Università degli Studi di Torino
Scuola di Dottorato in Scienza e Alta Tecnologia

Tesi di Dottorato di Ricerca in Scienza e Alta Teconologia
Indirizzo: Informatica

Opportunistic Networking in a Urban Scenario

Giuseppe Corrente

Advisors:

Prof. Matteo Sereno, Prof. Rossano Gaeta, Prof. Marco Grangetto

XXV Ciclo, July, 2013

Università degli Studi di Torino
Scuola di Dottorato in Scienza e Alta Tecnologia

Opportunistic Networking in a Urban Scenario

Giuseppe Corrente

Acknowledgments

This work has been partially supported by the M.A.S.P. project, a research project supported by the Piedmont Region, Italy, (framework program POR-FESR Asse I, I.1.3) and was published and presented in PE-WASUN 2011 (mainly showed in Chapter 4) in WONS 2012 (mainly showed in Chapter 6), and submitted to Elsevier *Special Issue on Opportunistic Networks* (mainly showed in Chapter 7).

I also want to thank my tutor Prof. Matteo Sereno, Prof. Marco Grangetto and Prof. Rossano Gaeta to help me in this research in the last three years. I also need to thank the whole Computer Science Department of the University of Turin that helped me deepen my passion for Computer Science and Science in general.

Finally I would like to thank my wife for having always been very supportive.

No problem can be solved from the same level of consciousness that created it.

(Albert Einstein)

Contents

List of Figures

List of Tables

Abstract

This thesis represents a research study about an example of opportunistic networking used as leverage to diffuse environmental or pervasive information from surroundings or target urban areas towards individual nodes.

This research results have been published in PE-WASUN 2011, WONS 2012, and submitted to Elsevier *Special Issue on Opportunistic Networks*.

Opportunistic networking based on hand-held mobile devices is turning into a viable and efficient opportunity to locate, collect, route and share information within a swarm of collaborative nodes. In this thesis we consider mobile (pedestrian and cars) and fixed terminals in a urban area that are interested in collecting the information originated from several sources. In particular, each terminal aims at retrieving the data items in a limited *region of interest* centered around the node position. Since data items may change over time, all nodes must strive for having access to the latest version. Furthermore, for mobile terminals the region of interest is a time varying concept due to the dynamic behaviour of nodes. The goal of the first step of this thesis is to evaluate the amount of information each node is able to gather resorting to simple distributed data collection and sharing through opportunistic communications among neighboring nodes. This goal is measured by means of a well defined performance index which is referred to as *coverage* (this will be explained in further details in Chapter 3). For instance, we analyze the impact of node density, different mix of cars and pedestrian, and amount of node memory. In addiction, we evaluate the improvement of using location aware memory management policies as well as the effect of adding a few ideal nodes whose mobility is described by an unconstrained Brownian motion. The preliminary findings highlight that simple location aware memory management schemes effectively exploit nodes with limited amount of memory. Furthermore, increasing randomness of nodes movement has a beneficial impact on the average performance of all node types.

Then we test the validity of the results found despite the variation of some simulation features. Our results are also confirmed by applying a more realistic communication scheme as local broadcast communication. These are also confirmed by varying added random mobile node features, pedestrian device communication rates and environmental variability types.

We also compare the performance of different types of nodes, and thus we need to exclude not accessible areas in performance index calculation because these areas could penalize different type of nodes in a different way, each one with its own mobility features and its own region of interest extension.

Then we explain the different performance obtained by pedestrians and vehicles. We also discover that a large fraction of transmissions from terminals operated by pedestrians are redundant, so dramatic savings in energy consumptions can be obtained by slowing down transmission rates at an acceptable cost of reducing the coverage of all node types. This result is also true under environmental variability conditions. We show that performance index can be further improved adding contextual awareness to pedestrian devices transmission policies without great overhead.

Moreover, we study how variable information propagates inside and outside buildings with and without the aid of some infrastructural additional elements, such as powered wireless relay nodes in elevators; indeed the addition of such infrastructural features to the scenario strongly increases the system performance in all the conditions.

Finally, we release the constraint by which the regions of interest must always be near their associated node position and study conditions in which the average target zone knowledge is satisfactory enough. The result is that if a mobile node wants to gather information of its target zone strongly removed from its surroundings it has to share its region of interest with many other nodes.

For implementing these studies we developed a simulator based on mobility and radio propagation traces obtained from the UDelModels tools.

1

Introduction

1.1 Internet of Things

Some decades ago Mark Weiser, in [51] and [52], described the beginning of a new coming era, the Ubiquitous Computing also called Calm Technology. He thinks that past and present computing history can be summarized as mainframe computing, personal computing, and network computing/Internet. In his vision the recent Internet era is a transient point from mainframes and networked personal computers toward a next era in which technology will be pervasive, oriented to supporting people in an almost invisible manner. Indeed in this vision the pervasive computing environment is directly connected to the mind of the user. With his words: ”Specialized elements of hardware and software, connected by wires, radio waves and infra-red, will be so ubiquitous that no one will notice their presence”. A similar dream was described by Arthur C. Clarke, famous futurist and science fiction author, thinking generically about a far future technological era: ”Any sufficiently advanced technology is indistinguishable from magic”.

To evaluate how far we are from similar dreams, I shall initially report some market data:

- At the end of 2011, there were 6 billion mobile subscriptions, estimates

The International Telecommunication Union (2011). That is equivalent to 87 percent of the world population. And is a huge increase from 5.4 billion in 2010 and 4.7 billion mobile subscriptions in 2009.

- In 2011 there were 9 billion connected devices and by 2020 that number will be 24 billion devices, according to GSMA statistics. The number of mobile connected devices is estimated to grow strongly from 6 billion in 2011 to 12 billion by 2020.

- The number of vehicles in operation worldwide surpassed the 1 billion-unit mark in 2010 for the first time ever. According to Wards research, which looked at government-reported registrations and historical vehicle-population trends, global registrations jumped from 980 million units in 2009 to 1.015 billion in 2010. Although each car doesn't have on-board computer, probably the most part of them has more than 40 embedded processors. These are less powerful than the PC processors, but enough to perform the work that must accomplish to safeguard user security, improve the efficiency of the vehicle and reduce the pollution that generates.

- The latest RFID market research from IDTechEx finds that in 2012 the value of the entire RFID market will be $7.67 billion, up from $6.51 billion in 2011. This includes tags, readers and software/services for RFID cards, labels, fobs and all other form factors. It includes passive and active RFID. In retail, RFID is seeing rapid growth for apparel tagging - this application alone demands 1 billion RFID labels in 2012, with 1.35 billion tags forecast for 2013. RFID in the form of tickets used for transit will demand 500 million tags in 2012. In total, 3.98 billion tags will be sold in 2012 versus 2.93 billion in 2011.

- Most of recent cell and smart phones, portable and hand-held devices, have wireless communication capabilities also of device-to-device type and a great part of them is also equipped with camera as well as memory and computational capabilities.

- Many recent urban districts in great cities are equipped with numerous public and private surveillance cameras.
- On 10 July 2012, the European Commission launched the *Smart Cities* and Communities European Innovation Partnership. The partnership proposes to pool resources to support the demonstration of energy, transport and information and communication technologies (ICT) in urban areas. The energy, transport and ICT industries are invited to collaborate with cities in a synergistic way. This will enable innovative, integrated and efficient technologies to enter the market more easily, while making cities more and more smart. The *Smart Cities* European funding will be about 365 million of Euro for 2013.
- According to a new report from Marketsandmarkets, a marketing research company, the global smart cities market is expected to top $1 trillion by 2016. Currently, the report estimates the value of the smart cities market at $526.3 billion, with a compound annual growth rate of about 14.2 percent from 2011-2016. According to the report: globally, there are some 700 cities, each with population exceeding 500,000 and are growing faster than the average growth rate of cities. This opens up the market for industry players to grow their business in new and emerging smart cities. The infrastructure investment for these cities is forecasted to be $30 trillion to $40 trillion, cumulatively, over the next 20 years. These projections follow along with another report estimating growth in smart cities. Last year, Pike Research said that smart city investment would grow to more than $100 billion by 2020.
- Consumer electronics suppliers, as well as service providers and telecommunication companies, are the catalysts of *Domotics* and home automation market. Worldwide market will reach about 115Bn$ in 2016

from about 45Bn$ in 2010 with an average growth rate of 16% per year. Security, energy management, healthcare and wellness, information, entertainment and remote working are among the fastest growing services in the market (cmtresearch).

- Global healthcare cloud computing market will reach $5.41 billion by 2017 (MarketsandMarkets).

From these facts we note an heterogeneity of hardware, software and communication technologies that are spreading pervasively in the physical world bringing closer the dream of ubiquitous computing. However, the way in which smart objects communicate and cooperate with each other and support people needs is stilll far from that outlined in Weisener's vision. One of the key point in order to realize such a dream is the evolution of the Internet into the Internet of Things as well as the transition from the Web to Web 3.0. The key point for the realization of this evolution is to leverage the collective intelligence and various other forms of communication.

As well as from a certain point of view the Web should mainly evolve towards an understanding of the contents to realize the so called Semantic Web, is not lesser importance the focus on the Web fusion with the future Internet of Things ([2], Figure 1.1)- as the figure shows this latter should be realized in three paradigms: internet-oriented (middleware), things oriented (sensors) and semantic-oriented (knowledge).

We will use contents starting from many devices (absolutely not only servers) and also through a multitude of personal devices. Web 3.0 will link different types of networks together, connecting as well as person-to-person, machine-to-person and vice-versa (person-to-machine) and also machine-to-machine. Different contents from different media will be consumed through multiple devices. If conventional Internet protocols has been a so successful and there is a trend to uniform all future evolutions towards them, there is a parallel trend to try different schemes of communication and eventually to integrate or connect these in the next Future Internet.

The IoT (Internet of Things) can be described as an integrated digital environment in which connected devices will be uncountable: some tens of billions of devices, most of which very small, worn by persons and/or diffused into buildings, streets, equipped vehicles, etc.. These small devices will be of various types and capabilities and will increasingly used in communication with each other. The Future Internet design will probably exploit a multitude of communication networks among intelligent devices connected with each other and with sentient things in the environment, as well as a current Internet-like protocol. A multitude of networking technologies could be used to gather information as RFID, WSN, ZigBee, and others, while the Internet infrastructure will not lose importance. On the contrary, it will retain its main role as global backbone for worldwide data sharing and diffusion. When very different, very far or anyway different subnets nodes will need to communicate each other, gateways will be provided to connect different types of network together to enable inter-network communications. So a traditional Internet like network could be in communication with a multitude of less conventional networks used to gather and diffuse local environmental information. A main role will be played by wireless communication technology to exchange ubiquitous data among smart objects, RFIDs and sensors. ”When objects can both sense the environment and communicate, they become tools for understanding complexity and responding to it swiftly. What is revolutionary in all this is that these physical information systems are now beginning to be deployed, and some of them even work largely without human intervention” (Copyright 2010 McKinsey & Company, The Internet of Things).

A similar multitude of sensors, tags and smart objects gathering and diffusing data produce an enormous amount of data. Such a great number of pervasive environmental and personal smart small objects raise new and heavy questions about the guarantee of the personal privacy. Surely important open questions are the management of huge masses of data from

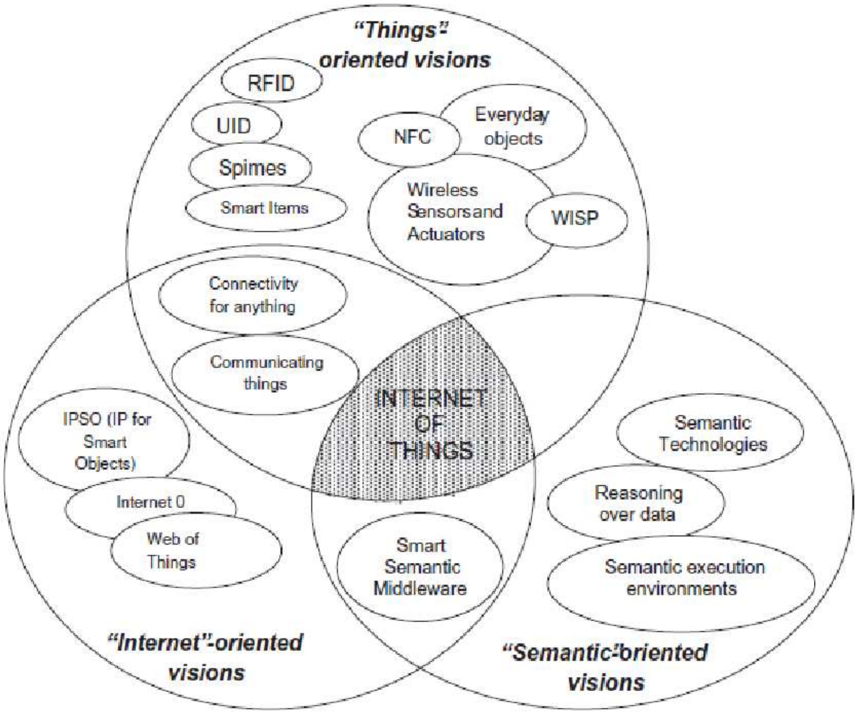


Figure 1.1: A view of Internet of Thing from [2].

pervasive smart objects and tags, as well as software and hardware security issues. However, these points, despite the great relevance, are beyond the focus of our research, which mainly concentrates on the pervasive forms of communication between objects. Surely IoT will have a great impact on business and everyday life of worldwide people.

1.2 Opportunistic Networks and Cyber Physical Convergence

At present, the social and cyber world are already partially connected with the physical world. The famous modern communication media expert H. Rheingold describes this trend: "When they connect the tangible objects and places of our daily lives with the Internet, hand-held communication media mutate into wearable remote-control devices for the physical world".

Besides classical Internet like protocols which base themselves on the assumption of infrequent topology changes, in recent years other network types have been studied. These can cope better with the increasing degrees of node mobility and other causes of intermittent connectivity.

Some networks called *MANETs* (Mobile Ad-hoc Networks) have been studied to cope topology changes, disconnections or disruption because of the mobility of router nodes or other causes. In this case the end-to-end connectivity assumption is maintained, but not stable in time. So routing tables in nodes have to be refreshed often, either periodically or at each topology change in the case of proactive routing, and on demand at each sender node request, in the case of reactive routing. A famous example of this latter is AODV (Ad-Hoc On-Demand Distance Vector). The reactive protocols have less overhead but higher delay than proactive routing.

When the predictability of disruption becomes difficult and its frequency high, some networks called *DTNs* (Delay or Disruption-tolerant networks) that leverage on mobility have been designed by using a store and forward transmission paradigm tolerating some delay. No end-to-end connectivity is supposed at each instant of time. A protocol layer over tcp/ip (the bundle layer) has been designed to manage and implement the store-and-forward. Each node can have the roles: host (source or destination), router (routes bundles or messages inside the same dtn region), gateway (routes the messages between different dtn regions).

When the occasional contacts with other mobile nodes in temporary proximity are the only possibility to exchange and diffuse data, the protocols studied are called opportunistic. The *Opportunistic Networks* (OppNets) can be used in lack of traditional connectivity. The protocols studied are of Store-Carry-and-Forward type, where the term carry underlines node mobility other than temporary contacts used as leverage. OppNets can be supported by an infrastructure (fixed or mobile), but this study focuses more on OppNets protocols without infrastructure that can be classified as those based on dissemination, such as some types of controlled flooding, and those based on context, that utilize context awareness to decide the next hop. On the other hand, a OppNets protocols classification mainly based on context awareness would result as follows:

- context oblivious: are dissemination-based routing, no information on behaviour user is utilized;
- partially context-aware: are mobility based, utilize user context information;
- fully context-aware: are social-context aware, they acquire information on the context and utilize it for message delivery.

This thesis mainly deals with the so-called partially context-aware OppNets protocols, because it specifically focuses on the type of awareness that derives from local sensing and communication between nodes.

The OppNets can be used as network layer to provide environmental knowledge to upper layer by gathering and diffusing information from a myriad of small smart objects diffused in the environment. But the information could also go from upper layers towards OppNets as described in ([9], Figure 1.2).

Equipped human and vehicle networks can be considered as electronic social networks and can be a base for opportunistic communication protocols. However, upper virtual social networks could be considered as a starting point for the development of opportunistic computing services. The opportunistic computing paradigm can be viewed as an evolution of distributed computing in the direction of mobile and pervasive computing paradigms. Opportunistic communication protocols could take context awareness into consideration, where for context awareness we do not only mean information

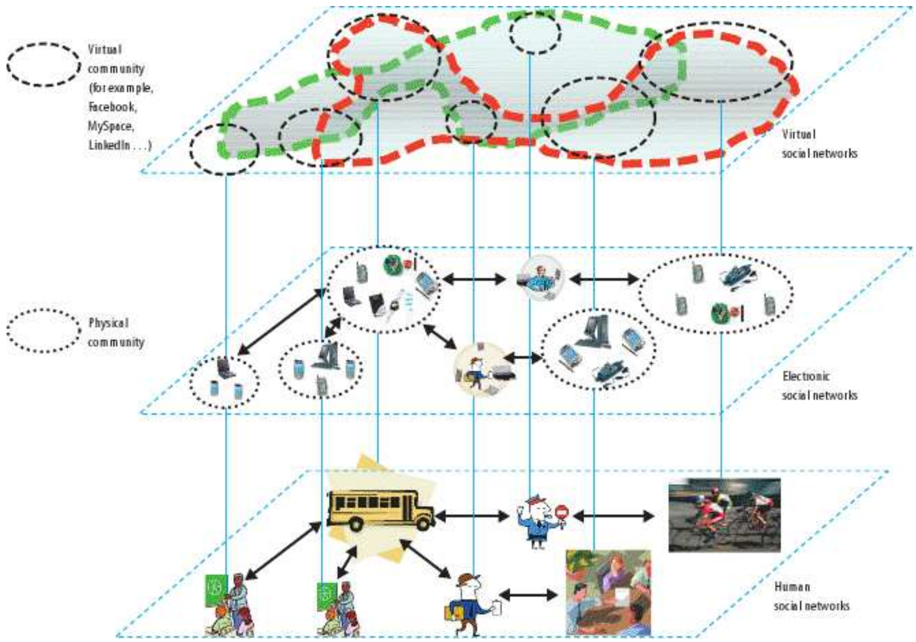


Figure 1.2: Opportunistic Networks figure from [9].

from the surrounding environment, but also virtual context data and relationships such as those acquired from upper social network layer. Also trust and friendship relationships could be imported from upper virtual layer and used in middle-ware and communication layer as a base for security and/or privacy rules.

For pervasive computing we do not only refer to hand-held, portable devices or sensor-rich smart-phones but increasingly more to a wide variety of smart objects, such as RFID tags, sensors, proximity sensing technologies and actuators, the latter are particularly enabled to influence the physical world. In the near future, beside designing good opportunistic networking protocols gathering data from the environment for higher virtual layer support, it could be important to design protocols to change and adapt the behaviour of a myriad of small pervasively diffused devices in a intelligent and coordinated way.

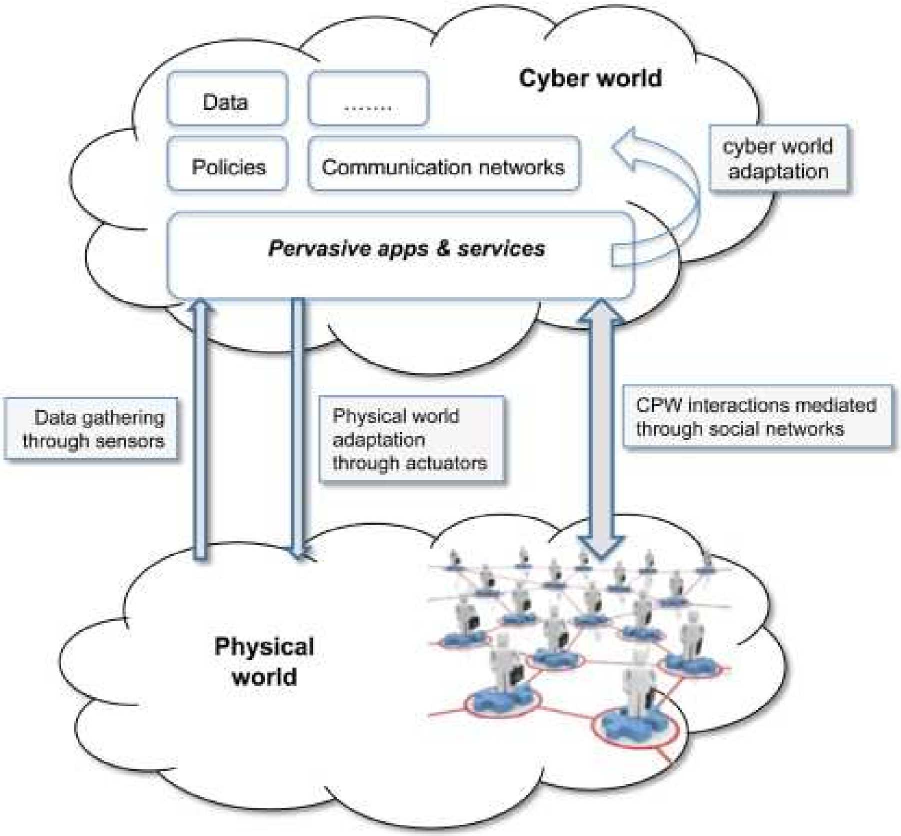


Figure 1.3: Cyber Physical convergence figure from [10].

