys PhD Forum*, Florence, Italy, May 2015, pp. 1-2. DOI: 10.1145/2752746.2752788

- A. Abrardo, M. Barni, A. Gorrieri, G. Ferrari "An information-theoretic analysis of dirty paper coding for informed audio watermarking", *Proc. of Information Theory and Applications Workshop* (ITA 2014), pp. 1-6, UCSD, San Diego, CA, USA, February 2014. DOI: 10.1109/ITA.2014.6804212.

- A. Gorrieri and G. Ferrari "Simulative analysis of saturation condition in a multi source ad-hoc network", *Proc. 9th the 9th International Wireless Communications and Mobile Computing Conference* (IWCMC 2013), Cagliari-Sardinia, Italy, July 2013, pp. 189-193. DOI: 10.1109/IWCMC.2013.6583557.

Bibliography

[1] Y. Gongjun and S.Olariu. A probabilistic analysis of link duration in vehicular ad hoc networks. *IEEE Trans. Intelligent Transportation Systems*, 12(4):1227–1236, Dec. 2011.

[2] C. Bormann Z. Shelby, K. Hartke and B. Frank. RFC 7252: The Constrained Application Protocol (CoAP). 2014.

[3] D Locke. MQ Telemetry Transport (MQTT) v3.1 protocol specification. https://www.ibm.com/developerworks/webservices/library/ws-mqtt/. In *IBM developer Works Technical Library*, 2010.

[4] IEEE 802.15.4b standard, Wireless medium access control and physical layer specification for low rate wireless personal area networks, 2006.

[5] K. Akkaya T. Imboden and Z. Moore. Performance evaluation of wireless mesh networks using ieee 802.11s and IEEE 802.11n. In *IEEE International Conference on Communications (ICC), Ottawa, Canada*, pages 5675–5679, June 2012.

[6] S. Yamaguchi, D. Arai, T. Ogishi, and S. Ano. Short paper: experimental study of long-term operation of BLE tags for realizing indoor location based service. In *18th International Conference on Intelligence in Next Generation Networks (ICIN), Paris, France*, pages 136–138, Feb 2015.

[7] Z Shelby and C. Bormann. *6LoWPAN: The Wireless Embedded Internet*. Wiley Publishing, 2010.

[8] Insitute of Electrical and Electronics Engineers. IEEE Std 802.11TM-2007. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, 2007.

[9] F. Shahin. *ZigBee Wireless Networks and Transceivers*. Newnes, Newton, MA, USA, 2008.

[10] IEEE 802 website. `http://www.ieee802.org/`,.

[11] K. Shuaib, M. Boulmalf, F. Sallabi, and A. Lakas. Co-existence of zigbee and wlan, a performance study. In *Wireless Telecommunications Symposium, 2006. WTS '06*, pages 1–6, 2006. `doi:10.1109/WTS.2006.334532`.

[12] L. Angrisani, M. Bertocco, D. Fortin, and A. Sona. Experimental study of coexistence issues between ieee 802.11b and ieee 802.15.4 wireless networks. *Instrumentation and Measurement, IEEE Transactions on*, 57(8):1514–1523, 2008. `doi:10.1109/TIM.2008.925346`.

[13] M. Petrova, Lili Wu, P. Mahonen, and J. Riihijarvi. Interference measurements on performance degradation between colocated ieee 802.11g/n and ieee 802.15.4 networks. In *Networking, 2007. ICN '07. Sixth International Conference on*, pages 93–93, 2007. `doi:10.1109/ICN.2007.53`.

[14] Telosb - Telosb Mote Platform. Technical Report 0094. URL: `http://www.xbow.com/Products/Product_pdf_files/Wireless_pdf/MICAz_Datasheet.pdf`.

[15] Jangkyu Yun, Byeongjik Lee, Jilong Li, and Kijun Han. A channel switching scheme for avoiding interference of between ieee 802.15.4 and other networks. In *Computer and Computational Sciences, 2008. IMSCCS '08. International Multisymposiums on*, pages 136–139, 2008. `doi:10.1109/IMSCCS.2008.41`.

[16] S. Pollin, M. Ergen, M. Timmers, A. Dejonghe, L. Van der Perre, F. Catthoor, I. Moerman, and A. Bahai. Distributed cognitive coexistence of 802.15.4 with

802.11. In *Cognitive Radio Oriented Wireless Networks and Communications, 2006. 1st International Conference on*, pages 1–5, 2006. doi:10.1109/CROWNCOM.2006.363456.

[17] Jean-Philippe Vasseur and Adam Dunkels. *Interconnecting Smart Objects with IP - The Next Internet*. Morgan Kaufmann, 2010. URL: http://TheNextInternet.org/.