If we identify human centred electronic networks as the physical world and upper virtual social networks as the cyber world, we could strongly assert that the convergence between these two worlds ([10], Figure 1.3) will be one of the main worldwide technology drivers of the future.

1.3 Driving Ideas

In this section we shall outline the topics specifically studied in our research.

A central point of interest is that each smart object should be conscious of its context. One way to reach this aim is by using sensing capabilities or, in a more effective way, by exchanging opportunistically data with other encountered smart devices or sensor nodes. There are two aims in gathering contextual information: to provide information to software upper levels for supporting user applications, and to adapt own behaviour to external conditions.

Another main point is that we do not view IoT (*Internet of Things*) as an isolated thing but integrated with Web 3.0 and as a central part of the *Smart City* design ([20] and [47]).

The driving ideas in our research may be summarized as follows:

- urban area, included inside areas, viewed as a potential source of informations, obtained by sensing environmental data as well as by wireless communications with a swarm of small tags and/or smart objects pervasively diffused in it
- leverage on urban mobility (pedestrians and vehicles) to help device-to-device networking
- increasing knowledge of surrounding or target zones of each smart node
- local/environmental awareness memory policies and transmission protocols
- different constraints of persons and vehicle equipments to represent heterogeneity of nodes
- communication protocols compatible with hand-held devices limited resources and energy, considering that mounted car equipment is potentially less constrained
- using simulations and statistics as main methodology
- supporting our research with simulations based on a realistic mobile and radio propagation model, instead of using a simplified one

- contributing to the Internet of Things and Smart City advances and realizations

In a few words we focus mainly on the *IoT* and the *opportunistic networking* aspects in an urban scenario.

For this reason our research will focus on a solution delivered by environmental local public data networking, scalable, with bottom-up architectural design without using any infrastructure and based on opportunistic communication. The only implementation of these features is enough to allow us to illustrate the development of systems with interesting and socially useful targets, showing that heterogeneity, mobility, context-awareness and cooperation can be leveraged for performance increasing and energy saving.

Useful details on the types of networks considered in our work are given in the following list:

- *MANET(Mobile Ad-hoc NETwork) and VANET(Vehicular Ad-hoc NETwork)* : ad hoc mobile networks with router mobility as a main feature. Hosts are also associated with related routers in a wireless and flexible way. Some topics of these types of networks are the auto-organization and multi-hop communication capabilities, routing protocols that afford mobility and dynamic configuration problems, eventual integration of Internet connectivity when it is possible.

- *PAN(Personal Area Network) and BAN(Body Area Network)* : networks strictly connected with their users. PANs can be used for communication among the personal devices themselves, or for connecting to a higher level network; while BAN field is an interdisciplinary area which could allow inexpensive and continuous health monitoring with real-time updates of medical records through the Internet.

- *DTN(delay and disruption tolerant network)* : networks outside the Internet context that work in strong mobile and difficult environments, such as in outside space, as well as certain areas in developing countries, or in all those contexts in which there is no guarantee of connectivity. This type of network is mainly based on a store and forward approach because often it is not possible to trace a predetermined routing path, not even using ad hoc algorithms for dynamic contexts.

- *Opportunistic Network* : in this type of network no assumption is made about source-destination connectivity. The only chance of exchanging data among mobile nodes is during short time contacts when they come in strict proximity. With the widespread of portable and handheld devices use with wireless capabilities, the opportunistic networking paradigm has become an interesting possibility. Indeed in this case vehicular and human mobility becomes a possible leverage for networking and spreading of information.

- *WSN (Wireless Sensor Network), RFID(Radio Frequency IDentification)* : in our study context these types of device and relative networks represent potential terminal nodes by which we detect, gather and monitor information and data, such as environmental measurements. The final aim is to have a detailed view of a whole scenario, or some regions of it, from different points of view also in contexts without connectivity.

- *VSN (Vehicular Sensor Network)*: a specific VANET in which vehicles are equipped with sensors or, for example, cameras without the strict resource limitation typical of WSN. This type of network has also a greater degree of mobility predictability because it is constrained from features of urban traffic. This type of network has been explored in [33] and [32] (see Figure 1.4) as a base for some communication protocol studies and middleware development.

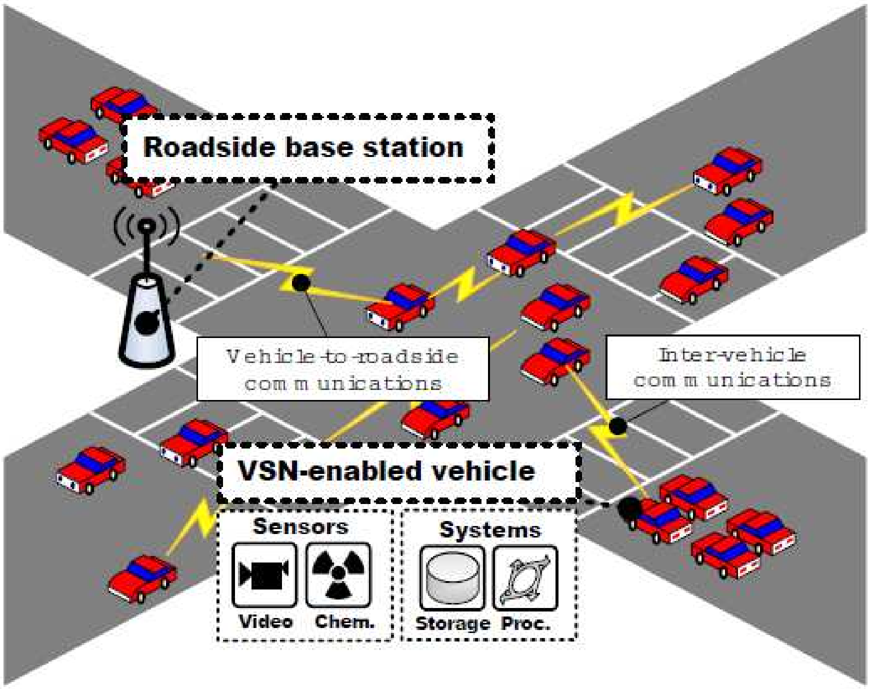


Figure 1.4: A Vehicular Sensor Network figure from [33].

1.4 Motivation

Portable computing devices are nowadays available for both people, e.g., smartphones and netbooks, and cars. The widespread diffusion of battery-operated smart terminals equipped with powerful multi-core processors, large random access memories, and wireless communication capabilities make it possible to develop sophisticated distributed applications for content dissemination to metropolitan mobile users. In fact wireless data communication capabilities of these devices make it possible to devise innovative distributed applications based on the creation of opportunistic mobile ad-hoc networks in urban scenarios.

So data collection and distribution lie at the heart of such distributed applications based on opportunistic networking: these two abstractions are implemented using protocols whose performance are influenced by numerous factors. To name a few:

- physical characteristics of the environment where the application is run: the ability of a terminal to communicate data to others within radio range is dependent on the presence of obstacles, e.g., terminals within buildings in a urban scenario may not receive data transmitted by terminals moving along the adjacent streets outside or residing on different floors of the same building.
- terminals heterogeneity: either terminals with different transmission powers (hence different radio ranges) can communicate with a number of others that depends on it, or terminals moving at different speeds and following different mobility patterns can count on a different neighborhood for communication.
- information dynamics: if data to be disseminated is time-dependent, then special storage and forwarding policies have to be designed.

To make the design of these applications even more complex, performance must be traded-off against energy consumption since mobile terminals are battery-operated. This very short and incomplete list of factors shows that the performance evaluation of protocols to support distributed applications for content dissemination as well as spreading information to metropolitan mobile users in realistic scenarios is absolutely essential.

1.5 Thesis Outline

Such applications can be based on the paradigm of data diffusion and collection, or both. In our study we focus on the following abstraction: we consider mobile and fixed terminals that are interested in collecting data items originating from several sources in a specific urban area. The area under investigation is ideally partitioned into a grid where each cell is the source of one data item.

In particular, each terminal aims at retrieving the data items in a limited *region of interest* (ROI) which is either centered around the current node position or is detached from it. In the latter case, the association between nodes and ROIs can be one to one or many to one- i.e. some nodes may share the same ROI. Clearly, for mobile nodes ROI is a time varying concept due to the dynamic behavior of pedestrians and vehicles. Since data items may change over time all nodes must strive to have access to the latest version. So we also study the influence of environment variability on the system index: obviously the more frequent are the environmental data changes the more is difficult it is to maintain the performance index high.

Terminals can basically perform two operations:

the first operation is *acquire* the data item originating in a specific location of the urban area when their current position is in its close proximity, e.g., autonomous sensing of environmental signals, reading from active sensors, passive information conveyed by RFID tag, commercial advertisements or traffic data, etc. Another potential source of local information is the presence of a grid of smart objects and/or sensors that can provide useful information associated to their position according to the Internet of Things paradigm [2].

The second operation is *exchange*: terminals can also exchange data items among them as long as the radio channel loss is lower than a given threshold. Terminals are equipped with memory to store the collected data items. The memory size depends on the type of terminal: hand-held computing devices are assumed to reserve a small memory to the application, while fixed terminals or those mounted on vehicles are assumed to use a larger memory. Furthermore, mobile nodes whose computing device is battery operated (pedestrians are assumed to use this kind of terminals) are required to cautiously transmit their data items in order to preserve energy and prolong the device operating time.

The analysis we present is based on the performance index we call *coverage*, defined as the fraction of updated data elements collected by a node compared to the total amount of data elements (eventually limited to only those accessible) within its ROI at a specific time. We present results for static and dynamic data items generation, various memory management schemes and for different classes of terminals- i.e. fixed nodes and mobile nodes (pedestrians and cars). Furthermore, we investigate the impact of increasing the randomness of mobility patterns in urban scenarios by adding artificial terminals that follow a completely random path within the considered simulation area characterized by Brownian motion with constant average speed.

We study the diffusion of information updates in the whole area, evaluate the impact of energy saving policies in the protocol version run by pedestrian devices, and the impact of contextual awareness about location and motion of nodes in the forwarding policies. One of the results we find is that most of the transmissions of these terminals are redundant and do not help in increasing the coverage of ROIs of neighbor nodes. Therefore we consider the performance of nodes when nodes operated by pedestrians transmit at a much lower rate. We show that a huge saving in overall number of transmissions (and hence in energy consumption) can be achieved at the cost of a reasonable reduction of the coverage of all node types. We also show that users within buildings may have poor performance due to the difficult of spreading information inside a building because of limited radio coverage. We evaluate the impact of the introduction of a very simple form of communication infrastructure represented by relaying nodes placed in the elevators of buildings.

Finally we analyze what happens if we dissociate ROI positions from node coordinates, removing it from node surroundings, and if we also release the constraint of the one-to-one association between nodes and relative ROI. The performance index is not good if ROIs and nodes are strongly dissociated unless few ROIs are in common among many nodes. So in some conditions each user has the possibility, simply by means of an opportunistic communication scheme, to gather enough information data about his surroundings or about a target zone.

The results we present are obtained from a simulator developed in C++ on top of the UDelModels tools [27, 49] (see Figure 1.5). These tools are used to define three-dimensional maps of urban areas and to obtain traces of mobility and radio propagation. The mobility characteristics of the nodes are based on statistical studies of population and traffic dynamics. Mobility traces generated by UDelModels are very detailed. For instance, pedestrians exhibit a motion that is representative of people in an urban scenario, with different mobility distributions for outdoor and indoor walking respectively. Moreover, a typical daily human activity cycle is taken into account. Cars mobility patterns take into account speed limits and traffic lights. The UDelModels propagation simulator is used to estimate the point to point channel loss between each pair of nodes in the three-dimensional space, taking into account the urban three-dimensional profile. The usage of such a detailed multi-dimensional mobility and propagation model for the analysis of content dissemination policies represents one of the most innovative points of our work, allowing us to identify critical features that may remain hidden when working with a simplified flat world with random mobility.

The thesis outline follows:

- *Introduction (Chapter 1)*: the current chapter;
- *Related Works (Chapter 2)*: a summary of other studies near our;
- *System Description (Chapter 3)*: a general description of system and simulation methodology as defined for our study while specific features of each set of simulations will be described at each correlated chapter;
- *Local Data Gathering (Chapter 4)*: a study of the influence of memory size, memory policy, motion type, population density, environment

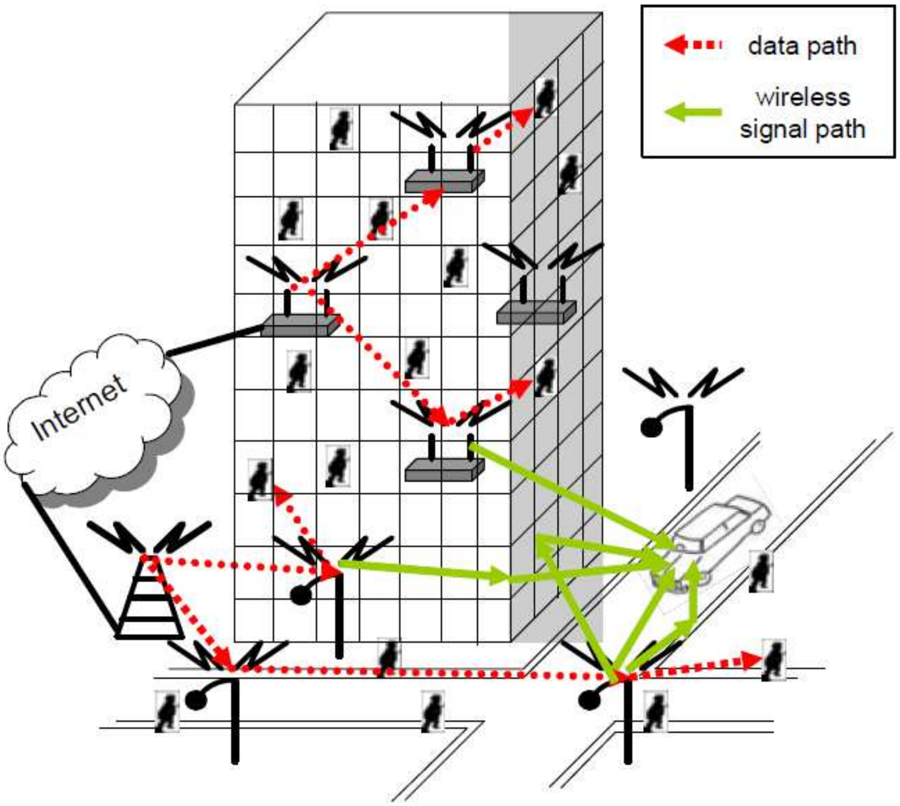


Figure 1.5: A view of urban wireless communication from [49].

variability on gathering local data using opportunistic communications;

- *Broadcasting, Delaying and Adding Nodes (Chapter 5)*: some other results with different communication protocol and transmission rates, different mobile nodes features, and environment variability type to show that our results are solid. We also study a more realistic transmission protocol based on local broadcasting;
- *Collaborative Data Retrieval with Energy Savings (Chapter 6)*: we define communication protocol and performance index in more detail and show that lowering pedestrian devices rates is a correct way to obtain energy saving with low performance index loss, starting from a study

on population behaviour classification;

- *Content dissemination (Chapter 7)* we explore the following points:
    - *Information dynamics and energy savings* : here we show that lowering pedestrian device is a good way for energy savings also under environmental variability conditions;
    - *Contextual Transmission Policies* : here we show that we can further optimize the ratio of energy savings and the loss of performance index adding context awareness to lowering transmission policy;
    - *Enhancing the building infrastructure* : here we show that some simple additional infrastructural scenario elements, such as equipping elevators with relay nodes, can significantly increase the performance index because they help overcome transmission obstacles inside buildings;
    - *ROI dynamics* : here we give some indications of conditions under which we obtain a good performance index also releasing the constraints between regions of interest and gathering nodes. In general we find that sharing target zones is a good strategy to gather information correlated to far regions of interest using opportunistic communication and local sensing;
- *Conclusions (Chapter 8)*: here the main significant results of our study and possible future evolutions are summarized.

2

Related Works

2.1 WSNs and Information Diffusion

Many fields of computer science have recently studied the problem of gathering or diffusing information. The wireless sensor networks field initially had similar aims and results to those explored in our study.

In [24] (see Figure 2.1) a three tiers architecture called MULE (Mobile Ubiquitous LAN Extensions) was proposed. MULE leverages on light sensors of limited capacity carried by people, animals, and vehicles to buffer and forward data gathered by a set of sensors to some access points. The nodes mobility is exploited to propagate the information while saving sensor energy. In the MULE architecture we have sensors with very limited memory and battery that read environmental data and can communicate only with MULE mobile nodes equipped with devices of medium capability and that have the aim by means of their mobility to forward the data to fixed nodes. The sensors have a low duty cycle to save energy. The article shows that there is a good energy saving with this architecture compared with a classical ad hoc network that does not leverage mobility of nodes, while the latency remains high enough. This study is supported by analytical and simulation methods, whereas our study is supported only by the second. The energy saving is a focus point of both studies. The aim of [24] is to transport information in an

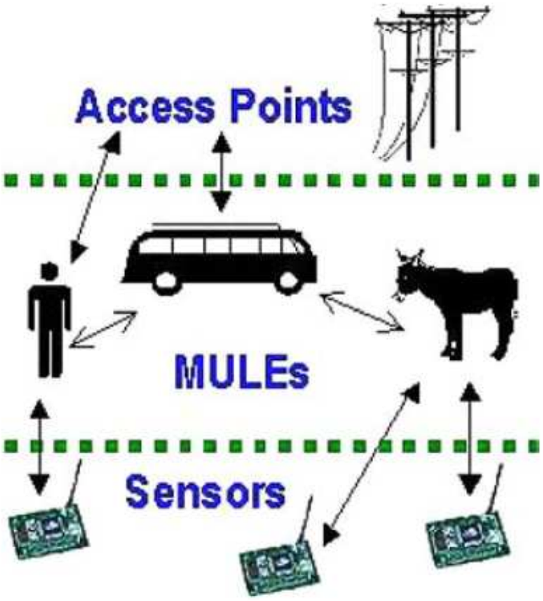


Figure 2.1: The MULE WSN architecture from [24].

rational way from some points (sensors) to others (access points) using the mobility of some nodes instead of, for example, a multi hop communication protocol, while the aim of our study is to maximize the knowledge of a specific region of interest for each node exchanging information in a opportunistic manner. Further main differences are that in [24] the scenarios are strictly bidimensional; the mobility model is not specific for urban life; the radio propagation model is simply based on distance (radial); the MULE mobile nodes are all of the same type and they have limitless buffers. All these limits make [24] not applicable to a dense urban scenario such as that used in this thesis.

The other studies ZebraNet [25] and SWIM [48] (Shared Wireless Infostation Model) employ only two layers consisting in mobile sensors and gathering points. These are similar to [24], applied to gathering biological wildlife tracking data. By exploiting animal mobility and peer-to-peer communication, the messages are passed between sensors only if sink delivery probability is increased because of resource limitations. These two works are supported by real experiments in biological systems with simulations [25] and both analysis and simulation [48], and apply their communication models respectively to zebras in a park in Kenya and to whales in open sea. Unlike in our study, their aim is to transport data to sink nodes in a multi-hop and delay-tolerant way, while our goal is to maximize a region of interest knowledge for each node and there is no special sink node. They limit the amount of traffic data with a history-based policy [25] and data packets TTL [48], whereas we use memory policy and location and context awareness. A common issue is energy consumption that in their case is limited by traffic-limiting data, while in our case it is limited by transmission delay. Their mobility model is that which is typical of zebras or whales, while our model is that of typical urban life.

Another similar article is DFT-MSN [50] (Delay/Fault-Tolerant Mobile Sensor Network), which deals with a study (supported by simulations and a few test-beds) on the gathering of data schemes that involve sensor mobility. The latter take into account their resources, mainly energy and buffer, according to delivery probability and minimizing flooding overhead. In [50], as in previous cited works, unlike our sudy, this model describes sink nodes; the mobility model is of a probabilistic type without any urban map; the radio propagation model is based on a fixed radius and the memory node is modelled simply as FCFS (First Come First Served) queue.

Similarly to wildlife motion, human and vehicular mobility is not completely random: nodes that have met frequently in recent past are likely to meet again in near future. This is the starting point for PROPHET [35] (Probabilistic ROuting Protocol using History of Encounters and Transitivity), a routing protocol which is based on a history of encounters and their transitivity property. This protocol was designed for an intermittently connected network of mobile nodes in which the aim was to use a more economic routing protocol than epidemic one. The simulation model in [35] is two-dimensional and mobilitywise it is less detailed than ours, while as a node memory management, they use a simple FIFO policy and a basic radial propagation model. Therefore in general their model is very simple compared to ours. A further improvement of PROPHET is PROPHET+ [21] in which there is the idea of taking into consideration not only the probability of encountering the destination nodes, but also some other parameters, such as energy level and free buffer size, at each transmission step. [21] uses real Bluetooth traces called iMote as a simulation model. IMote trace is a human mobility trace logged at the 2005 IEEE INFOCOM four days conference.

A distributed technique to collect data in a given area of interest is proposed in APR [40] (Amorphous Placement and Retrieval). To this end node memory is managed so as to maximize the mutual geographical distance among the stored data in order to maximize coverage area. Then each node aims at having a view of the whole field in its own cache, or by querying its neighbours. Each data also has its own TTL, so the nodes can often refresh their cache by exchanging data with neighbours and by dropping data that do not maximize distance between local information in the contents of each cache. Here energy saving is obtained lowering the need of multi-hop routings. Also this system provides a good resilience for loosely connected networks and node or packet failures, and a good performance increase enlarging memory size of node caches. Compared with [40], in our study we also use opportunistic exchange of samples between pairs of neighbours to diffuse information. However, we have a different target. We aim at maximizing the knowledge of a region of interest associated with nodes, so the memory policy adopted is designed in this direction, and not for having the whole view of the field encapsulated in a few near nodes as in [40]. Also for us the memory size of cache nodes is considered a precious resource. Further in [40] the authors also suggest applying their algorithm to fields in which some zones are of high interest, but they always consider all the nodes as having the same levels of interest for the different zones of the field, while we are interested in a more various map of interests among nodes. In [39] an information theory approach is presented to optimize the caching mechanism used by mobile devices. In this case every node is assumed to query for the information in a given point in space, assuming a certain probability distribution for location of the queries. In [40] and [39] the scenario is a simple two-dimensional random mobility model and also the radio propagation model is very simple, whereas the cache management, which is very sophisticated, has a different aim from ours.

In STDCS [1] (Spatio-Temporal Data-Centric Storage scheme) the problem of answering efficiently real-time geo-centric ad-hoc queries is considered. In this proposal a sensor for local storage is elected and may vary with time for energy load-balancing; a point-to-point routing scheme is used to deliver the reading of any sensor to its storage sensor. The simulation model in [1] is based on fixed sensors read by mobile user devices, simple radial radio propagation model and a two-dimensional area.

In MobiQuery [36] temporal and spatial constraints of sensors queries are considered as well. As in our study also in [36] one of the main objectives is to continuously gather real-time information in one's vicinities. However, in this study temporal constrains are explored more deeply. Examples of indicated implementation are: a fireman requesting a periodic update of a temperature map of the surrounding area or an autonomous intelligent mobile device in a search and rescue operation, such as in a disaster zone, querying continuously for information of its surroundings as it moves in a unknown dynamic environment. Similarly to our work, in [36] nodes can operate under low duty cycles; both analysis and simulation are used to support the model. The mobility simulation in [36] is simply random with velocity between human walk and moderate speed vehicle, but the nodes are all equally equipped. The radio propagation model is simply distance based.

2.2 DTNs and Node Mobility

Several papers have analyzed the reliability of information diffusion that aims at exploiting the mobility and the random contacts among mobile terminals. Among these, one similar to the spirit of our study is exposed in [34] where there is a study of a content distribution application in which mobile nodes are interested in acquiring information from a few fixed nodes used as source. The authors compare some distribution schemes by simulation, using mobility traces obtained by tracing movements of Cambridge University students. In the end, they find that the best results can be achieved when the mobile nodes collaborate between themselves in spreading information. This result can be further improved by using occasional mobile contacts with other portable devices outside the target group of students. In this study the authors measure the diffusion and delay of a digital content whose source are a few fixed nodes, while our performance index is the average knowledge that nodes have of a given region of interest environmental information, however we also aim at showing that collaboration among nodes improves the final result.

The importance of cooperation among nodes in information spreading is also confirmed in [43] and [26]. In [43] the authors theoretically study the packet delivery rate within a time TTL in DTN under different routing protocol (Epidemic, Two-hops, Binary Spray and Wait) and degrees of node cooperation (from fully cooperative, to cooperative with some probability, to defective in forwarding) and they also show that a modest level of cooperation considerably increases network delivery rate. On the other hand, in [26] it is shown that in mobile DTNs, although consistent percentage of mobile nodes are not cooperative, meaning that they do not contribute in message delivery in order not to consume their resources, there is no dramatic loss in network performance. Therefore controlled non-cooperative behaviour may be used to reduce resource consumption.

In [45] the authors individuate and comment some common errors that are generally made in many trace-based simulations of opportunistic networks. In fact many of these type of studies do not take into consideration cache limits of mobile devices and other technological limitations, which bring to network performance overestimation and to additional backbone usefulness underestimation. Our study underlines the importance of technological limits of mobile devices, for example by using a minimal memory for each portable device, and supporting the additional backbone usefulness by adding some semi-fixed infrastructural nodes in order to overcome communication obstacles.

In [7] the authors study a node location system based on directional gossiping rather than on flooding, obtaining a good energy saving. They also improve this system by considering the random mobility of some nodes which carry information around in order to save more wireless transmission energy. In such system each node periodically exchange information with neighbours about its own position: in this way each node cache contains a list of contacted nodes and recently updated data of their position. So queries about a node-destination location can be gradually directed to node destination. The latter can answer its own position in an updated way. In this paper ([7]) the information contained in node caches is the contacted nodes position and not an environmental data as in our study. Indeed in this case the target of the queries is the position of a given node, and not the knowledge of a region of interest as in our thesis. However, also in this study cache and transmission policies help to obtain searched information, and mobility helps to spread information. Further on mobility helps transmission energy savings, whereas our study also focuses on the design and evaluation of transmission policies for energy saving. On the other hand in ([7]) the node mobility is based on a simple probabilistic model in a two-dimensional scenario and the the radius of radio propagation model is fixed, which is very different from our more sophisticated models.

Theoretical bounds on gossip performance and information diffusion speed in mobile systems were obtained in [46],[41] and [23], respectively. The capacity of ad hoc wireless networks is evaluated under general mobility conditions in [16]. Other approaches for the modeling and exploitation of nodes mobility in wireless networks are presented in [29] and [3], where random graph theory is employed.

2.3 DTNs and Buffer Management

In Hamlet [14], the authors try to maximize the hit ratio of queries issued by mobile nodes searching for cached content in other nodes memory. They adopt a clever memory policy especially in the nodes in which information passes through during data hop paths. Each node decides to memorize information in transit and to associate a drop time to it, on the base of its presence in the neighbouring nodes. The latter is estimated according to the distance in hop count between itself and the source, and itself and the destination. The result is the creation of a content diversity within the nodes neighborhood. The simulation models used in [14] are more detailed than those in previous cited articles, and in some features more similar to those used in our study. However, the mobility models used in [14] are either of only pedestrian motion or only vehicle motion type, both only of an outside type. In addition, the radio range is fixed, so a very simple radio propagation model is used.

In LACMA [31] (Location-Aided Content Management Architecture) the authors propose a system that associates the information of nodes of a mobile system to a location by means of layers of grids of different side sizes according to information of different popularity. This system tries to maintain a constant density of information in each cell grid by using opportunistic communication among nodes, and information replication when this density varies because of the mobility. The node mobility is supposed to be checked by GPS. In [31] the simulation models used are less sophisticated than our models. There is only one class of mobile nodes, and no urban map is described. The aim of both Hamlet and LACMA is to maximize the cache hit probability while minimizing the number of hops of the content queries, whereas the goal of our study is to maximize the nodes knowledge of an associated region of interest, but in all these studies the goal is reached by means of the content of the nodes and their neighbours memory.

A similar approach is adopted in [28] where epidemic dissemination and buffer management policies are used to optimize the average delivery rate and delay. In this article ([28]) it is shown that by using epidemic dissemination with an oversimplified drop memory policy, especially when the memory is full, it is not possible to achieve good results in delivery rate and delay. This is the case when some limits are assumed into a memory device, as it is typical of most DTN networks. So a sophisticated history based drop memory policy is elaborated in order to obtain network performance optimization. This drop memory policy is tested with simulations based on simple probabilistic mobility models as well as with two different type of real traces, collected previously from wildlife and taxis tracking, with the result of increasing the network performance. In our study we show that, even if we adopt a different target and performance index, elaborated drop memory policy can substantially help performance index, above all if memory limits are strong.

2.4 Opportunistic Social Networks

Another field of interest that has given us some good point of view is that of OSNs (Opportunistic Social Networks). For example in [15], starting from real traces of contacts collected in a university workplace, a classification of nodes based on frequency and duration of contacts among people is proposed. This taxonomy includes friends, strangers, and others classes and reflects the degree of relations. On this basis a social awareness forwarding scheme is designed to maximize the efficiency of collaborative data dissemination. In a similar way in a social based forwarding algorithm, BUBBLE [22], real traces are analyzed in order to underline the importance of the centrality degree and the belonging to a community, with the aim at individuating a good social awareness forwarding scheme. In the same vein, based on WiFi, GPS and Second Life traces, the work in [53] classifies nodes in two main classes: socials, which are frequently in a determined area, and vagabonds, whose presence in that area is more sporadic and less regular. The contribution to information dissemination of each class is considered and the key role played by vagabonds is highlighted. In [42] the role of OSNs nodes with high contact rate but low degree of membership of a community in time is underlined. Communities in time are defined in this work by means of a temporal contact graph model. In [42] the experimental data are four Bluetooth contact traces in some campuses or conferences. On the other hand, in our study we do not consider social features of nodes but their position, mobility and change of context in order to increase network performance. The emergence of opportunistic computing [9], as a way of exploiting human networks in network communications, indicates a growing interest in the relationships between social networks and opportunistic networking/computing.

2.5 The UDel Models Tools

In [27] e [49] the authors describe a realistic model for pedestrian and vehicle mobility, which is based on essays and statistical studies of real motion and people's habits. Such studies are implemented in UDel Models tools. We use these tools as a realistic mobility basis for our work. Their work includes a realistic propagation model that studies the wireless propagation very deeply. It values the effect on propagation and channel loss of a lot of factors such as walls, construction materials, corners, diffraction, reflection, mutual position and height and presence of obstacles. Moreover, the internal structure of buildings, according to the type of building, is taken into consideration in order to provide a radio propagation model that is able to calculate the channel loss between nodes, in a sufficient detailed way also inside buildings. The outside and inside radio propagation are simulated in two different integrated ways, taking into consideration that frequencies used for mobile networking are partly absorbed, partly reflected off and partly passed through the building peripheral walls. Therefore, the outside and inside environments are not fully separated as far as radio propagation is concerned, and the two propagation models have to communicate with each other. The outside propagation model is based on a more accurate beam tracing technique, a computational simplified model compared to 3-D ray tracing model, whereas the indoor propagation model is based on an attenuation factor model, taking into consideration the attenuation of the ray each time it passes through an interior wall or a floor. The whole relative tool is computationally very heavy because it is based on a very detailed model. Based on this tool, we establish if the channel loss among objects enables or not the communication among them, for each second of mobility simulation. In [27] great importance is given to the influence of map features on human and vehicle mobility, for example the dislocation of aggregation points, such as offices, restaurants and traffic lights. Also timing is taken in account by considering typical time frames, for instance, the time span in which people commute to and from their workplace, and the breaks they take. In addition, the behaviour of a number of people who are not working is also considered. City-map features and the influence these have on human and vehicle mobility are also considered. Vehicles move around roads regulated by traffic lights, whereas humans move both outside, along pavements, and inside, on different levels of buildings.

Moreover, this study takes into consideration the tendency that both vehicles and humans have of forming small clusters. The former group up because of traffic constraints (red traffic lights, speed limits and congestion).

The latter have a natural propensity for gathering together in small groups (conversations, queues, etc.). What is more each specific time of day has a specific density of vehicles and people, respectively. These are calculated on the basis of both time and initial input data, which refers to the total numbers of people and vehicles.

2.6 Further Human and Vehicle Motion Models

To base our study on such e realistic motion model as [27] is important because some features of real world motion can be of great influence to network performances as studied in [44]. In the latter it is shown that in an outdoor scenario, the human walks can be succesfully simulated by means of a truncated form of Levy walks also commonly noticed in animals. This is based on some tens of GPS traces. The fact that a realistic human mobility model is important in mobile networks is also the starting point of other simulation tools, such SLAW [30] (Self-similar Least Action Walk) and SWIM [38] (Small World In Motion). In [30] truncated power-law distributions of flights, pause-times and inter-contact times, fractal way-points, and heterogeneously defined areas of individual mobility are included in the mobility model called SLAW, which can generate artificial human walk traces that are effective in representing social contexts present among people sharing common interests, or those in a single community. Similar features are implemented in SWIM. Human mobility is also studied in its general form in [19], in this the trajectories of 100,000 anonymized mobile phone users tracked for some months are studied. In [19] it is shown that each person is characterized by a typical travel distance and a trend to return to a few highly frequented places, unlike in random paths generated with random and Levy flight based models. In [4] it is underlined that CDRs (Call Detail Records) from a cellular network are a useful mean to collect a huge amount of data on human mobility. Maximizing network performance by exploiting some typical features of the motion of real world mobility objects (pedestrians as vehicles) motion, such as the trend of forming a group, is also justified in [8]. In this, the performance are delay and/or throughput and not a degree of knowledge of a local, environmental information. In [8] it is shown that correlated node movements sometimes lead to better performance than the one achievable under independent node movements. Therefore the node motion correlation is not a feature of secondary importance. Indeed, UDel Models mobility model described earlier, used as a starting point in our study, takes into consideration the trend of grouping of both vehicles, due to traffic constraints, and of people, due to their behavioural habits.

Finally we take in great consideration MobEyes platform ([33], [32] and Figure 1.4), in which is considered a VSN (Vehicular Sensor Network) that has as main example of applicative aim the surveillance of an urban area by means of cameras and sensors on cars. The general problem described is the speed diffusion of concise summaries of the whole scenario and, on demand, of the integral data about a single critical event. The main aims of MobEyes are to leverage car mobility in the city, by considering non-restrictive computational and memory limits, and to avoid congestion problems because of massive data diffusion, in order to have an effective possibility of monitoring an urban area for security reasons. In fact computational vehicle on-board resources are used to extract concise significant summaries from integral memorized information, which are then spread by a Bloom filter based protocol in an effective manner.

3

General System Description

In this chapter we outline the general system model, the simulation methodology, the performance index used and the common settings in their main features, while their variations and particularities for each different set of simulations are exposed in each related chapter.

3.1 Our System

In this section the general characteristics of the system under analysis are presented. In an urban area, the performance of data exchange protocols among mobile nodes (pedestrians and vehicles) and fixed terminals is examined by characterizing each communication node in according to its mobility characteristics, radio coverage and amount of memory devoted to the data exchange protocol.

We consider the following classes of nodes:

- *fixed* (F) nodes, i.e., wireless relay nodes or access points;
- *pedestrians* (P) nodes, that carry portable devices with limited power and memory capacities;
- *vehicular* (V) nodes that move faster along trajectories constrained by roads.

The nodes are placed in a three-dimensional space where roads and buildings with multiple levels are modelled.

We focus on the collection of data items located in a urban area U. We assume that a two-dimensional data set is associated with U. In particular, each $\delta \times \delta$ tile of U is characterized by a data item $i(x, y, t)$, where the integer coordinates (x, y) are the tile indexes and t represents a time-stamp associated with the information. The information $i(x, y, t)$ can be composed by a set of environmental measurements taken autonomously by the mobile nodes or communicated by an infrastructure of active sensors, passive information conveyed by RFID tag, commercial advertisements or traffic data, etc. A possible source of data could be the myriad of sensors and smart objects with computational and wireless communication capabilities that are described in the Internet of Things paradigm. As an example, we can think about information of social utility, e.g., traffic and mobility management, public services and events advertising, etc. The area U is populated by wireless fixed and mobile nodes whose objective is to retrieve the data in a limited *region of interest* ROI centered around the node position. For simplicity in the following we assume that the ROI is represented by all the information $i(x, y, t)$ within a $\Delta \times \Delta$ tile around each node.

As mentioned above, our model includes several classes of nodes, namely *fixed* (F) nodes and mobile nodes which are further divided into *pedestrian* (P) and *vehicular* (V). By F nodes we mean wireless relay stations placed at fixed location, e.g. at intersections between blocks and building entrances. Nodes P and V are mobile wireless terminals characterized by different mobility patterns. P nodes move along streets at walking speed and typically concentrate within buildings, while V nodes are cars moving faster along the streets.

The interactions, i.e. the mobility and the radio coverage of F, P and V nodes, are modelled by the UDelModels tools [27]. UDelModels is a suite of tools for simulating urban mesh networks and includes a simulator of urban radio propagation and a simulator of realistic urban mobility. UDelModels simulate the mobility of P and V nodes in a 3D space exploiting statistical studies of population and traffic dynamics. This approach is far more precise than simple 2D models, where nodes are generally placed at random on a grid and move by using some variant of Brownian motion. The UDelModels simulator takes as inputs a three-dimensional map of the urban area, the number of nodes in the different classes and the statistical parameters of the mobility.

Pedestrians exhibit a motion that is representative of people in an urban scenario, with different mobility distributions for outside and inside walking, respectively. Moreover, a typical daily human activity cycle is taken into account.

Car mobility patterns take into account speed limits and traffic lights.

The UDelModels propagation [49] simulator is used to estimate the point to point channel loss between each pair of nodes in the three-dimensional space, taking into account the urban three-dimensional profile. These estimates are used to predict the radio contacts that each node is able to establish with others. Two nodes are assumed to be able to communicate with one another, if the channel loss between them is lower than a threshold α.

In Figure 3.1, there is an example of simulated urban area where node ROIs are represented as squares centered around sample P and V nodes. Note that the squares are of different dimensions; in Figure 3.2 the whole simulated area is shown as seen from two different UDel Models applications. Further details on UDelModels tools are described in Chapter 2.

The data stored in the buffer $\mathcal{B}$ of a given node come from two sources: the information $i(x, y, t)$ directly acquired from the environment when the node steps inside the tile at coordinates (x, y) and the information gossiped by other nodes during occasional radio contacts thanks to their mobility.

Different buffer management strategies have been included in our model. The simplest is based on a First In First Out (FIFO) approach, with updat-

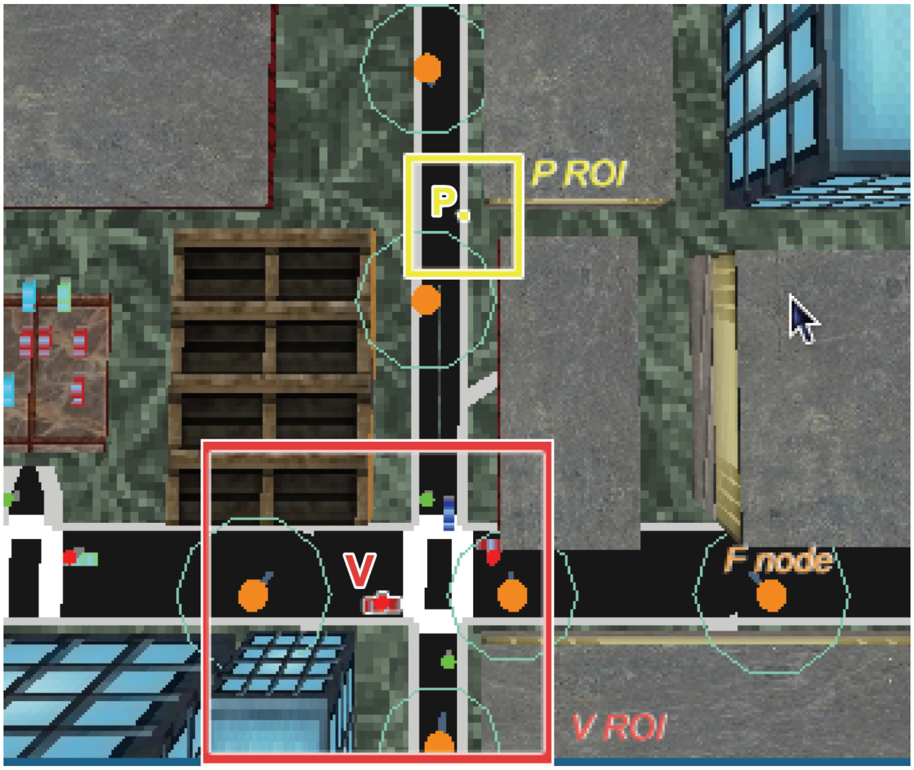


Figure 3.1: A zoom of the simulated urban area: node ROIs are depicted as squares.

ing of the already known information at a given location (x, y). With this strategy each piece of received information $i(x, y, t)$ is pushed into $\mathcal{B}$ with a FIFO policy if no information about the position (x, y) has been included yet; on the contrary, if $i(x, y, t_0) \in \mathcal{B}$ with $t_0 < t$ then, the corresponding record is updated. In this way the system can represent both static and dynamic information associated with the environment.

Clearly, this basic approach is quite simple but does not enforce any locality awareness check on the content of $\mathcal{B}$. In this research we analyze two improvements, referred to as *selective dropping* (SD) and *selective insertion* (SI), that aim at prioritizing the storage of the local data. When using SD the data can be popped out of $\mathcal{B}$ only if they refer to a location outside the node's ROI. Similarly, the SI prevents from pushing into $\mathcal{B}$ any data outside the node ROI. The memory policy obtained applying both SD and SI is called SD/SI (or SI/SD).