[18] O. Abedi, M. Fathy, and J. Taghiloo. Enhancing AODV routing protocol using mobility parameters in VANET. In *IEEE/ACS International Conference on Computer Systems and Applications (AICCSA)*, pages 229–235, Doha, Qatar, March 2008.

[19] L. Pelusi, A. Passarella, and M. Conti. Opportunistic networking: data forwarding in disconnected mobile ad hoc networks. *IEEE Communications Magazine*, 44(11):134–141, November 2006.

[20] S.R. Das, C.E. Perkins, and E.M. Royer. Performance comparison of two on-demand routing protocols for ad hoc networks. In *Proceedings of the 19th IEEE INFOCOM, Tel Aviv, Israel*, volume 1, pages 3–12, March 2000.

[21] A. Tuteja, R. Gujral, and S. Thalia. Comparative performance analysis of DSDV, AODV and DSR routing protocols in manet using ns-2. In *Proceedings of International Conference on Advances in Computer Engineering (ACE)*, pages 330–333, Bangalore, Karnataka, India, June 2010.

[22] Charles E. Perkins and Pravin Bhagwat. Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers. *ACM SIGCOMM Computer Communication Review*, 24:234–244, oct 1994.

[23] S. Murthy and J. J. Garcia-Luna-Aceves. A routing protocol for packet radio networks. In *Proceedings of the 1st annual international conference on Mobile computing and networking*, pages 86–95, Berkeley, CA, USA, 1995. ACM.

[24] Y. Hu. D. Johnson and D. Maltz. The dynamic source routing protocol (DSR) for mobile ad hoc networks. http://tools.ietf.org/html/rfc4728.

[25] C.E. Perkins and E.M. Royer. Ad-hoc on-demand distance vector routing. In *Second IEEE Workshop on Mobile Computing Systems and Applications WMCSA, New Orleans, LA, USA, February 1999*, pages 90–100.

[26] R. H Jhaveri and M Patel. Mobile Ad-hoc Networking with AODV: A Survey. *International Journal of Next-Generation Computing*, 6(3):165–191, 2015.

[27] S. Lee, W Su, and M. Gerla. On-demand multicast routing protocol in multihop wireless mobile networks. *Mobile Networks and Applications*, 7:441–453, 2002.

[28] Yu-Chee Tseng, Sze-Yao Ni, Yuh-Shyan Chen, and Jang-Ping Sheu. The broadcast storm problem in a mobile ad hoc network. *Wireless Networks*, 8:153–167, 2002.

[29] S. Panichpapiboon and G. Ferrari. Irresponsbile forwarding. In *Proc. IEEE Intl. Conf. on Intelligent Transport System Telecommunication* (ITST), pages 311–316, Phuket, Thailand, Oct. 2008.

[30] Y. Sasson, D. Cavin, and A. Schiper. Probabilistic broadcast for flooding in wireless mobile ad hoc networks. In *Proceedings of Wireless Communications and Networking, WCNC*, volume 2, pages 1124–1130, March 2003.

[31] Zygmunt J. Haas, Joseph Y. Halpern, and Li Li. Gossip-based ad hoc routing. *IEEE/ACM Transactions on Networking, Piscataway, NJ, USA*, 14:479–491, jun 2006.

[32] Amnon Aharony and Dietrich Stauffer. *Introduction to percolation theory*. Taylor and Francis, London, 1994.

[33] Chien-Chung Shen, Zhuochuan Huang, and Chaiporn Jaikaeo. Directional broadcast for mobile ad hoc networks with percolation theory. *IEEE Transactions on Mobile Computing*, 5:317–332, 2006.

[34] Jeffrey E. Wieselthier, Gam D. Nguyen, and Anthony Ephremides. Energy-efficient broadcast and multicast trees in wireless networks. *Mobile Networks and Applications*, 7:481–492, dec 2002.

[35] Birsen Sirkeci Mergen, Student Member, Anna Scaglione, and Gökhan Mergen. Asymptotic analysis of multistage cooperative broadcast in wireless networks. *IEEE Transactions on Information Theory*, 52:2531–2550, 2006.

[36] S. Busanelli, G. Ferrari, and S. Panichpapiboon. Efficient broadcasting in IEEE 802.11 networks through irresponsible forwarding. In *Proc. IEEE Global Telecommun. Conf.* (GLOBECOM), pages 1–6, Honolulu, HI, USA, Dec. 2009.

[37] A. Udaya Shankar, Cengiz Alaettinoglu, Klaudia Dussa-Zieger, and Ibrahim Matta. Transient and steady-state performance of routing protocols: Distance-vector versus link-state. *Internetworking: research and experience*, 6, 1996.