As already mentioned, the simulation is based on the UDelModels tools [27, 49] for the computation of the nodes mobility and radio propagation. In particular, UDelModels is used to generate realistic node trajectories and to compute the communication channel loss between any two radio stations in the three-dimensional urban scenario with the granularity of 1 s. The UDelModels results are fed into an ad-hoc simulator, designed by us in C++, whose basic functions are:

- to simulate the motion of nodes following the relative UDelModels trace;
- to simulate the data sensing by each mobile object in the current position;
- to simulate the radio contacts among the nodes as far as the channel loss is below a threshold α;
- to implement the management of the buffer $\mathcal{B}$ depending on memory policy selected;
- to simulate the exchange of collected items among the nodes in radio contact following a given protocol;
- to estimate the system performance in terms of the coverage of the node ROIs;

further particular features are added to each chapter depending on specific system model and simulation settings.

Indeed each chapter representing a set of simulations has an associated system whose general characteristic are described above, but with other special features. These are:

- in Chapters 4 and 5 there are some additional random walk nodes with Brownian motion;
- in Chapter 4 there is an exchange data protocol of unicast type, while in Chapters 5, 6 and 7 a broadcast type is used;
- in Chapters 5, 6 and 7 delaying or stopping pedestrians device transmissions is experimented because of energy saving;
- in Chapter 7 some infrastructural node are positioned in simulated elevators to help inside information spreading and different ROI dynamics are experimented;

obviously these additional features are associated with some specific parameters and settings.

3.2 Performance index : coverage

The performance index in which we are interested is represented by the percentage of the ROI covered by items stored in the local buffer $\mathcal{B}$ of each node. In the following we will refer to such percentage as *coverage* ($0 \leq C \leq 1$). The value of the coverage clearly varies from node to node and it is time dependent. In simulations with environmental variability of information we add to the calculation of coverage the constraint of taking into consideration only updated information.

In Chapters 6 and 7 the cooperation between nodes of different class types is underlined. Thus we refine the *coverage* definition taking into consideration only *accessible* grid cells of each ROI both in static and dynamic environmental scenarios.

In fact there is no certainty for all parts of our map to be accessible. Excluding from *coverage* calculation inaccessible sub-areas generally involves a more significant value for those nodes that often pass close to unapproachable zones.

Our analysis is generally based on the estimate of the average ROI coverage experienced in different node classes; the transient time is excluded from the computation to avoid biasing the estimates, and the averages are calculated on a sufficient number of independent trials to obtain statistically meaningful results.

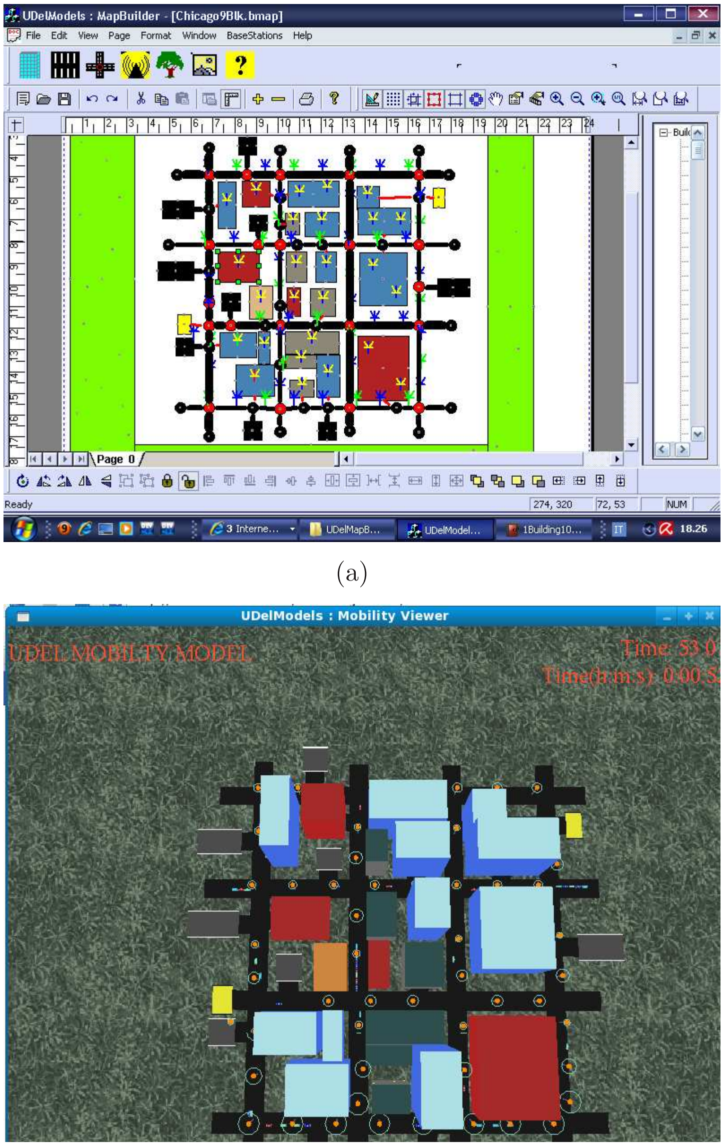
3.3 Common settings

Some settings that are common to all simulations are described in the following.

We consider an urban area of 550 m × 500 m, representing 9 blocks of Chicago[1]. The fixed wireless infrastructure comprises 54 F nodes placed at road intersections and building entrances. The size of the ROI depends on the node class, with $\Delta = 400, 200$ and 100 m for F, V and P nodes, respectively; while all scenarios will be partitioned in a bidimensional informational grid with each cell of side $\delta = 25$ m. The memory constraint B depends on the size of the corresponding ROI: fixed nodes have the largest amount of resources while pedestrians are equipped with a small sized memory. The mobility of V and P nodes is simulated with UDelModels: the cars speed is in the range $(25, 67)$ km/h, whereas the pedestrian walk in the range $(2.5, 6.5)$ km/h. All the other mobility parameters are set to the default values. The simulations are worked out in a time interval of three hours and the initial 15 minutes are always considered as transient and therefore skipped in the computation of the average coverage C.

---

[1] The map data are available at *http://udelmodels.eecis.udel.edu/*



(a)

(b)

Figure 3.2: The simulated urban area views in two different UDelModels applications: a) Map Builder to edit the map of the scenario and b) Mobility Viewer to visualize node motions of a simulation.

4

Local Data Gathering Using Opportunistic Networking

This chapter is about a set of simulations whose main aim is to study how our local data gathering system works with opportunistic networking according to population density, presence of cars, presence of nodes with Brownian motion, different hand-held devices memory size and simple policies of memory management usage, with and without the presence of environmental variability.

4.1 Specific system

In this section the characteristics of the specific system under analysis are presented along with the assumptions we made to model the interactions among the fixed communication relays and the tetherless nomadic terminals, while a more general system description is in Chapter 3.

We remember here only that each $\delta \times \delta$ tile of the whole scenario U is characterized by a data item $i(x, y, t)$, where the integer coordinates (x, y) are the tile indexes and t represents a time-stamp associated with the information. The area U is populated by wireless fixed and mobile nodes whose objective is to retrieve the data in a limited ROI centered around the node position. As in the general case we assume that the ROI is represented by all the information $i(x, y, t)$ within a $\Delta \times \Delta$ tile around each node.

Our model consists of various classes of nodes:

- *fixed* (F): wireless relay stations placed at fixed location as street intersections and building entrances;
- *mobile nodes*:
  - *pedestrian* (P): transit along streets at walking pace and usually concentrate within buildings;
  - *vehicular* (V): move faster because they are mounted on cars;
  - *random walk* (R): the R nodes are artificial entities that follow a completely random path.

These nodes are characterized by Brownian motion with average speed equal to V_R. This latter class is introduced in order to study the effects of nodes with unconstrained mobility on the performance of opportunistic data exchange and collection.

The interactions, i.e. the mobility and the radio coverage of F, P and V nodes, is modelled by the UDelModels tools [27], as already seen in Chapter 3. In some simulation in this chapter as said above we add R nodes that have a specific communication pattern of behaviour; in fact the R nodes represent the only exception to our system because they are included in our model as ideal entities. Thus they do not use the general precise beam tracing and attenuation model, but an R node is assumed to be able to communicate with all nodes within the range of a fixed radius A_R around the R node's current position.

4.1.1 Data collection strategies

The goal of the present study is the analysis of the efficiency of the data collection strategies used by the nodes to cover their respective ROI in presence of limited buffering capability B, defined as the number of data items that a node is able to store in its local memory $\mathcal{B}$. Each node can collect all the information of interest, provided that $B \geq \left(\frac{\Delta}{\delta}\right)^2$. The two sources of data stored in each node's buffer $\mathcal{B}$ are:

- the information $i(x, y, t)$ directly acquired from the environment when the node steps inside the tile at coordinates (x, y);

- the information gossiped by other nodes during occasional radio contacts thanks to their mobility.

In all our analysis we assume that the time is slotted. In each time slot the node acquires the information $i(x, y, t)$ corresponding to its position and randomly communicates with neighbours to acquire other potentially useful information. The data collection efficiency, i.e. the percentage of data collected by the node that fall within its current ROI, are clearly influenced by the strategy used to exchange data and the policy adopted to refresh the buffer.

In this set of simulations the exchange policy we adopt is simple and completely random: each node randomly selects k_n nodes in its radio coverage and pushes to each of them k_d randomly chosen data items stored in its $\mathcal{B}$. This is only one of many possible patterns of communication. In the following chapter, we show that our results are also valid for a different and more realistic pattern of communication among neighbours based on broadcasting. During each time slot a given node receives a variable number of items from other nodes in the vicinity and collects one item associated to its position.

As already shown in Chapter 3 the different memory policies that we simulate are : FIFO (*First In First Out*), SD (*selective dropping*), SI (*selective insertion*) and both SI/SD.

As previously described, the latter three introduce some sort of locality-awareness check, while the first is one of the simplest possible.

It is worth pointing out that the locality check we introduced does not guarantee a priori a good performance to mobile nodes, as the ROI is not statically known but varies according to the node mobility.

4.1.2 Settings

The performance index we are interested in is represented by the percentage of the ROI covered by the items stored in the local buffer $\mathcal{B}$ of each node. We have called this index *coverage* ($0 \leq C \leq 1$) and it is described in Chapter 3.

Indeed, when the simulation begins $\mathcal{B} = \emptyset$ and $C = 0$ for all the nodes. The goal of the collection policies we analyze is to keep C as high as possible all the time, excluding the initial transient phase. All our analysis will be based on the estimation of the coverage C, defined as the average ROI coverage experienced in the different node classes: fixed, vehicular and pedestrian, respectively; the transient is excluded from the computation to avoid biasing the estimates. Finally, all the reported results are always averaged on 20 independent simulation trials so as to obtain statistically meaningful results.

UDelModels calculate trajectories of the nodes and radio channel loss between them for each second. The UDelModels results are used as input by our ad-hoc simulator whose major functionalities are:

- to simulate the radio contacts among the nodes providing channel loss is below a threshold α;
- to implement the management of the buffer $\mathcal{B}$;
- to simulate the exchange of collected items among the nodes in radio contact;
- to simulate the mobility and radio propagation of R nodes, which are not available in UDelModels; the direction of the trajectory of the R nodes is updated randomly every second and their speed is drawn according to a uniform distribution;

Table 4.1: System settings.

Parameter	Value		
U	550 m × 500 m		
δ	25 m		
Δ	F nodes	V nodes	P, R nodes
	400 m	200 m	100 m
B	$b\left(\frac{\Delta}{\delta}\right)^2$, with $b \geq 1$		
α	-30 dB		
V_R	$\mathcal{U}(9, 11)$ m/s		
A_R	20 m		
k_n	5		
k_d	5		
Simulated time	from 7 to 10 a.m		
Time slot	1 s		
Memory policy			
FIFO	SD	SI	SI/SD

- to estimate the system performance in terms of coverage of node ROIs.

The simulation results shown in the following are worked out using the settings summarized in Table 4.1. Here, the main settings of this chapter set of simulations are specified, while common settings are exposed in Chapter 3.

The simulated area is an urban district of Chicago. The size of the ROI depends on the node class, with $\Delta = 400, 200$ and 100 m for F, V and P/R nodes, respectively. The memory constraint B depends on the size of the corresponding ROI: fixed nodes have the largest amount of resources while pedestrians are equipped with a small sized memory. The R nodes move according to Brownian motion with a random speed distributed according to a uniform distribution $\mathcal{U}(9, 11)$ m/s. Each node pushes $k_d = 5$ data items in its buffer to $k_n = 5$ neighbours, at the most. All the other mobility parameters are set to the common or default values. The simulations are worked out in a time interval between 7 and 10 a.m. and the initial 15 minutes are always considered as transient and therefore skipped in the computation of the average coverage C.

4.2 Results

In this section we report on a set of experiments, whose goal is the analysis of the sensitivity of the proposed data collection policies to crucial feature of the system, i.e. the density and the kind of nodes mobility, the amount of nodes memory, and the usage of buffer management policies.

The first phenomenon that we study is about the sensitivity of C with respect to the density of mobile nodes, in particular P nodes. Locality awareness is not used and the simple FIFO buffer management policy is adopted. In Figure 4.1 the values of C for P and F nodes are reported versus the number of P nodes; no vehicles are simulated in this case. In Figure 4.1(a) an ideal situation when all nodes have no memory constraints ($b = \infty$) is shown, whereas Figure 4.1(b) refers to the opposite condition when all nodes use the minimum amount of memory ($b = 1$). From the figures it can be noted that increasing the density of P nodes improves the performance of all node classes. Nonetheless, the performance gain turns out to be rather limited when not enough memory is available. In terms of absolute values P nodes always achieve the best coverage mainly because for now we do not have excluded inaccessible zones to the calculation of performance index coverage, as we do for example in Chapter 6.

In Figure 4.2 the coverage achieved by the nodes is studied as a function of the memory used by the P nodes. In this case we make the realistic assumption that F nodes do not have any memory limits whereas the pedestrians, being equipped with hand-held devices, can use a limited amount of buffering space for opportunistic communications. Figure.4.2(a) and Figure 4.2(b) refer to two different scenarios with 125 and 500 P nodes, respectively. This experiments show that all node classes benefit from a higher availability of memory at the mobile nodes.

We also consider the effect of different mobility patterns of the mobile nodes. To this end, in Figure 4.3 we show some experiments worked out in a scenario where the overall number of mobile nodes is kept fixed to 250, with a variable percentage of vehicles and pedestrians. In particular, we let the number of vehicles be the 0, 10 and 20 % of the mobile nodes. In Figure 4.3(a) no memory limits are used and it can be noted that vehicles help in improving the coverage of all node classes. In Figure 4.3(b) the same experiment is repeated in the case $b = 1$, i.e. when all the nodes use the minimum amount of memory. In this latter situation it can be noted that the coverage of P nodes tends to decrease when a higher number of cars is around. This is clearly due to the fact that in all these experiments the nodes do not implement any form of locality awareness when managing their buffers. Since the V nodes move faster than P nodes, these latter are likely to store in their buffer useless data item transported by V nodes coming from a different area that in turns decreases their coverage.

The adoption of the SD and SI locality aware policies allows one to overcome the issue we have just commented on. Table 4.2 compares the values of C obtained in the different node classes in the scenario with 200 pedestrians and 50 vehicles as a function of different system settings, namely the FIFO, SD and SI buffer management with and without memory limitation. The SD/SI column refers to the case when the two buffer management strategies are activated jointly. It is worth noting that the adoption of SD or SI highly improve the coverage yielding a good performance even in the presence of most severe memory constraints, i.e. $b = 1$ for all node classes. As an example, from the results in Table 4.2 it turns out that using SD or SI both P and V nodes are able to achieve a coverage in the case $b = 1$ that is very close to what they would get using unconstrained memory.

To complete the study on different mobility pattern we added to the system a limited number of R nodes with unconstrained mobility. In Table 4.3

Table 4.2: C for different buffer management policies in the scenario with 200 P and 50 V nodes.

Class	Policy				
	FIFO $b = 1$	SD $b = 1$	SI $b = 1$	SD/SI $b = 1$	FIFO $b = \infty$
P	0.30	0.55	0.57	0.58	0.68
V	0.15	0.45	0.49	0.49	0.48
F	0.53	0.56	0.57	0.57	0.56

Table 4.3: C with the presence of 0,1 and 5 R nodes in the scenario with 200 P and 50 V nodes.

Class	Settings		
	FIFO, $b = \infty$ no R nodes	SD, $b = 1$ 1 R node	SD, $b = 1$ 5 R nodes
P	0.68	0.71	0.77
V	0.48	0.57	0.64
F	0.56	0.68	0.77

the values of C obtained by adding 1 and 5 R nodes when $b = 1$ are compared with those obtained without R nodes and $b = \infty$. The scenario is the same as in Table 4.2. By comparing the two tables, it can be noted that adding a single R node increases the coverage of P, V and F nodes from 0.55, 0.45, 0.56 up to 0.71, 0.57, 0.68. Using 5 R nodes enhances the performance further.

In all previous experiments we assumed that the information $i(x, y, t)$ associated to the urban area is kept fixed for all the 3 hours simulation. Clearly, adding a temporal dynamic to the information has a negative impact on the overall system performance. To take this effects into account, we performed a last set of simulations where one information item is updated every d s. The computation of the coverage index C is slightly modified so as to consider only the *updated* information. In other words, the coverage is now defined as the percentage of updated information collected by the node buffer that falls within the ROI. The obtained results are shown in Figure 4.4, that reports the coverage as function of the update interval d in the scenario with 200 P, 50 V and 5 R nodes using SD and $b = 1$. As expected, one can observe that frequent updating of the information negatively impact on C. Nonetheless, we can point out that in the presented scenario the proposed system is able to sustain 1 update every 30 s without a noticeable performance impairment.

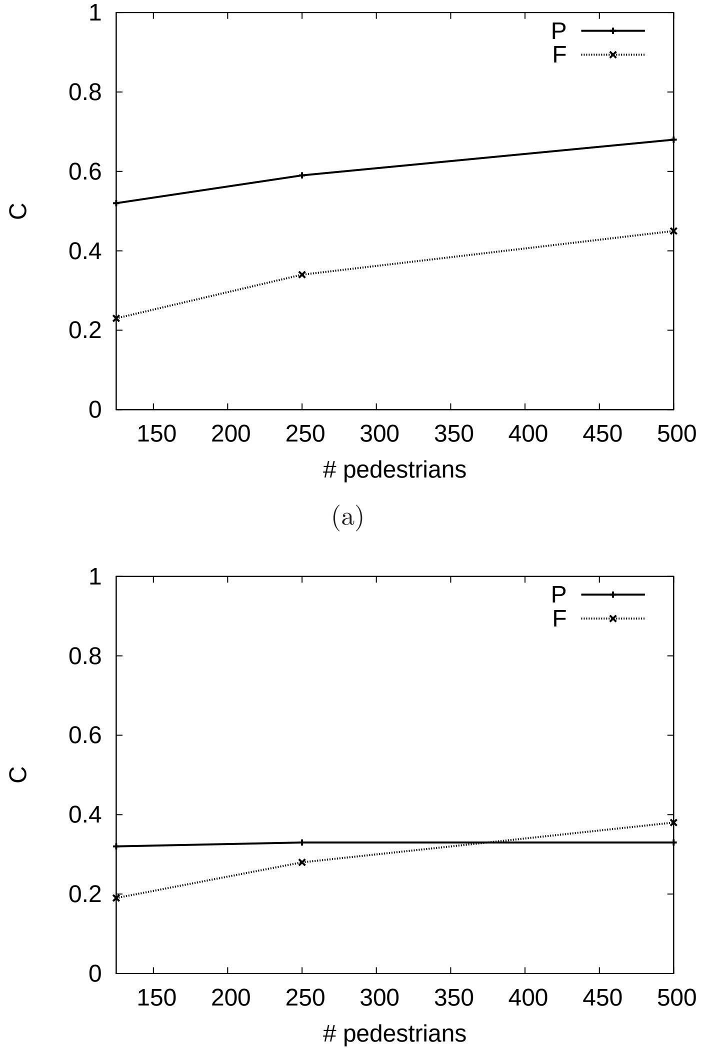
4.3 Summary

In this chapter we describe our results published in [12] (PE-WASUN 2011). This was our first study on the argument of the thesis.

We analyze the performance of the system by defining and estimating the percentage of the ROI covered by the items stored in the buffer of each node (the node coverage) for both static and dynamic information. Nodes memory is limited and it is proportional to the nodes computing resources, i.e., hand-held terminals are assumed to have small buffers while vehicles and fixed nodes use larger storage. Each node senses and gathers information data within the range of its current position. The communication scheme is as simple as possible and of a unicast type. We underline the positive correlation among index performance, called *coverage*, and memory devices sizes and population density by means of initial simulations.

The preliminary findings highlight that simple location-awareness memory management schemes effectively exploit nodes with limited amount of memory. Furthermore, by adding a few ideal nodes whose mobility is described by an unconstrained Brownian motion, increasing randomness of node movement proved to have a beneficial impact on the average coverage of all node types.