[38] Athanasios Papoulis. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill Companies, 3rd edition, Feb 1991.

[39] Anbao Wang and Wenrong Jiang. Research of teaching on network course based on ns-3. In *Int. Workshop on Education Technology and Computer Science*, volume 2, pages 629–632, Wuhan, China, Mar. 2009.

[40] F.K. Karnadi, Zhi Hai Mo, and Kun chan Lan. Rapid generation of realistic mobility models for VANET. In *Proc. IEEE Wireless Commmunications and Networking Conference* (WCNC), pages 2506–2511, Hong Kong, China, Mar. 2007.

[41] Chitraxi Raj, Urvik Upadhayaya, Twinkle Makwana and Payal Mahida. Simulation of VANET using ns-3 and SUMO. *International Journal of Advanced Research in Computer Science and Software Engineering*, 4:563–569, April 2014.

[42] C. Sommer and F. Dressler. Progressing toward realistic mobility models in VANET simulations. *IEEE Communications Magazine*, 46(11):132–137, November 2008.

[43] M. Haklay and P. Weber. OpenStreetMap: User-generated street maps. *IEEE Pervasive Computing*, 7(4):12–18, Oct. 2008.

[44] J.W. Robinson and T.S. Randhawa. Saturation throughput analysis of IEEE 802.11e enhanced distributed coordination function. *IEEE Journal on Selected Areas in Communications*, 22:917–928, June 2004.

[45] Uichin Lee and Mario Gerla. A survey of urban vehicular sensing platforms. *Elsevier Computer Networks*, 54(4):527–544, Mar. 2010. doi:DOI:10.1016/j.comnet.2009.07.011.

[46] Uichin Lee, E. Magistretti, M. Gerla, P. Bellavista, and A. Corradi. Dissemination and harvesting of urban data using vehicular sensing platforms. *IEEE Trans. Veh. Technol.*, 58(2):882–901, Feb. 2009.

[47] Bret Hull, Vladimir Bychkovsky, Yang Zhang, Kevin Chen, Michel Goraczko, Allen Miu, Eugene Shih, Hari Balakrishnan, and Samuel Madden. CarTel: a distributed mobile sensor computing system. In *ACM Int. Conf. on Embedded Networked Sensor Systems (*SenSys*)*, pages 125–138, Boulder, Co, USA, Nov. 2006.

[48] C. R. Lin and M. Gerla. Adaptive clustering for mobile wireless networks. *IEEE Journal on Selected Areas in Communications*, 15(7):1265–1275, Sep. 1997.

[49] L. Bononi and M. Di Felice. A cross layered MAC and clustering scheme for efficient broadcast in VANETs. In *IEEE Conf. Mobile Ad-hoc and Sensor Systems (*MASS), pages 1–8, Montreal, Canada, Oct. 2007.

[50] M. Fiore and J. Härri. The networking shape of vehicular mobility. In *Proc. of the ACM Intl. Symposium on Mobile Ad-hoc Networking and Computing* (MobiHoc), pages 261–272, Hong Kong, China, May 2008.

[51] M. Martalò, C. Buratti, G. Ferrari, and R. Verdone. Decentralized detection in IEEE 802.15.4 wireless sensor networks. *EURASIP Journal on Wireless Communications and Networking*, 2010, 10 pages, 2010.

[52] G. Ferrari, M. Martalò, and R. Pagliari. Decentralized detection in clustered sensor networks. *IEEE Trans. Aerosp. Electron. Syst.*, 47(2):959–973, Apr. 2011.

[53] Arne Kesting, Martin Treiber, Martin Schönhof, and Dirk Helbing. Adaptive cruise control design for active congestion avoidance. *Elsevier Transportation Research Part C: Emerging Technologies*, 16(6):668–683, 2008.

[54] Arne Kesting, Martin Schönhof, Stefan Lämmer, Martin Treiber, and Dirk Helbing. *Managing Complexity: Insights, Concepts, Applications*, chapter "Decentralized Approaches to Adaptive Traffic Control", pages 189–199. Springer-Verlag, Heidelberg, Germany, 2008. Eds.: D. Helbing.

[55] S. Busanelli, G. Ferrari, and S. Panichpapiboon. Cluster-based irresponsible forwarding. In *Tyrrhenian International Workshop on Digital Communications*, pages 59–68, Pula, Sardinia, Italy, Sep. 2009.

[56] S. Busanelli, M. Martalò, and G. Ferrari. Clustered vehicular networks: Decentralized detection "on the move". In *Proc. Int. Conf. ITS Telecommunications (ITST)*, pages 744–749, St. Petersburg, Russia, Aug. 2011.