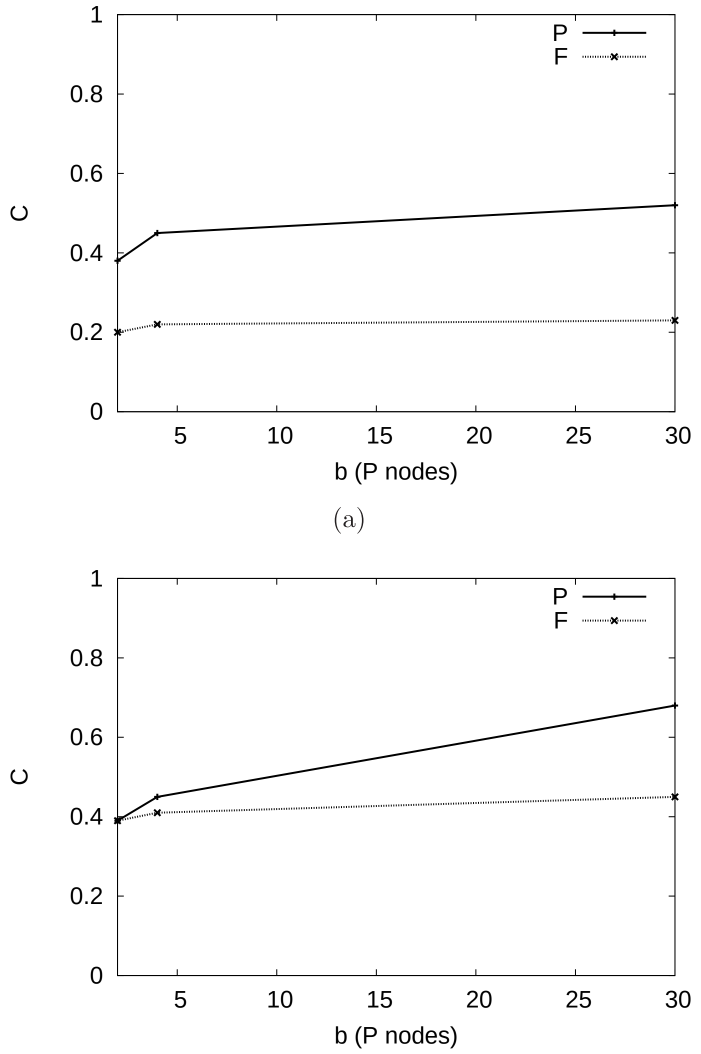
Another result described is the negative correlation between environmental data variability and our performance index in the case of dynamic environment, or changes in data associated to the scenario.



(a)

(b)

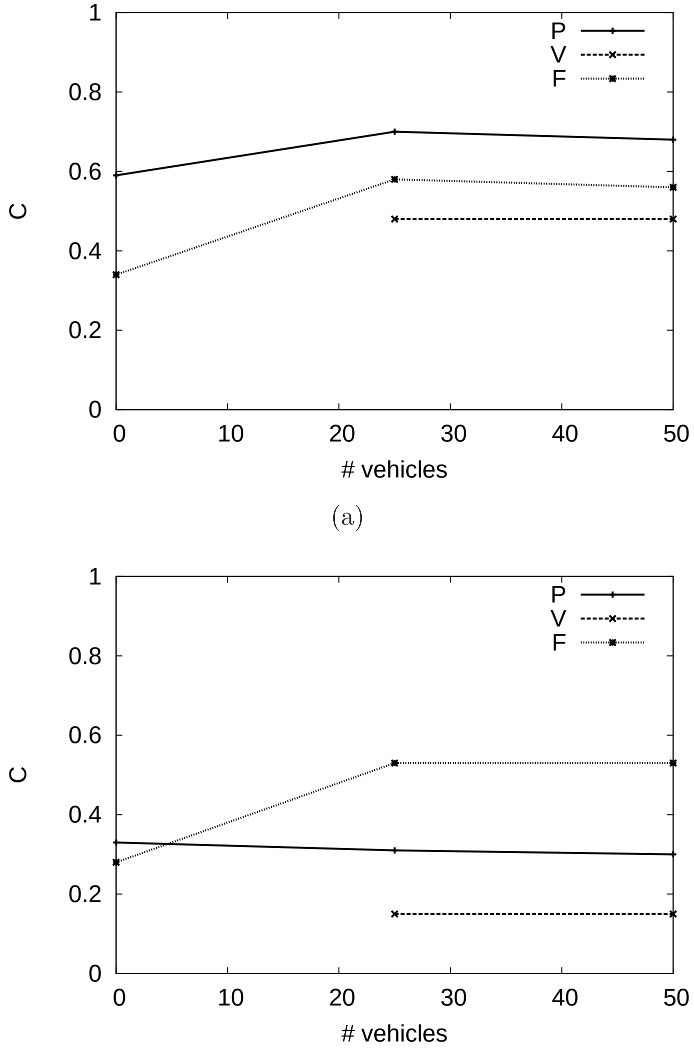
Figure 4.1: C for P and F nodes as a function of the number of P nodes when $b = \infty$ (a) and $b = 1$ (b) and using FIFO buffer management.



(a)

(b)

Figure 4.2: C for P and F nodes as a function of b for the P nodes ($b = \infty$ for F nodes) in the 125 (a) and 500 (b) pedestrians scenarios.



(a)

(b)

Figure 4.3: C for P, V and F nodes as a function of the number of cars over a total number of 250 mobile nodes using $b = \infty$ (a) and $b = 1$ (b).

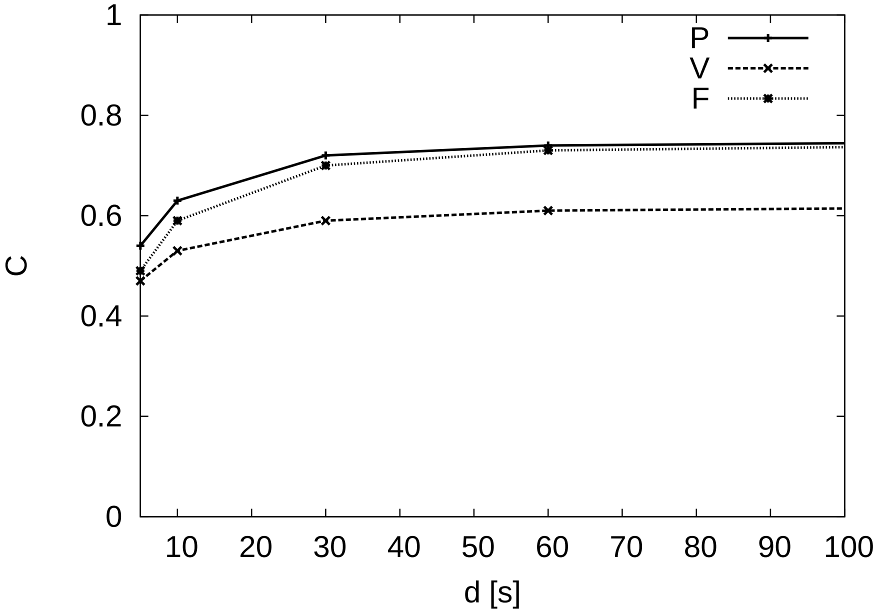


Figure 4.4: C for all node classes versus the update interval d, in the scenario with 200 P, 50 V and 5 R nodes using SD and $b = 1$.

5

Broadcasting, Delaying and Adding Nodes

In this chapter, we want to show that the results found in Chapter 4 are reliable, by showing that similar results are also obtained in different settings.

We also want to experiment with some other communication protocols and/or infrastrucural changes, which will serve as a starting point for the next chapters regarding more realistic and/or energy saving transmission policies and other features.

5.1 Specific system and settings

In this section the system model and settings differences between Chapter 4 and this chapter are underlined, in order to provide evidence for previous results and new material for further studies, while a more general system description is in Chapter 3.

We vary the transmission policy by broadcasting to neighbours every $T = 0.2$ (basic T) seconds for a generic node, while the pedestrian devices transmissions can be delayed. So T is 0.2 as a base, whereas it ranges from 0.2 to ∞ for delayed pedestrian devices. This type of transmission is nearer to wireless nature of transmission than that described in the previous chapter.

We also adopt it in the following chapters.

We do not explicitly consider collisions among concurrent transmissions of nearby nodes. We argue that this approximation is based on the following reasoning:

> Let us assume that N transmitters share a wireless channel whose capacity is M b/s. Each transmitter has D bits to transmit in one transmission that may occur at any point in time in an interval of S seconds. Transmission of D bits requires $H = \frac{D}{M}$ seconds to complete. Clearly, the probability of two colliding transmissions is equal to $p_c = \frac{2H}{S}$. A more accurate expression can be obtained by considering border effects that show up if transmissions of either nodes start within the first or the last D seconds in the time interval of length S: in this case we obtain $p_c = \frac{6HS-4H^2}{2S^2}$. We can now derive the probability that N transmitters do not collide as $p_{ok} = (1-p_c)^{\binom{N}{2}}$. In our case, as calculated in detail in Chapter 7, is $p_{ok} = 0.973$ which is rather high and confirms that neglecting collision phenomena represent an acceptable approximation.

The urban scenario is the same studied in Chapter 4 and generally described in Chapter 3. It is a 9 block district of Chicago from 7 to 10 a.m., while the total mobile population consists of 200 pedestrians and 50 vehicles and eventually some additional random walk nodes. Here the random walk nodes parameters have some differences in respect to settings in Chapter 4 in Δ, V_R, A_R as reported in Table 5.1. We also apply some differences to environment variability, by adding the parameter *ncc*(*number of cluster's cells*), the number of contiguous cell values that vary each d seconds. By default this parameter (*ncc*) can be assumed as 1 unless specified differently.

Coverage C is calculated in the same way as in Chapter 4: the ratio for each node of updated information elements of its own memory that are related to grid cells within its ROI(*region of interest*) over the total number of

Table 5.1: System settings.

Parameter	Value			
U	550 m × 500 m			
δ	25 m			
Δ	F nodes	V nodes	P nodes	R nodes
	400 m	200 m	100 m	200 m
B	$b\left(\frac{\Delta}{\delta}\right)^2$, with $b \geq 1$			
α	-30 dB			
V_R	$\mathcal{U}(9,11)$ or $\mathcal{U}(4.5,5.5)$ m/s			
A_R	10 m or 20 m			
basic T	0.2			
delayed T	0.2 , 1 , 10 , ∞			
k_d	5			
Simulated time	from 7 to 10 a.m			
Environmental variability	d		ncc	
	30,60,300, ∞ s		0,1,5 cells	
Memory policy	FIFO , SD , SD/SI			
# nodes	54 F	200 P	50 V	0, 1, 5, 10 R

cells of the same ROI at each moment. The performance index is the average for each node class of the so defined coverage after the first 15 minutes of simulated time.

The memory policies used (FIFO (*First In First Out*), SD (*selective dropping*) , SI (*selective insertion*), SD/SI) are those described in Chapter 3.

5.2 Results

The results of these tests confirm the main previous results, already published in PE-WASUN 2011 and showed in Chapter 4, also for broadcasting propagation.

In particular SD/SI memory policy is better than SD or FIFO policy with minimal memory (see Tables 5.2 and 5.3); we experiment that random

Table 5.2: C for different buffer management policies in the scenario with 200 P and 50 V nodes.

Class	Policy			
	FIFO $b=1$ $T=0.2$	SD $b=1$ $T=0.2$	SD/SI $b=1$ $T=0.2$	FIFO $b=\infty$ $T=0.2$
P	0.31	0.56	0.58	0.69
V	0.13	0.46	0.49	0.50
F	0.48	0.56	0.57	0.57

Table 5.3: No trasmission by pedestrian devices ($T=\infty$), C for different buffer management policies in the scenario with 200 P and 50 V nodes.

Class	Policy			
	FIFO $b=1$ $T=\infty$	SD $b=1$ $T=\infty$	SD/SI $b=1$ $T=\infty$	FIFO $b=\infty$ $T=\infty$
P	0.27	0.35	0.40	0.50
V	0.13	0.33	0.33	0.34
F	0.42	0.43	0.43	0.43

Table 5.4: C with the presence of 0,1,5 and 10 R nodes in the scenario with 200 P and 50 V nodes.

Class	Settings			
	FIFO, $b=\infty$ no R nodes $T=0.2$	SD, $b=1$ 1 R node $T=0.2$	SD, $b=1$ 5 R nodes $T=0.2$	SD, $b=1$ 10 R nodes $T=0.2$
P	0.69	0.72	0.78	0.81
V	0.50	0.62	0.68	0.80
F	0.57	0.73	0.82	0.88

Table 5.5: No trasmission by pedestrian devices ($T = \infty$), C with the presence of 0,1,5 and 10 R nodes in the scenario with 200 P and 50 V nodes.

Class	Settings			
	FIFO, $b = \infty$ no R nodes $T = \infty$	SD, $b = 1$ 1 R node $T = \infty$	SD, $b = 1$ 5 R nodes $T = \infty$	SD, $b = 1$ 10 R nodes $T = \infty$
P	0.50	0.58	0.68	0.72
V	0.34	0.54	0.63	0.77
F	0.43	0.66	0.77	0.86

Table 5.6: One cell element change each d secs, C with the presence of 5 R nodes in the scenario with 200 P and 50 V nodes.

Class	Settings		
	SD, $d = \infty$ 5 R nodes $T = 0.2$	SD, $d = 60$ 5 R nodes $T = 0.2$	SD, $d = 30$ 5 R nodes $T = 0.2$
P	0.78	0.76	0.73
V	0.68	0.65	0.62
F	0.82	0.78	0.75

Table 5.7: Changing ped. devices trasmission delay (T), one cell element change each d secs, C with the presence of 5 R nodes in the scenario with 200 P and 50 V nodes.

Class	Settings			
	SD, $d = \infty$ 5 R nodes $T = 1$	SD, $d = \infty$ 5 R nodes $T = 10$	SD, $d = \infty$ 5 R nodes $T = \infty$	SD, $d = 60$ 5 R nodes $T = \infty$
P	0.78	0.75	0.68	0.66
V	0.68	0.66	0.62	0.59
F	0.81	0.80	0.77	0.73

Table 5.8: Changing ped. devices trasmission delay (T), C with no R nodes in the scenario with 200 P and 50 V nodes.

Class	Settings			
	SD, 0R $T = 0.2$	SD, 0R $T = 1$	SD, 0R $T = 10$	SD, 0R $T = \infty$
P	0.57	0.56	0.50	0.36
V	0.46	0.44	0.42	0.32
F	0.56	0.55	0.53	0.43

Table 5.9: Changing ped. devices trasmission delay (T), C with the presence of 10 R nodes in the scenario with 200 P and 50 V nodes.

Class	Settings			
	SD, 10R $T = 0.2$	SD, 10R $T = 1$	SD, 10R $T = 10$	SD, 10R $T = \infty$
P	0.81	0.81	0.78	0.72
V	0.80	0.80	0.78	0.77
F	0.88	0.88	0.87	0.86

Table 5.10: Changing R node properties , C with the presence of 5 R nodes in the scenario with 200 P and 50 V nodes.

Class	Settings		
	SD, 5R V_R=$\mathcal{U}(9, 11)$ m/s A_R=20 m $T = 0.2$	SD, 5R V_R=$\mathcal{U}(9, 11)$ m/s A_R=10 m $T = 0.2$	SD, 5R V_R=$\mathcal{U}(4.5, 5.5)$ m/s A_R=20 m $T = 0.2$
P	0.79	0.77	0.74
V	0.67	0.67	0.62
F	0.82	0.81	0.74

Table 5.11: No trasmission by ped. devices ($T = \infty$), changing R node properties , C with the presence of 5 R nodes in the scenario with 200 P and 50 V nodes.

Class	Settings		
	SD, 5R V_R=$\mathcal{U}(9, 11)$ m/s A_R=20 m $T = \infty$	SD, 5R V_R=$\mathcal{U}(9, 11)$ m/s A_R=10 m $T = \infty$	SD, 5R V_R=$\mathcal{U}(4.5, 5.5)$ m/s A_R=20 m $T = \infty$
P	0.69	0.65	0.61
V	0.62	0.60	0.53
F	0.77	0.75	0.64

Table 5.12: ncc(number of near cells in cluster) elements change each d secs, C with the presence of 5 R nodes in the scenario with 200 P and 50 V nodes.

Class	Settings			
	SD, $d = \infty$ $ncc = 0$ 5 R nodes $T = 0.2$	SD, $d = 60$ $ncc = 1$ 5 R nodes $T = 0.2$	SD, $d = 300$ $ncc = 5$ 5 R nodes $T = 0.2$	SD, $d = 60$ $ncc = 5$ 5 R nodes $T = 0.2$
P	0.78	0.76	0.76	0.67
V	0.68	0.65	0.65	0.57
F	0.82	0.77	0.78	0.66

Table 5.13: No trasmission by ped. devices ($T = \infty$), ncc(number of near cells in cluster) elements change each d secs, C with the presence of 5 R nodes in the scenario with 200 P and 50 V nodes.

Class	Settings			
	SD, $d = \infty$ $ngc = 0$ 5 R nodes $T = \infty$	SD, $d = 60$ $ngc = 1$ 5 R nodes $T = \infty$	SD, $d = 300$ $ngc = 5$ 5 R nodes $T = \infty$	SD, $d = 60$ $ngc = 5$ 5 R nodes $T = \infty$
P	0.69	0.66	0.67	0.57
V	0.62	0.59	0.60	0.52
F	0.77	0.73	0.74	0.63

walk nodes have a positive influence (see Tables 5.4 and 5.5) also with a different Δ value in respect to that used in Chapter 4 (now it is set to 200 instead of 100). Moreover, the random walk nodes positive effect increases by augmenting the number of R nodes (see Tables 5.7, 5.8 and 5.9), their radius of transmission and/or velocity (see Tables 5.10 and 5.11).

What can be also noticed in the results shown in Tables 5.13 and 5.12 is that, for some intervals of parameter values, the *coverage* remains almost constant if the ratio between d (interval between contiguous environmental changes) and ncc (number of cluster's cells of each environmental change) does not vary.

Here we also give a first clue that lowering pedestrian device transmissions by increasing the delayed T parameter, for example for energy saving purposes (see Table 5.8), we only have moderate loss in system performance measured in terms of *coverage*. A deeper insight into this matter is found in the next chapters.

In addition, we have studied our system with pedestrian device transmissions set to zero ($T = \infty$) (see Tables 5.3, 5.5, 5.11 and 5.13), having obviously a similar but more stressed results than those obtained only by lowering pedestrian device transmissions.

Furthemore, the time of simulation was extended to all working day, but the conditions were restarted every three hours. This test is a full transmission test ($T = 0.2$) with no environmental change ($d = \infty$), and with a full starting population of 200 pedestrians and 50 cars. However, the active populations of both groups are different depending on the time of day, so also the coverage changes.

Figure 5.1 shows a different performance at different times of day. Each point highlighted in the graph represents the average of a three-hour time slot. In this case, and in those that follow, *coverage* is calculated taking into account only accessible subareas of each ROI zone, to avoid penalizing nodes that often pass by inaccessible areas.

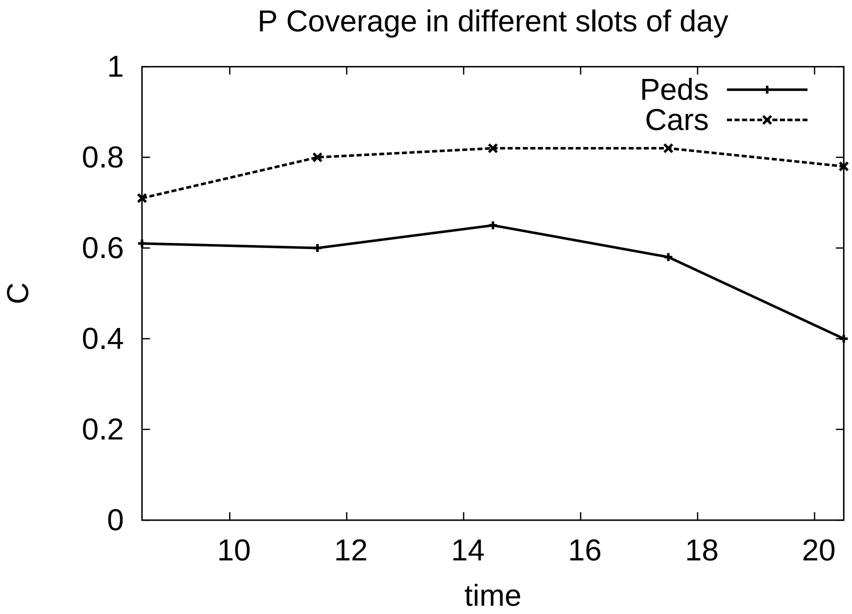


Figure 5.1: Average coverage for pedestrians and vehicles for the population mix composed of 200 pedestrians and 50 vehicles at different times of a working day, with different density of active populations in different times, restarting conditions every 3 hours and calculating averages for consecutive periods of 3 hours each.

So far ROIs are strictly linked with their nodes and, as shown in Figure 5.3, the coverage distribution vs time is regular, so in these cases the average can be a useful performance index.

For a visual scheme of indoor radio propagation the reader can see Figure 5.2 from [49], bearing in mind that in our model nodes communicate with each other only if channel loss between them is less than -30 dB.

Finally we tested some trials inserting a new infrastructural element that will be fully developed only in Chaper 7: elevators equipped with powered relay nodes. The *coverage* increases as elevators node sensitivity (measured in dB) increases, as shown in Figure 5.4.

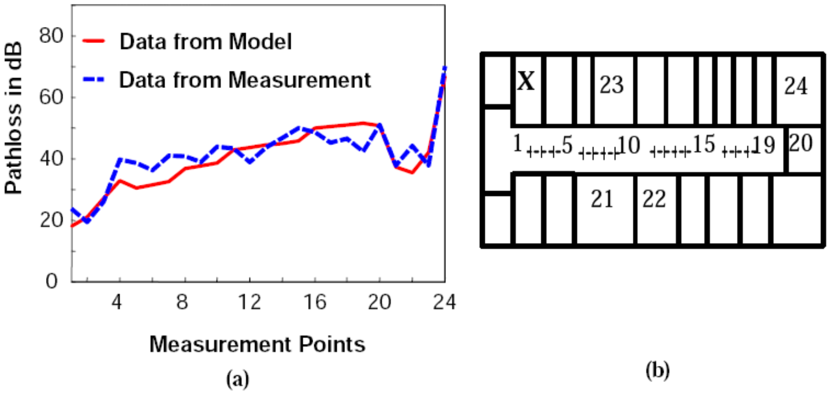


Figure 5.2: Observed and estimated path loss in an indoor environment. The measurement points shown on the x-axis correspond to the numbered locations in the map. The map is of the third floor of Evans Hall in University of Delaware campus. See [49].

5.3 Summary

In this chapter we enforce the results of Chapter 4 using other tests in similar conditions. For example, we change the setting conditions of random walk nodes and environmental data variability.

Here we also experiment an opportunistic transmission policy based on *local broadcasting*. Indeed it is nearer to real wireless device-to-device communication. In such a protocol we decide to ignore collision problems, after a mathematical analysis of collision probability resulting almost zero.

Moreover, we delay pedestrian devices transmissions with the aim of energy savings for hand-held and portable devices, and carry out tests with elevators equipped with relay nodes of growing sensitivity. All these tests will be useful as a starting point in the following chapters.

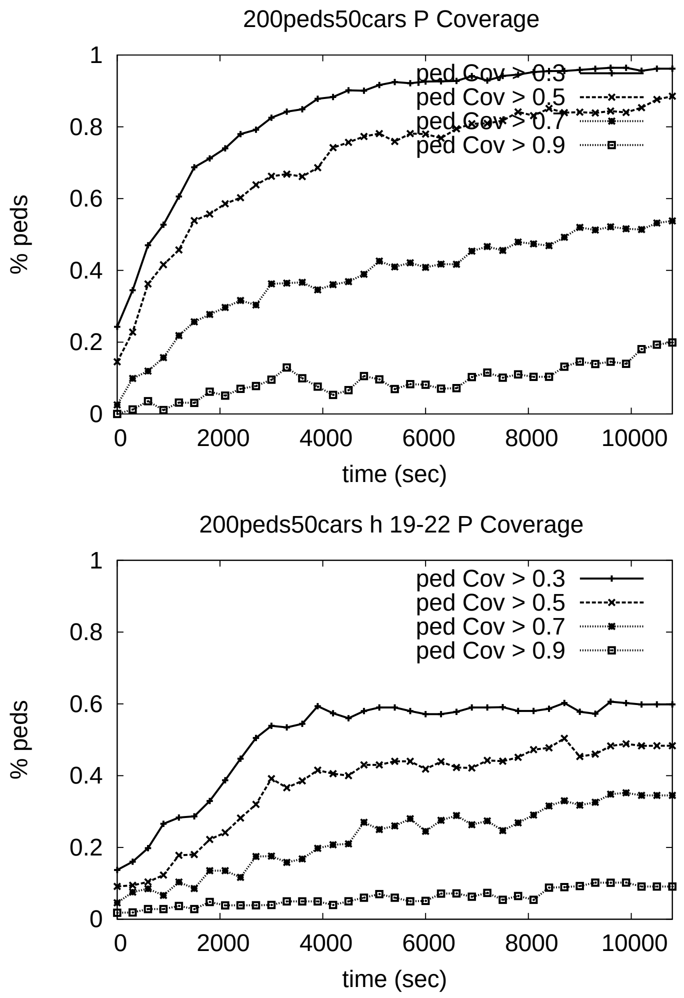


Figure 5.3: Average active population percentage of coverage limit reached for pedestrians (peds) during hours 7-10(up) and 19-22(down) in the scenario with 200 pedestrians and 50 vehicles.

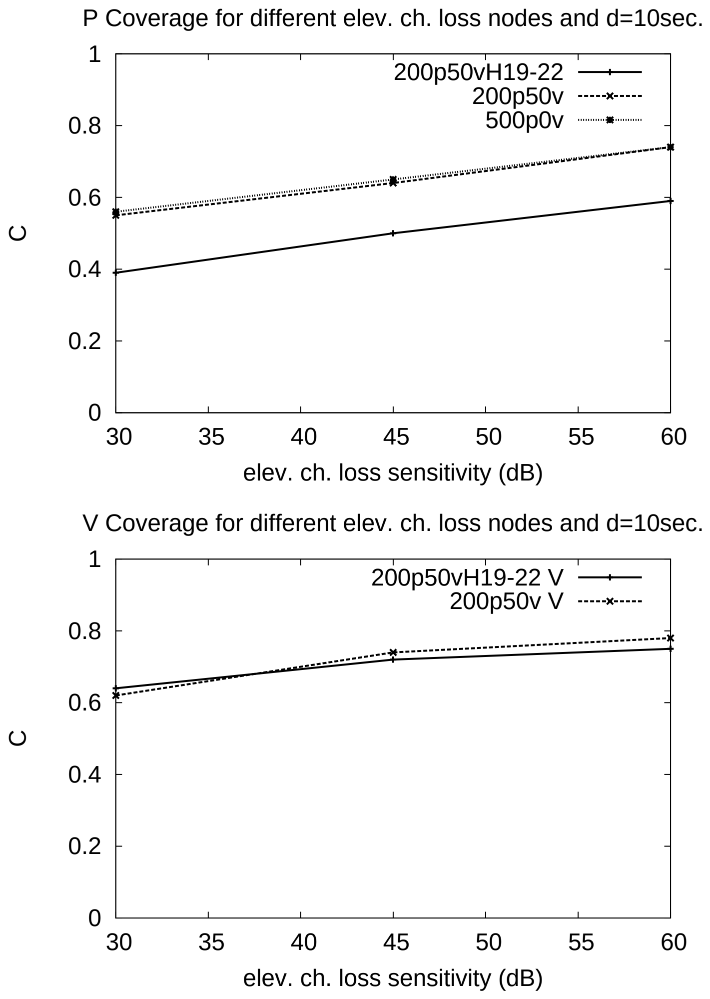


Figure 5.4: Average coverage for pedestrians(up) and vehicles(down) with environmental changes (d=10sec) with elevator nodes working at alternate floors with 1 minute average stopping time and different sensitivity of equipped nodes. Three different scenarios are shown: 200 pedestrians and 50 vehicles at 19-22 and 7-10 respectively and 500 pedestrians without vehicles at 7-10.

6

Collaborative Data Retrieval and Energy Saving Using Opportunistic Networking

In this chapter we define *coverage* as the fraction of updated data items collected by a node with respect to the total amount of *accessible* data items within its ROI at a specific time. We have added the notion of accessibility to general definition in order to better compare the resulting index of different classes of nodes.

We also show that vehicles perform better than pedestrians and that the main reason lies in the different mobility models describing these two types of nodes.

Finally we obtain as a simulation result that pedestrian device transmission rates can be lowered with great energy saving and not excessive loss of coverage. This is because a great number of pedestrian device transmissions does not contribute to coverage increase.

6.1 Specific system

In this section we present some assumptions made on the opportunistic communication model in a urban scenario.

The goal of present work is to analyze the performance of opportunistic data gathering policies exploiting both mobile nodes and fixed access points distributed in a urban area. In particular, a rectangular grid is considered and a piece of information $i(x, y, t)$ is associated to each grid position (x, y) an time-stamp t. The time-stamp is introduced to let the information vary from time to time. The integer coordinates (x, y) refer to a square area of $\delta \times \delta$ meters. The information $i(x, y, t)$ can be composed by a set of environmental measurements taken autonomously by the mobile nodes or communicated by an infrastructure of active sensors, passive information conveyed by RFID tag, commercial advertisements or traffic data, etc. A realistic urban map is superimposed to the rectangular grid considering parking lots, residential blocks and roads.

Different node classes are then considered. *Fixed* (F) nodes, i.e., wireless relay nodes or access points, are placed at road intersections and building entrances. Two kind of mobile nodes with very different mobility and communication/memory/CPU capacity features are included in the model, namely *pedestrians* (P) and *vehicles* (V). P nodes carry portable devices with limited power and CPU capacities and move along streets at walking speed, typically concentrating inside buildings. V nodes move faster but their trajectories are constrained on roads; moreover, we can assume that this class of nodes has no strict limitations in terms of power and CPU since car-installed devices can be considered.

The goal of all the node classes is to retrieve as much as possible information around their current position. In particular, we assume that each node exploits a limited memory space capable of storing only B information items and aims at knowing all the information within a $\Delta \times \Delta$ square *Region* *of Interest* (ROI) centered on the node position. Clearly, the ROI of mobile nodes is time varying.

6.1.1 Data collection and sharing strategies

The information items can be collected in two ways. Each mobile node can collect and buffer the information $i(x, y, t)$ acquired from the environment thanks to its mobility. More importantly, all nodes can broadcast some buffered items to the others exploiting their occasional radio contacts. In all our analysis we assume that the time is slotted. In each time slot a node acquires the information $i(x, y, t)$ corresponding to its position (if it is still missing). Moreover, each node can use a simple and random broadcasting policy to propagate its own information to the neighborhood. To limit power consumption broadcasting can be activated only every T seconds. In this simulation we try to extend T mainly for pedestrian devices because for them energy savings are very important. When broadcasting, each node randomly selects k_d information items stored in its buffer and propagate them. Therefore, during each time slot a given node may receive a variable number of items from other nodes in the vicinity.

Different buffer management strategies have been considered in previous chapters and described in Chapter 3; here we mainly use SD/SI (both selective dropping and selective insertion) that is associated with better performance.

The performance index we are interested in is represented by the percentage of accessible ROI covered by the items stored in the local buffer of each node. We have called this performance index as *coverage* ($0 \leq C \leq 1$) and already described. The value of the coverage clearly varies from node to node and it is time dependent. Furthermore, some subareas of the urban scenario are not accessible to nodes, i.e., due to some physical obstacle, therefore C is computed as the number of items in local buffer over the total number of *accessible* subareas in the node ROI.

Table 6.1: System settings.

Parameter	Value		
Simulated area	550 m × 500 m		
δ	25 m		
B	$\left(\frac{\Delta}{\delta}\right)^2$		
k_d	5		
Simulated time	from 7 to 10 a.m		
	F nodes	V nodes	P nodes
Δ	400 m	200 m	100 m
T	0.2 s	0.2 s	0.2, 10, 60, 600, ∞ s
$Memorypolicy$	SD/SI		
$\#mobilenodes$	250		

6.1.2 Simulation methodology

UDelModels simulate the mobility of P and V nodes in a three-dimensional space exploiting statistical studies of population and traffic dynamics. The number of nodes in the different classes and the statistical parameters of the mobility model can be properly configured. Pedestrians exhibit a motion that is representative of people in an urban scenario, with different mobility distributions for outside and inside walking, respectively. Moreover, a typical daily human activity cycle is taken into account. Cars mobility patterns take into account speed limits and traffic lights.

The UDelModels propagation [49] simulator is used to estimate the point to point channel loss between each pair of nodes in the three-dimensional space. The channel loss is used to model the radio contacts among the communicating nodes. Any two nodes are assumed to be able to communicate each other if the channel loss is below a threshold α (in all our experiments α was fixed to -30 dB).

The UDelModels results are used as input by our ad-hoc C++ simulator whose major functionalities are:

- to simulate the radio contacts among the nodes as far as the channel loss is below a threshold α;
- to implement the management of the nodes buffer;
- to simulate the broadcast of collected data items among the nodes in radio contact;
- to estimate the system performance in terms of the coverage of the nodes ROI.

We do not explicitly consider collisions among concurrent transmissions of nearby nodes, in fact the broadcast system implemented is equivalent to that described in Chapter 5.

6.1.3 System settings

Indeed, when the simulation begins the local buffer of nodes is empty and $C = 0$ for all the nodes. The goal of the data collection system we analyze is to keep C as high as possible all the time, excluding the initial transient phase. All our analysis will be based on the estimation of the average coverage C, defined as the average ROI coverage experienced in the different node classes: fixed, vehicular and pedestrian, respectively; the transient is excluded from the computation to avoid biasing the estimates. Not accessible subareas also are excluded from coverage computation to improve comparisons between nodes of different type. Finally, all the reported results are always averaged on 20 independent simulation trials so as to obtain statistically meaningful results.

The simulation results shown in the following are worked out using the settings summarized in Table 6.1.

The memory constraint B depends on the size of the corresponding ROI: fixed nodes have the largest amount of resources while pedestrians are equipped with a small sized memory. Each node broadcasts $k_d = 5$ data items in its buffer to its neighbors. Two successive transmission are separated from T seconds. The mobility of V and P nodes is simulated with UDelModels and is described in Chapter 3 together with other common settings.

All the other mobility parameters are set to the common or default values. The simulations are worked out in a time interval between 7 and 10 a.m. and the initial 15 minutes are always considered as transient and therefore excluded from the computation of the average coverage C.

6.2 Results

Figure 6.1 shows the coverage distribution for pedestrians (up graph) and vehicles (down graph) in all three population mixes. It turns out that both pedestrians and vehicles display similar results for different mixes. It can be noted that vehicles always obtain better coverage than pedestrians: the average coverage for vehicles ranged from 0.70 to 0.74 while pedestrians obtained results ranging from 0.60 to 0.61.

It is worth to take a closer look to the class of pedestrians. These nodes can be further classified into *indoor* and *walking* pedestrians.

Figure 6.2 shows the coverage distribution for these two subclasses (indoor on the up and walking on the down).

It can be noted that performance of indoor pedestrians is much worse than walking but since indoor pedestrians represent about 95% of the total population then the whole class performance is weighted down. The reason for worse performance of indoor pedestrians lies in two factors: first, limited nodes movement reduces the amount of novel useful information and second buildings act as shields that isolate indoor pedestrians into clusters that do not receive fresh information brought by vehicles passing by.

The ROI of two close nodes almost coincide. If these two nodes are rather static (as the indoor pedestrians use to be) it is likely that a fraction of data received by nodes is redundant because it is already stored in the nodes memory. To better gain insights into this phenomenon we considered the population mix composed of 200 pedestrians and 50 vehicles (the other population mixes yield similar results). In this case, the average number of data reception per second for pedestrians is 21.3 while it is 120.6 for vehicles; clearly, these numbers are a function of the average number of neighbors of the nodes in one class. For pedestrians 84% of all receptions is from other pedestrians, 6% from vehicles and remaining share from fixed nodes. For vehicles this distribution is different: 3% of receptions comes from pedestrians, 64% from other vehicles, and the rest from fixed nodes. A further refinement shows that only 25% of receptions for pedestrians and 29% for vehicles are not redundant, i.e., are not already stored in the nodes buffer. Referring to pedestrians, only 12% of receptions are not duplicates and come from other pedestrians. For vehicles, only 0.2% of receptions are not duplicates and come from pedestrians. This raises a natural question: are all transmissions made by pedestrians really useful? To answer this question we experimented by slowing down transmission rates of pedestrian devices. In particular, we separate two successive transmissions of a pedestrian device by a fixed amount of T seconds. Figure 6.3 shows that shutting down pedestrians (they never transmit) makes the average coverage drop from about 0.60 to about 0.42. On the other hand, slowing down transmission by one minute yields an average coverage of about 0.52. For vehicles the same delay interval decreases the average coverage to about 0.62. Slowing down transmissions from pedestrians lowers energy consumption which is a critical issue for hand-held devices: our experiments show that sixty times lower transmissions from pedestrians yield acceptable reduction in the average coverage of both vehicles and pedestrians.

6.3 Summary

In this chapter we describe the results published in [11] (WONS 2012).

Mobile nodes (operated by pedestrians and vehicles) and fixed relays cooperate to achieve the goal of retrieving the data items in their region of interest (ROI), centered around the current node position. Nodes use a very simple local broadcast protocol to exchange data items.

We analyzed the performance of the system by defining the node coverage, i.e., the percentage of the ROI covered by the items related to *accesible subareas* stored in the buffer of each node.

Terminals held by pedestrians are battery operated and energy consumption is an issue to be dealt with when analyzing the system performance.

We studied the coverage distribution of each class of nodes and the main findings are:

- vehicles always obtain better coverage than pedestrians. Indeed, pedestrians can be further classified into indoor and walking nodes. The coverage distribution for these two subclasses differ (the performance of indoor pedestrians is much worse than walking nodes) but since indoor pedestrians represent about 95% of the total population, the whole class performance is weighted down. The reason for worse performance of indoor pedestrians lies in the fact that buildings represent a more difficult environment for radio propagation.

- a large fraction of transmissions from terminals operated by pedestrians are redundant; because of energy consumption problems we show that dramatic savings can be obtained by slowing down transmission rates at an acceptable cost of coverage loss for of all node classes.

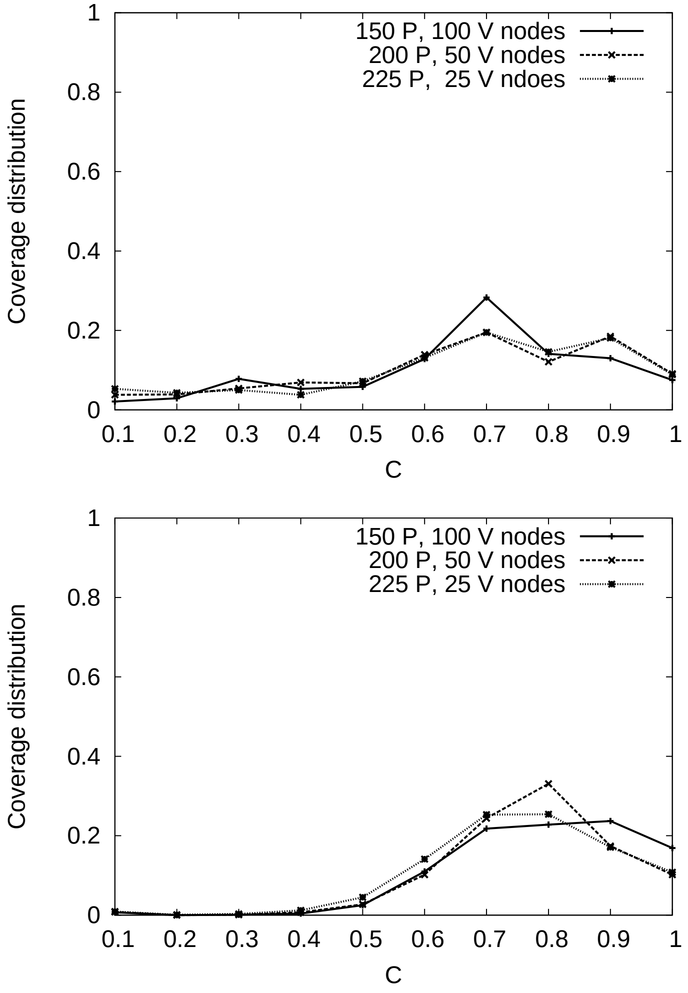


Figure 6.1: Coverage distribution for pedestrians (up) and vehicles (down) in the three population mixes.

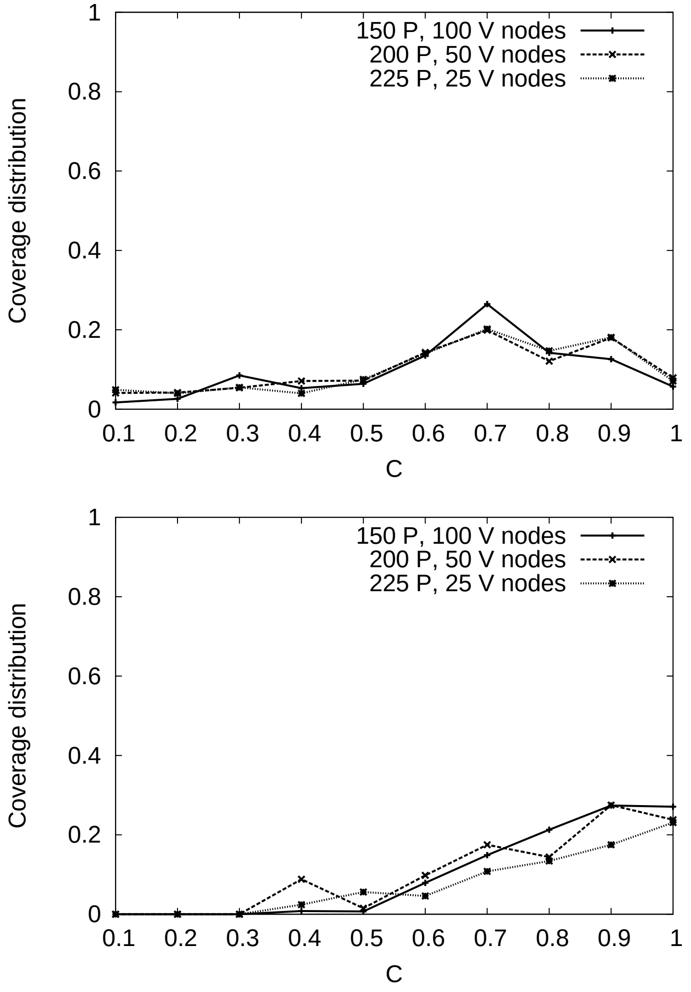


Figure 6.2: Coverage distribution for pedestrians inside buildings (up) and walking outside (down) in the three population mixes.