[57] N. Wisitpongphan, F. Bai, P. Mudalige, V. Sadekar, and O. K. Tonguz. Routing in sparse vehicular ad hoc wireless networks. *IEEE J. Select. Areas Commun.*, 25(8):1538–1556, Oct. 2007.

[58] K. Fall and K. Varadhan. *The ns Manual*. The VINT Project, May 2010.

[59] G. Ferrari and O. K. Tonguz. Impact of mobility on the ber performance of ad hoc wireless networks. *IEEE Transaction on Vehicular Technology*, 56:271–286, Jan 2007.

[60] S. Busanelli, F. Rebecchi, M. Picone, N. Iotti, and G. Ferrari. Cross-network information dissemination in vehicular ad hoc networks (VANETs): Experimental results from a smartphone-based testbed. *MDPI Future Internet*, 5(3):398–428, 2013.

[61] Eureka project 6252 X-NETAD. http://www.eurekanetwork.org/project/-/id/6252.

[62] F. A. Teixeira, V. F. Silva, J. L. Leoni, D. F. Macedo, and J. M. S. Nogueira. Vehicular networks using the IEEE 802.11p standard: An experimental analysis. *Elsevier Vehicular Communications*, 1(2):91–96, Apr. 2014.

[63] B. E. Bilgin and V. C. Gungor. Performance Comparison of IEEE 802.11p and IEEE 802.11b for Vehicle-to-Vehicle Communications in Highway, Rural, and Urban Areas. *International Journal of Vehicular Technology*, 2013:10 pages, Oct. 2013.

[64] Stefan Krauß. *Microscopic Modeling of Traffic Flow: Investigation of Collision Free Vehicle Dynamics*. PhD thesis, 1998.

[65] Daniel Krajzewicz. Traffic simulation with SUMO – simulation of urban mobility. In Jaume Barceló, editor, *Fundamentals of Traffic Simulation*, volume 145 of *International Series in Operations Research & Management Science*, pages 269–293. Springer New York, 2010.

[66] S. Asmussen and L. Rojas-Nandayapa. Asymptotics of Sums of Lognormal Random Variables with Gaussian Copula. *Statistics and Probability Letters*, 78(16):2709–2714, 2008.

[67] N.A. Marlow. A normal limit theorem for power sums of independent normal random variables. *Bell System Technical Journal*, 46(9):2081–2089, 1967.

[68] S. Busanelli, G. Ferrari, and R. Gruppini. Recursive analytical performance evaluation of broadcast protocols with silencing: application to VANETs. *EURASIP Journal on Wireless Communication and Networking,*, 2012.

[69] A. Gorrieri and G. Ferrari. Simulative analysis of saturation condition in a pedestrian ad-hoc network. In *9th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 189–193, Cagliari, Italy, July 2013.

[70] Joanna Kulik, Wendi Heinzelman, and Hari Balakrishnan. Negotiation-based protocols for disseminating information in wireless sensor networks. *Wireless Networks*, 8(2/3):169–185, mar 2002.

[71] M. Zorzi and R.R. Rao. Geographic random forwarding (GeRaF) for ad hoc and sensor networks: Multihop performance. *IEEE Transactions on Mobile Computing*, 2(4):337–348, Oct 2003.

[72] M. Zorzi and R.R. Rao. Geographic random forwarding (GeRaF) for ad hoc and sensor networks: energy and latency performance. *IEEE Transactions on Mobile Computing*, 2(4):349–365, Oct 2003. doi:10.1109/TMC.2003.1255650.

[73] Arjan Durresi, Vamsi K. Paruchuri, S. Sitharama Iyengar, and Rajgopal Kannan. Optimized broadcast protocol for sensor networks. *IEEE Transactions on Computers*, 54(8):1013–1024, 2005.

[74] Xing Fan, Bo Yang, R. Yamamoto, and Y. Tanaka. Road side unit assisted stochastic multi-hop broadcast scheme for instant emergency message propagation. In *17th International Conference on Advanced Communication Technology (ICACT), Gangwon-do, South Korea*, pages 450–457, July 2015. doi:10.1109/ICACT.2015.7224909.

[75] Hongseok Yoo and Dongkyun Kim. ROFF: RObust and Fast Forwarding in vehicular ad-hoc networks. *IEEE Transactions on Mobile Computing*, 14(7):1490–1502, Jul 2015. doi:10.1109/TMC.2014.2359664.