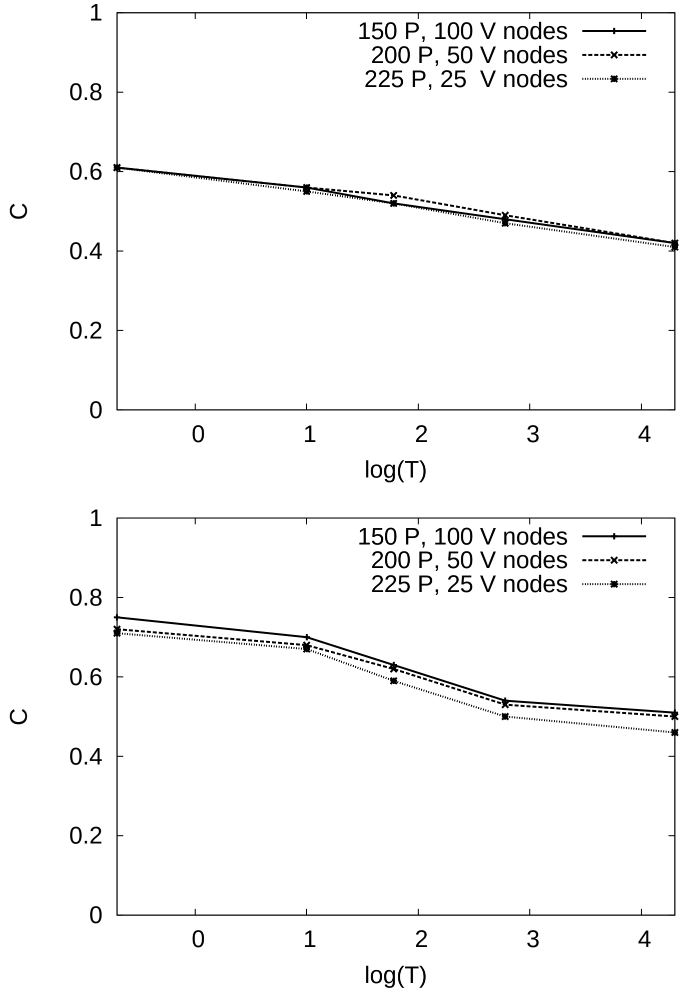


Figure 6.3: Average coverage for pedestrians (up) and vehicles (down) in the three population mixes for increasing values of the transmission delay T.

7

Content Dissemination to Metropolitan Mobile Users

In this chapter we deepen the concept of delaying pedestrian devices transmissions for energy savings. For example, we study the system behaviour when such delay is applied in presence of environmental variability. We also study some additional context-awareness policies to increase coverage.

For the same reason we add some extra relay nodes in elevators obtaining a more efficient spreading of information between different floors inside buildings.

Finally, we carry out some simulations releasing the constraint of one-to-one association between nodes and relative *Regions of Interests*. Likewise, we cease to consider the condition by which each ROI is centered on its own node position. By doing so, we find that sharing among many nodes the same few ROIs is a good strategy to have enough coverage when there is strong dissociation between nodes and ROIs positions.

7.1 Specific system

In this section the assumptions that lead to the characterization of the abstract model used in the following simulations are presented.

We analyze the performance of data exchange protocols among mobile nodes (operated by pedestrians and vehicles) and fixed terminals in a urban area. Each communication node is characterized by its mobility characteristics, radio coverage and amount of memory devoted to the data exchange protocol. For the sake of our analysis a node is able to communicate a message in area covered by its radio using broadcast transmissions. In this paper we consider the following classes of nodes:

- *fixed* (F) nodes, i.e., wireless relay nodes or access points, that are placed at road intersections and building entrances.
- *pedestrians* (P) nodes, that carry portable devices with limited power and memory capacities. Clearly, P nodes move along streets at walking speed, typically concentrating inside buildings.
- *vehicular* (V) nodes that move faster along trajectories constrained by roads; moreover, we assume that this class of nodes has no strict limitations in terms of power and memory since they represent car-installed hardware.

In our experiments we noted that the dissemination of information between different levels of a building can be very critical, due to the limited radio propagation between adjacent floors. To overcome this problem a new class of semi-fixed nodes has been introduced in the system. In particular, we assume to equip some elevators with a wireless relay able to exploit the vertical mobility pattern to disseminate information among different floors of a building. We call these nodes: *elevator nodes* (E) nodes.

The nodes are placed in a three-dimensional space where roads and buildings with multiple levels are modeled as described in Section 7.2. A piece of information is associated with every square tiles of $\delta \times \delta$ meters. In particular we denote as $i(x, y, t)$ the information carried by tile at location (x, y) and time-stamp t. This abstraction is used to model the presence of pieces of information associated to the location.The information $i(x, y, t)$ can be composed by a set of environmental measurements taken autonomously by the mobile nodes or communicated by an infrastructure of active sensors, passive information conveyed by RFID tag, commercial advertisements or traffic data, etc.

The aim of each node is to collect information associated to a given set of tiles that determine its *Region of Interest* (ROI). For the sake of simplicity a ROI is identified with a $\Delta \times \Delta$ square centered around a given target location. ROIs positions may be static or dynamic in both space and time.

Each node aims at maximizing the knowledge of the information associated with its ROI, constrained by the size of the memory (that we denote as B) that it has been allocated to the purpose. To allow complete coverage of the ROI we must assume that $B \geq \left(\frac{\Delta}{\delta}\right)^2$.

7.1.1 Data exchange strategies

Every nodes can complete its task, i.e. to collect the maximum amount of information associated with the ROI, either by directly reading an interesting item $i(x, y, t)$ from the environment, or by overhearing the same piece of information from the broadcast transmissions of other nodes. The former case occurs when a node hits an interesting location (x, y) thanks to its mobility. The latter case is more interesting to model and investigate, being dependent on the mobility of all nodes in the system and the policies they adopt to collect, share, and broadcast the information stored in their memory.

In our analysis we assume that the time is slotted. In each time slot a node acquires the information $i(x, y, t)$ corresponding to its current position (if it is not already contained in B). Moreover, each node can use a simple and random broadcasting policy to propagate its own information to its neighborhood. In particular, a node can randomly selects k information items stored in its local memory B and propagate them. The nodes adopt broadcasting policies aiming at limiting power consumption and at the same time avoiding to clutter the shared radio spectrum. To this end, broadcasting is activated only once every T time slots; moreover, we investigate context aware approaches, e.g., a mobile node starts broadcasting only when it enters a new area, trying to maximize the information diffusion.

Different buffer management strategies have been considered in previous chapters and described in Chapter 3; here we mainly use SD/SI (both selective dropping and selective insertion) that is associated with better performance. When it has a meaning we distinguish memory policy for sensing from that for communication.

The performance index we are interested in is represented by the percentage of ROI covered by the items stored in the local buffer of each node. In the following we will refer to such percentage as *coverage* ($0 \leq C \leq 1$). The value of the coverage clearly varies from node to node and it is time dependent. Furthermore, some subareas of the urban scenario are not accessible to nodes, e.g., due to some physical obstacle, therefore C is computed as the number of items in the local buffer over the total number of *accessible* elements in the node ROI. Finally, if the environmental information changes with time only the most up-to-date elements are used for the computation of the ROI coverage.

7.2 Simulation methodology

In our investigation we evaluate the impact of a realistic 3D environment where pedestrian and cars co-exist in a typical urban scenario; to this end, we adopt the precise mobility simulator UDelModels [27]. UDelModels are a set of tools developed to model urban mesh networks that emulates both the mobility and the radio signal attenuation between any two nodes. UDelModels simulate the mobility of P and V nodes in a 3D space exploiting statistical studies of population and traffic dynamics. In UdelModels the number of nodes in the different classes and the statistical parameters of the mobility model can be properly configured. Pedestrians exhibit a motion that is representative of people in an urban scenario, with different mobility distributions for outside and inside walking, respectively. Moreover, a typical daily human activity cycle is taken into account. Cars mobility patterns take into account speed limits and traffic lights.

The UDelModels propagation [49] simulator is used to estimate the point to point attenuation between each pair of nodes in the 3D space. The channel loss is used to model the radio contacts among the communicating nodes. Any two nodes are assumed to be able to communicate to each other if the channel loss is below a threshold α (in all our experiments α was fixed to -30 dB).

The UDelModels results are used as input to a C++ simulator we developed whose major functionalities are:

- to simulate the radio contacts among the nodes as far as the channel loss is below a threshold α;
- to implement the management of the nodes buffer;
- to simulate broadcast;
- to implement different strategies for energy saving based on transmission delays and context awareness;
- to simulate the E nodes;
- to estimate the system performance in terms of the coverage of the nodes ROI.

7.2.1 Modeling assumption

In our simulation we do not explicitly consider collisions among concurrent transmissions of nearby nodes. We argue that this approximation is acceptable based on the following reasoning:

Let us assume that N transmitters share a wireless channel whose capacity is M b/s. Each transmitter has D bits to transmit in one transmission that may occur at any point in time in an interval of T seconds. Transmission of D bits requires $H = \frac{D}{M}$ seconds to complete. Clearly, the probability of two colliding transmissions is equal to $p_c = \frac{2H}{T}$. A more accurate expression can be obtained by considering border effects that show up if transmissions of either nodes start within the first or the last D seconds in the time interval of length T: in this case we obtain $p_c = \frac{6HT-4H^2}{2T^2}$. We can now derive the probability that N transmitters do not collide as $p_{ok} = (1-p_c)^{\binom{N}{2}}$. In our settings, we assume that the relevant information regarding node coordinates and time stamps (which can be extracted from GPS records) plus the size of the data items requires 30 bytes to be stored. Since we consider transmissions of $k = 5$ data items we obtain $D = 240 \cdot k$ bits to be broadcast once every $T = 0.2$ seconds (see Table 7.1). We observed that the average number of potentially colliding nodes in all our simulations has been $N = 6$. If we assume a $M =$ 10 Mb/s channel we obtain $p_{ok} = 0.973$ which is rather high and confirms that neglecting collision phenomena represents an acceptable approximations.

7.3 Results

In this section we evaluate the performance of content dissemination to metropolitan mobile users under several scenarios. First of all Section 7.3.1 describes the system settings that are common to all experiments. In Section 7.3.2 we evaluate the impact of policies aimed at reducing the overall number of transmissions for energy-saving purposes when information is dynamic, in Section 7.3.3 we analyze how performance is improved for indoor users when elevators are equipped with terminals able to relay data items. Finally, in Section 7.3.4 we analyze coverage for several models of ROI.

7.3.1 System settings

We consider a metropolitan area whose size is 550 m × 500 m, representing 9 blocks of Chicago[1]. The 3D world is partitioned into flat square tiles whose side is $\delta = 25$ m and each tile is associated to a piece of information. The size of the ROI depends on the node class, with $\Delta = 400, 200$ and 100 m for F, V and P nodes, respectively. In practice we assume that V nodes are interested in a larger area with respect to P nodes because of their larger speed. F nodes have the largest ROI since we assume that there are no strict constraints on the memory of the fixed infrastructure.

The memory constraint of each node is equal to the overall number of pieces of information in the node ROI, i.e. $B = \left(\frac{\Delta}{\delta}\right)^2$. Following the results obtained in [12] the memory policy is SI/SD for both information sensing and communication. The duration of one time slot is equal to 0.2 second and each node transmits once every $T = 0.2$ seconds. Each transmission involves the selection of $k = 5$ data items in its buffer.

The fixed wireless infrastructure comprises $N_F = 54$ F nodes placed at road intersections and building entrances; the population we consider comprises $N_P = 200$ P nodes and $N_V = 50$ V nodes. The mobility of V and P nodes is simulated with UDelModels with the speed of V nodes in the range $(25, 67)$ km/h whereas the speed of P nodes is in the range $(2.5, 6.5)$ km/h. The channel loss is used to model the radio contacts among the communicating nodes; any two nodes are assumed to be able to communicate with each other if the channel loss is below a threshold α that in all our experiments was fixed to -30 dB.

The simulations are worked out in a time interval between 7 and 10 a.m. and the initial 15 minutes are always considered as transient and therefore

---

[1] The map data are available at *http://udelmodels.eecis.udel.edu/*

Table 7.1: System settings.

Parameter	Value		
Simulated area	550 m × 500 m		
δ	25 m		
B	$\left(\frac{\Delta}{\delta}\right)^2$		
k	5		
T	0.2 s		
α	-30 dB		
Simulated time	from 7 to 10 a.m		
	F nodes	V nodes	P nodes
Δ (in m)	400	200	100
population	54	50	200
speed (in Km/h)	0	$[25 - 67]$	$[2.5 - 6.5]$
	communication		sensing
Memory policy	SI/SD		SI/SD

excluded from the computation of the average coverage C. All the common system settings are summarized in Table 7.1.

7.3.2 Information dynamics and energy savings

The simplest policy that can be devised for reducing the overall number of transmissions (hence the overall energy consumption) is to increase the time between successive transmissions of all P nodes. To this end we consider increasing delays between transmissions; namely we analyze the average coverage for $T \in \{0.2, 10, 60, 600, \infty\}$ seconds. The value $T = \infty$ represents the extreme case when P nodes never transmit, i.e. they do not participate to the spreading of information; it represents a lower bound on the achievable performance. The analysis is carried out when information does not vary in time (the scenario is called *static*) as well as when one randomly chosen information item is updated once every d seconds.

Figure 7.1 shows the average coverage for different values of d and for increasing delays between successive transmissions of pedestrians. Frequent information updates negatively impact on the average coverage of both pedestrians and vehicles. Pedestrians are able to sustain 1 update every 30 seconds with an acceptable performance reduction. Vehicles are much more resilient to information change and are able to efficiently cope with updates that occur once every 10 seconds. In Table 7.2 we have reported the overall number of transmissions for P nodes as a function of T (first row). It can be noted that for $T = 10$ seconds there is one order of magnitude less transmissions and for $T = 60$ seconds the savings is two orders of magnitude.

We also devised more complex transmission policies. In particular, we leveraged the UDel Models capabilities of tracking the status of a node in the mobility traces; an attribute (called floor) is assigned to each node to classify its current position. The floor attribute for a P node can be set to a code to represent any of the following cases:

- inside, the floor number a node is on;
- outside simulated area;
- walking outside;
- driving in car;
- in parking lot;
- in subway station.

In practical deployments this context information can be obtained by the ever growing number of sensors being embedded in modern mobile phones, e.g., pressure, gyroscope, along with localization services provided by GPS, wifi and mobile networks. This context information can be exploited to improve the information diffusion, taking into account the fact that it is very unlikely to get a radio contact between two terminals separated by a building element, e.g. one node walking outside and one sitting inside a building, or two nodes laying in separate floors of the same building. According to this rationale it turns to be useful to force a node to spread the collected information as soon as its context changes, i.e. it is moving outside, inside or changing the floor, since it is very likely to contact novel nodes (unreachable in its previous context) increasing the probability to exchange fresh data. As a consequence we have devised the following context aware transmission policies:

- floor(w): whenever a P node changes the value of the floor attribute the device wakes up for w seconds and transmits once every $T = 0.2$ seconds;
- walk(w): whenever a P node leaves or enters the walking outside state, e.g. whether it exits/enters a building or stops driving, it wakes up for w seconds and transmits once every $T = 0.2$ seconds;

In the following experiment context awareness has been used in conjunction with previous approach based limiting the transmission opportunity. In Figure 7.2 we show the average coverage for P and V nodes when P nodes are allowed to transmit according to floor(w), with $w = 0.2$ s and $w = 1$ s respectively, and at least once every T s. It can be noted that, as we predict, there is an improvement of the average coverage for both P and V nodes for both static and dynamic information due to a better exploitation of P nodes mobility. In Table 7.2 the overall number of P nodes transmissions (per simulated scenario) is shown along with the additional number of transmissions performed according to the context aware policies; it can be noted that walk(1) requires the lowest number of additional transmissions. On the other hand, Figure 7.2 shows that floor(1) achieves the best coverage.

Finally, Figure 7.3 shows the comparison between floor(1) and walk(1) policies; although results in Table 7.2 clearly prove that the walk(1) policy can achieve a very small amount of transmissions for P nodes the gain in the average value of the coverage is negligible with respect to simply delaying successive transmissions.

Table 7.2: Overall number of transmissions for P nodes and overall additional transmissions for context aware policies.

	T=0.2	T=10	T=60	T=600	T=∞
Reference scenario	7000000	140000	23000	2300	0
Additional transmissions					
floor(1)	6700				
floor(0.2)	1340				
walk(1)	260				

7.3.3 Enhancing the building infrastructure

As we noted in the previous section, context awareness (implemented as additional transmission of P nodes based on the floor(w) and walk(w) policies) proved to be beneficial: there is an improvement of the average coverage for both P and V nodes for both static and dynamic information due to a better exploitation of P nodes mobility. These two policies rely on spontaneous nodes movements and context changing without the need of an extra infrastructure.

Nevertheless, we could improve the coverage of nodes by devising a simple solution based on elevators equipped with terminals able to relay data items. These terminals can be thought of as semi-fixed since their mobility is constrained to a vertical movement. We model the elevator mobility by means of two parameters: the average stopping time at a floor W (60 seconds in our experiments) and the elevator speed (one floor per second). Every time an elevator moves it changes floor by choosing a destination uniformly at random among all floors of the building. Furthermore, their channel loss threshold was fixed to -45 dB.

In Table 7.3 we show the average coverage of P and V nodes for dynamic information ($d = 10$) in the reference scenario. We also show the case of delayed transmissions ($T = 60$) with and without the use of the floor(1) policy. It can be noted that E nodes allow the average coverage of both node

Table 7.3: Average coverage for P and V nodes with elevator based infrastructure.

	no E nodes	W=180	W=120	W=60	W=15	W=5
	P nodes					
delayed	0.40	0.50	0.51	0.55	0.61	0.63
delayed + floor(1)	0.44	0.54	0.55	0.58	0.63	0.66
reference scenario	0.48	0.59	0.60	0.64	0.70	0.73
	V nodes					
delayed	0.53	0.64	0.64	0.66	0.67	0.68
delayed + floor(1)	0.56	0.67	0.67	0.68	0.70	0.71
reference scenario	0.61	0.72	0.72	0.75	0.77	0.78

types to increase in all cases. Indeed, when E nodes hold on at the lower floors of a building are able to collect information carried by V and P nodes passing close to the outer walls. These nodes hold data items that belong to the ROI of P nodes inside the building and vice-versa; in this case, E nodes act as trait d'union among P nodes inside the building and the surrounding external environment. Additionally, E nodes spread this information to all P nodes laying in upper floors as soon as they move up and down. In this case, E nodes act as trait d'union among P nodes laying at different floors of the same building. The key role of E nodes is further highlighted in Table 7.3 where it is shown that short average waiting times at a floor translate into higher average coverage of P and V nodes.

7.3.4 ROI dynamics

In all previous analysis we assume that every node is interested in retrieving the information located around its own position. As a consequence, the ROI has been linked with node position. In the following, this constraint will be removed letting the nodes pick up a random ROI according to several models. Table 7.4 compares the coverage obtained by P and V nodes in several scenarios varying the ROI model, the caching policies and transmission frequency. In particular, the results obtained in the reference scenario with a linked ROI are compared with those worked out when nodes let their ROIs floating according to following model: every 10 minutes a node selects the ROI corresponding to its instantaneous position, then the node moves according to the mobility model without updating the ROI location. The first two sets of experiments in Table 7.4 investigate the effects of the selective insert (SI) policy in the static scenario ($d = \infty$). In particular, we investigate the effect of using SI when reading a piece of information from the environment, i.e. associated to the node position. It can be noted that using SI when the ROI location does not coincide with the node position can be very critical, because it represents an egoistic behavior where data items that are not useful (in terms of ROI) for a node are not acquired and shared in the opportunistic network. This conjecture is confirmed by the simulations, where we notice a slight improvements in terms of coverage for the floating ROI in the case not using SI for environment sensing. As a consequence, SI is not used in all the remaining results of Table 7.4, where we analyze the performance when the information is updated with $d = 10$ s, delayed transmissions ($T = 60s$) and when elevators are used to help spreading the data. The simulation results show that the proposed approaches are effective also for the floating ROIs. In particular, for P nodes we report very limited differences with respect to the case of linked ROIs. This is due to the fact that P nodes mobility within 10 minute is limited, thus the linked and floating ROI models are quite similar in practice. On the other hand, V nodes significantly improve their coverage in the case of floating ROI. Indeed, fixing the ROI for 10 minutes, makes it easier for faster V nodes to collect the items of interest because these are kept constant for a certain period.