[76] C. Wu, X. Chen, Y. Ji, S. Ohzahata, and T. Kato. Efficient Broadcasting in VANETs Using Dynamic Backbone and Network Coding. *IEEE Transactions on Wireless Communications*, 14(11):6057–6071, Nov 2015. doi:10.1109/TWC.2015.2447812.

[77] Yi Song, Jiang Xie, and Xudong Wang. A novel unified analytical model for broadcast protocols in multi-hop cognitive radio ad hoc networks. *IEEE Transactions on Mobile Computing*, 13(8):1653–1667, Aug 2014.

[78] Xiaohuan Li, Bin-Jie Hu, Hongbin Chen, Bing Li, Huanglong Teng, and Manman Cui. Multi-hop delay reduction for safety-related message broadcasting in vehicle-to-vehicle communications. *Communications, IET*, 9(3):404–411, Dec 2015. doi:10.1049/iet-com.2014.0303.

[79] C. Campolo, A. Molinaro, A. Vinel, and Yan Zhang. Modeling prioritized broadcasting in multichannel vehicular networks. *IEEE Transactions on Vehicular Technology*, 61(2):687–701, Feb 2012. doi:10.1109/TVT.2011.2181440.

[80] *Amendment 6: Wireless Access in Vehicular Environments*, IEEE Std. 802.11p, 2010.

[81] A. Awad, T. Frunzke, and F. Dressler. Adaptive distance estimation and localization in WSN using RSSI measures. In *Digital System Design Architectures, Methods and Tools, 2007. DSD 2007. 10th Euromicro Conference on*, pages 471–478, 2007.

[82] K. Benkic, M. Malajner, P. Planinsic, and Z. Cucej. Using RSSI value for distance estimation in wireless sensor networks based on Zigbee. In *Systems, Signals and Image Processing, 2008. IWSSIP 2008. 15th International Conference on*, pages 303–306, 2008. doi:10.1109/IWSSIP.2008.4604427.

[83] A Kostin. Probability distribution of distance between pairs of nearest stations in wireless network. *Electronics Letters*, 46(18):1299–1300, September 2010.

[84] JeongGil Ko, A. Terzis, S. Dawson-Haggerty, D.E. Culler, J.W. Hui, and P. Levis. Connecting low-power and lossy networks to the internet. *IEEE Communications Magazine*, 49(4):96–101, April 2011.

[85] Biswanath Mukherjee Jennifer Yick and Dipak Ghosal. Wireless sensor network survey. *Computer Networks*, 52(12):2292 – 2330, 2008.

[86] O. K. Tonguz and G. Ferrari. *Ad Hoc Wireless Networks: A Communication-Theoretic Perspective*. John Wiley & Sons, 2006.

[87] D. Camara and A A F Loureiro. A GPS/ant-like routing algorithm for ad hoc networks. In *IEEE Wireless Communications and Networking Conference,* WCNC 2000, volume 3, pages 1232–1236, Sep 2000, Chicago Il USA.

[88] E. Kaplan. *Understanding GPS - Principles and applications*. Artech House, 2nd edition edition, December 2005.

Ringraziamenti

Ci sono davvero molte persone che devo ringraziare e senza le quali sarebbe stato molto difficile arrivare a questo traguardo. Parto ringraziando il mio relatore Gianluigi Ferrari che mi ha dato la possibilità di intraprendere questo percorso, guidandomi e incoraggiandomi nei momenti di difficoltà. Un grande ringraziamento va a Marco Picone e Simone Cirani che hanno rappresentato una guida fondamentale per il mio dottorato insegnandomi davvero molto. Un grazie particolare a Laura Belli, Luca Davoli e Federico Parisi con i quali ho condiviso la maggior parte del dottorato, grazie soprattutto per la pazienza, la competenza e la disponibilità. Grazie anche a tutto il WASNLab, Marco Martalò, Matteo Giuberti, Giovanni Spigoni, Pietro Gonizzi, Stefania Monica, Muhammad Asim, Mirko Mancin e Nicolò Strozzi, per il fondamentale aiuto tecnico, per tutto ciò che mi hanno insegnato, per i consigli e le risate. Un grazie speciale al MultimediaLab al completo, Carlo Tripodi, Davide Alinovi e Luca Cattani, per i fondamentali consigli in materia di elaborazione del segnale e probabilità e per il fatto che la pausa pranzo sarebbe stata notevolmente più noiosa senza di loro. Un grazie anche a Gabriele Ferrari e Mattia Antonini, sempre in grado di scherzare e fornire al tempo stesso preziosi consigli. Ultimi ma non per importanza, ringrazio Agnese e la mia famiglia, sempre pronti ad incoraggiarmi e sostenermi nel momento del bisogno.