Finally, a static ROI model has been simulated, where all nodes randomly select one out of N_{ROI} static ROIs, with a given time frequency. In Table 7.5 the coverage obtained in the limit case of $N_{ROI} = 1$, i.e. same area of

Table 7.4: Coverage of P and V nodes for linked ROI (l-ROI) and floating ROI (f-ROI) for different policies.

	scenario		C for P nodes		C for V nodes	
	N_P	N_V	l-ROI	f-ROI	l-ROI	f-ROI
$d = \infty$, with SI	200	50	0.62	0.62	0.71	0.73
	500	0	0.62	0.62	-	-
$d = \infty$, no SI	200	50	0.62	0.66	0.71	0.81
	500	0	0.62	0.66	-	-
$d = 10$ s, no SI	200	50	0.49	0.47	0.62	0.72
	500	0	0.51	0.48	-	-
$d = \infty$, no SI, delayed	200	50	0.54	0.59	0.62	0.73
	500	0	0.53	0.57	-	-
$d = 10$ s, no SI, delayed	200	50	0.41	0.40	0.53	0.65
	500	0	0.40	0.39	-	-
$d = \infty$, no SI, E nodes	200	50	0.80	0.84	0.86	0.91
	500	0	0.79	0.81	-	-
$d = 10$ s, no SI, E nodes	200	50	0.64	0.60	0.74	0.81
	500	0	0.65	0.61	-	-

Table 7.5: Coverage of P and V nodes with a single static ROI placed outside or inside buildings in the case $N_P = 200$, $N_V = 50$, time 7-10 (reference scenario with $d = \infty$ without and with E nodes).

	C for P nodes		C for V nodes	
	out	in	out	in
reference	0.55	0.70	0.89	0.55
with E nodes	0.65	0.87	0.92	0.86

interest for all nodes, is shown. In particular, we show the sensitiveness of C with respect to the position of the ROI inside and outside the buildings, respectively. It can be noted that P nodes are more effective in collecting information inside buildings, where the density of P nodes is higher. On the contrary, the V nodes coverage improves when the area of interest is located outside.

In presence of ROI dynamics it is very interesting to analyze the temporal behavior of the coverage. To this end, we define the performance index $F_X(\alpha, t)$ that represents the fraction of nodes of class X having $C \geq \alpha$ at time t. In Figure 7.4 we show $F_P(\alpha, t)$, i.e. the percentage of P nodes with $C \geq \alpha = 0.3, 0.5, 0.7, 0.9$ as a function of t. This plots allow one to appreciate the distribution of the values of C of the P nodes. Figure 7.4 refers to the the reference scenario with static information ($d = \infty$). The top-left figure is worked out in the case of a single ROI being updated randomly every 1 hour, synchronously by all nodes. The synchronous update is used as a limit case where all nodes change their interests at the same time and allow us to appreciate the worst case delays in the propagation of the information. As a consequence, in the plot the three transient behaviors due to the synchronous updates turns to be evident. The top-right figure refers to the same synchronous ROI update experiment when increasing the number of ROIs to $N_{ROI} = 4$; it can be observed that nodes coverage reduces when N_{ROI} increases. The positive effects yielded by the exploitation of E nodes is shown in the third figure (bottom-left), where the coverage distribution is highly improved in the case $N_{ROI} = 4$. Finally, in the bottom-right part of Figure 7.4 we show the coverage distribution when all nodes randomly updates their ROIs every hour in an asynchronous fashion. In this latter case, every node picks up a random ROI and keeps it fixed for a random period uniformly distributed between 50 and 70 minutes.

7.4 Summary

In this chapter we describe the results submitted to Elsevier *Special Issue on Opportunistic Networks*.

We study the behaviour of delaying pedestrian devices transmissions in the presence of environmental data variability. We also simulate adding context awareness transmission policies to delaying policies. Thus we obtain performance index increase with low overhead. Furthermore we study how adding some simple infrastructural relay nodes in elevators increases network performance, as radio propagation meets a multitude of obstacles inside buildings.

Finally we show the coverage variations against different ROIs positioning policies or different types of node-ROI association. We find good results only for low degree of node-ROI positions dissociation or for few ROIs shared among many nodes. So we confirm that the cooperation among nodes increases coverage.

The limit to having many nodes sharing few target zones can be accepted if we think that only few regions in a city are of real interest. It is an alternative use of our system in which the starting aim is for each node to gather as much knowledge as possible about its surroundings.

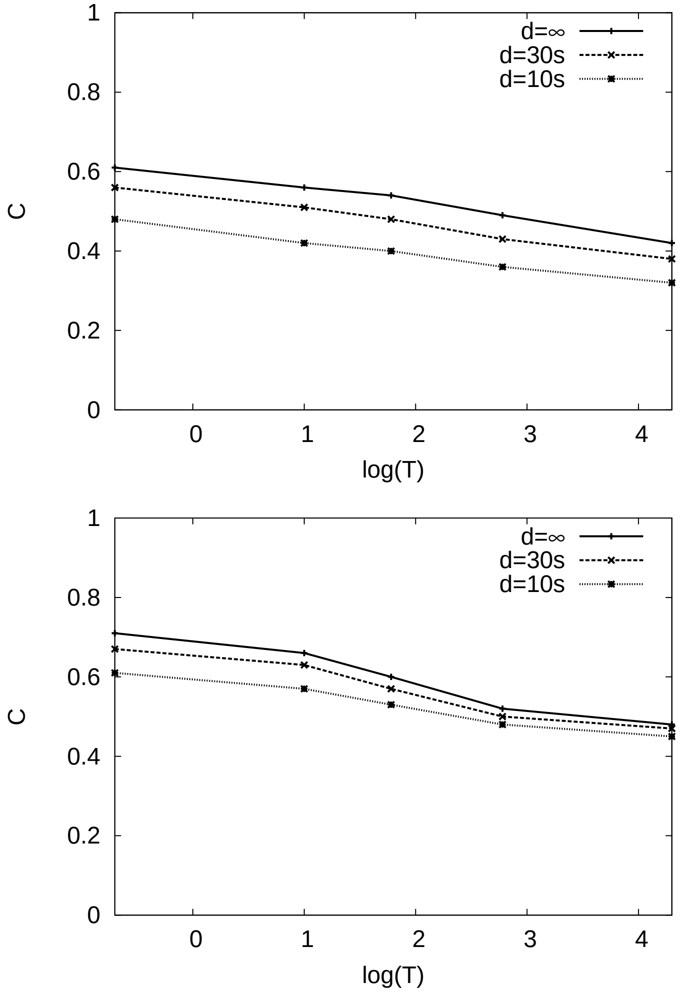


Figure 7.1: Average coverage for P nodes (up) and V nodes (down) for increasing values of the transmission delay T.

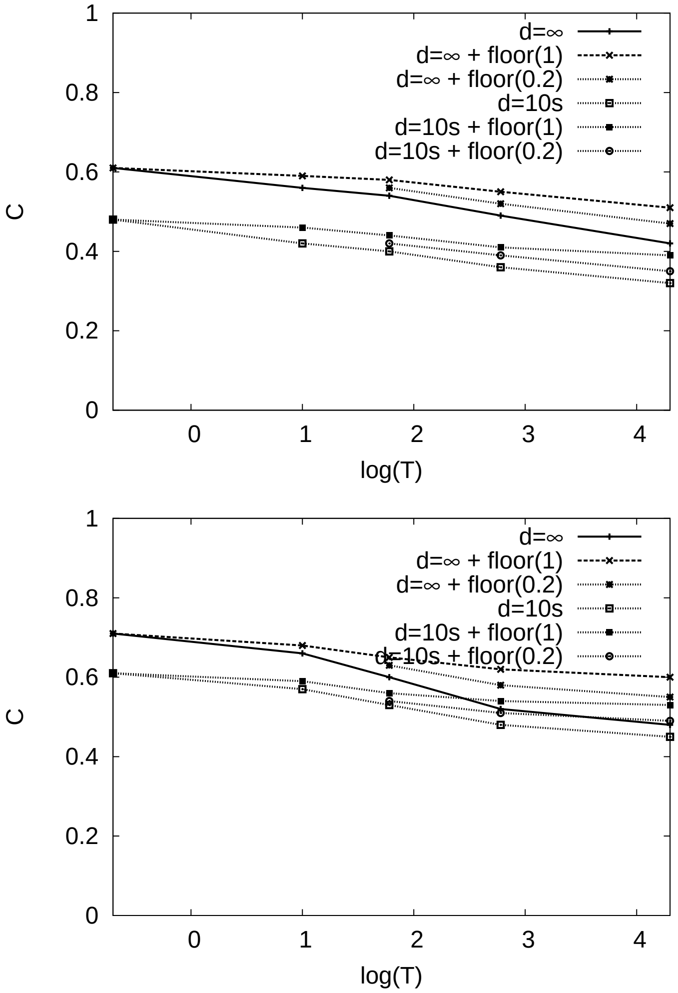


Figure 7.2: Average coverage for P nodes (up) and V nodes (down) for increasing values of the transmission delay T and for policies floor(1) and floor(0.2).

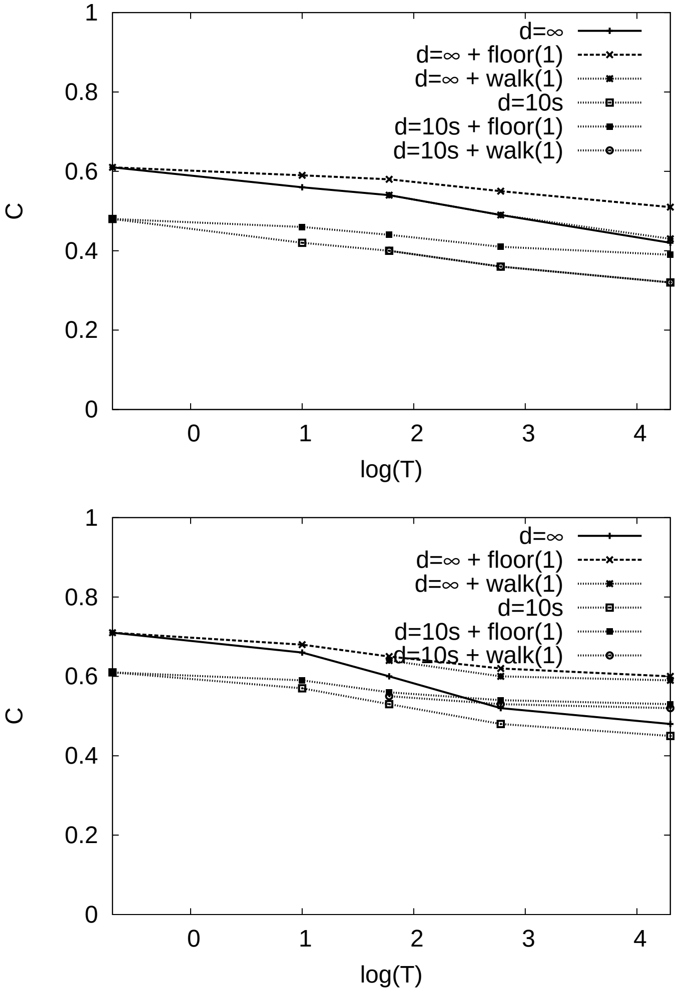


Figure 7.3: Average coverage for P nodes (up) and V nodes (down) for increasing values of the transmission delay T and for policies floor(1) and walk(1).

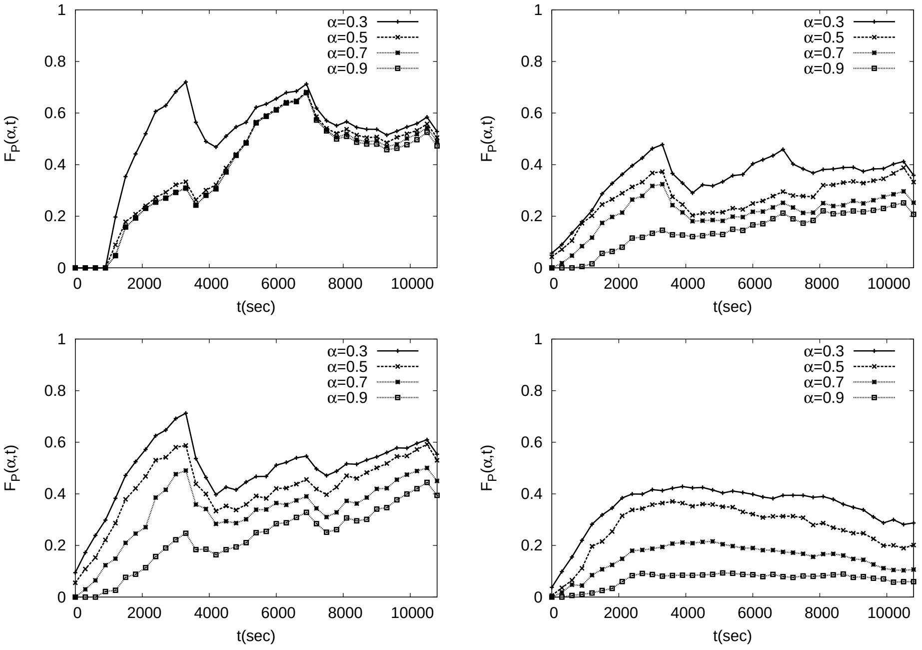


Figure 7.4: Temporal behavior of C for P nodes in the reference scenario with $d = \infty$ when $N_{ROI} = 1$ with 1 hour synchronous update (top left), $N_{ROI} = 4$ with 1 hour synchronous update (top right), $N_{ROI} = 4$ with 1 hour synchronous update and E nodes (bottom left), $N_{ROI} = N$ with 1 hour asynchronous update (bottom right).

8

Conclusions

8.1 Results Summary

In this thesis we consider data gathering in urban scenarios using opportunistic networking. The area under study is discretized in tiles, each defining a data item that nodes can acquire and share through opportunistic transmissions to other nodes in their proximity. The availability of an informational item for each grid cell comes from the Internet of Things paradigm.

Nodes memory is limited and is proportional to the nodes computing resources, i.e., hand-held terminals are assumed to have small buffers while cars and fixed nodes enjoy greater storage capabilities.

The distinctive features of our work lie in the exploitation of detailed mobility and radio propagation traces generated by the UDelModels tools and in the particular abstraction for data gathering: each node aims to retrieve the data items in its ROI centered around the current node position. Also more general ROI positioning policies are considered when some urban target zones are of interest.

Other focus points of our study are hand-held devices energy savings, the impact of eventual wireless infrastructural add-ons and the importance of cooperating among nodes to increase ROIs knowledge. Since data items may change over time all nodes must strive for having access to the latest version.

Furthermore, for mobile terminals ROI is a time varying concept due to the dynamic behavior of pedestrians and cars.

We analyze the performance of the system by defining the node coverage, i.e., the percentage of the ROI *(region of interest)* covered by the items stored in the buffer of each node (the node coverage) for both static and dynamic information. We develope an ad-hoc simulator exploiting the detailed mobility and radio propagation traces generated by the UDelModels tools to estimate this performance index for all node types.

The main findings are:

- simple location-awareness memory management schemes effectively exploit nodes with limited amount of memory;

- generally population density and number of vehicles have a positive impact on system performance;

- increasing randomness of nodes movement by adding a few ideal nodes whose mobility is described by an unconstrained Brownian motion proves to have a beneficial impact on the average coverage of all node types, with the limit that both mobility and propagation of these additional random walk nodes are calculated by our simulator in a simplified way instead than by UDelModels tools;

- the results are valid for a unicast as well as for a more realistic broadcast type of communication, then we show that for our simulations we can realistically assume no radio interference effect;

- vehicles always obtain better coverage than pedestrians, if we refine the definition of node coverage excluding not accessible cells in its calculation. This is mainly because the outside environment is more adapt to radio propagation. Pedestrians can be further classified into indoor and walking nodes, and obviously indoor pedestrians are more penalized;

- a large fraction of transmissions from terminals operated by pedestrians are redundant; because of energy consumption problems, we show that dramatic savings can be obtained by slowing down transmission rates at an acceptable performance cost;
- lowering pedestrian devices is a good strategy for energy savings without strong loss of coverage also in presence of environmental variability;
- adding some context-awareness additional strategy of waking up to delaying policies further improves the results without strong overhead; these strategies consist of waking up lowered pedestrian devices for short instants, when they are in an environment change situation;
- pedestrians are less penalized if we equip elevators with powered relay nodes especially if they are used frequently; these equipped elevators are simulated using both UDelModels tools and our simulator and using some series of internal static nodes vertically placed;
- the performance of the system is not penalized, if we dissociate ROI position from their nodes positions for short but frequent intervals of time ;
- the coverage decreases quickly, if we dissociate fully ROI positions from their nodes, unless we have few ROIs position in common among many nodes.

8.2 Possible Applications Impact

By considering the possible applications, we design an opportunistic networking system that can be useful as a base for distributed applications running in hand-held devices by which every citizen can query informational status about his surroundings. He may also successfully query informational status about far common target zones. These queries could be satisfied with minimal energy consumption and memory occupation. Some of the application areas that can use our communication system and more generally opportunistic communication used as network layer are :

- *Crisis management* to cope with unplanned and unexpected disruptive events, by means of restoring contacts and supporting communication services for first responders, adding more facilities to basic connectivity.
- *Infomobility services and intelligent transportation systems* to improve traffic efficiency and safety. Also applications for tourist information and assistance, parking availability notification, maps and entertainments could benefit from our system.
- *Pervasive healthcare* to cooperate with body-sensor networks to monitor and follow patients' lifestyles and environments. It could also be useful to monitor structures of buildings, risky workplaces and other zones important for ecological and health reasons.

Furthermore, we have shown that the awareness principles adopted by us (memory policies and transmission protocols) can be used successfully to increase performance and effectiveness, as well as leveraging mobility and additional nodes.

By extending our studies we expect that middle-ware could be designed to connect higher level functionality of opportunistic computing services, eventually on the top of virtual social networks, with human centric networks, based on opportunistic networking between small smart objects diffused everywhere.

8.3 Future Evolutions

Our study on opportunistic communication system is just a starting point, which needs to be further refined and extended in order to be successfully applied to network layer design in the above application areas.

There are many possible extensions to the current work. For example, we are considering the impact of coding techniques [37] on the coverage of nodes as well as the use of compressive sensing [13, 5] concepts for sparsely defined information items.

The idea of using network coding can be summarized in the following changes to the transmission protocol adopted.

In the current implementation at each broadcast transmission each device broadcasts k data to its neighbours. In the modified protocol this data diffusion phase could follow an ad-hoc designed light protocol. The latter would exchange meta-information about the data missing from the neighbours' ROIs. A similar light protocol is used in [33] for spreading summaries by leveraging Bloom filters. The available data, indicated in the previous phase, are fused together with network coding techniques and then broadcasted into the neighbourhood. The data transmitted would be chosen to maximize the chance of increasing average Region of Interest knowledge of nodes inside transmission range. In the following phase each node that receives some information fused in a single message could try to decode useful data using previously memorized items. Note that protocol in [33] does not use network coding but only Bloom filters to facilitate the search of missing information, whereas we could try by using both. Obviously some statistics could be calculated to analyse the advantages by using or not such techniques.

In addition, another possibility would be exploring the research area on spatio-temporal databases [6] to import ideas and methods to improve our work. For instances, we could apply them with the aim to increase the node Regions of Interest knowledge, eventually redefining it in a broader way, and to manage the whole knowledge circulating in the system in a useful way.

Moreover, another idea to explore could be deploying some prototypes of our system model in order to collect real cases traces and statistics for an in-depth study and analysis. For example, this could be carried out by managing the contacts among mobile nodes as a complex graph and exploiting new transmission protocols using graph-theory and complex network analysis.

Finally a broader next aim is the application of similar communication protocols and their evolution as a possible important part of Future Internet and Internet of Things, in particular when they are considered as an integral part and a main infrastructure of Smart City models. Another important aspect of Smart City targets is green energy management. In a different context, changing our target aims but not our methodology and core simulation tools, another possible option would be to evaluate energy needs in a urban scenario. These energy needs are associated to house or office usages as well as mobile and fixed devices computing requests sorted to cloud computing centres. Thus a Smart Grid paradigm similar to that depicted in [17] based on Energy Packet Network (EPN) concept, could be evaluated by simulations. Indeed, in the EPN the energy is thought as composed of packets each of one sufficient to satisfy certain needs for a slot of time. The energy distribution network can be seen as a graph in which the edges are weighted with dispersion factors. Local small renewable green energy sources are also represented, in addition to conventional energy power stations. Cloud data centres are represented with their great energy needs, as well as domestic or industrial energy needs. Another important factor in this model is the possible presence of energy storages, capable of storing energy when it is produced economically or in a green way, and supply it when it is strongly needed. For example, when electric cars are linked to the energy network, even though individually they represent only very small energy storages, taken as a whole their effect could become relevant, thus interesting to be evaluated. In [17] and [18] the author analyses a similar scenario with g-network, a mathematical model for probabilistic study of generalized queuing networks, obtaining some results such as the importance of energy storages or of an efficient associated informational communication network in order to have all energy requests satisfied at a reasonable cost. We could apply this mathematical model to urban life scenarios and validate them with UDelModels and our specifically customized simulation tools.

Bibliography

[1] M. Aly, A. Gopalan, J. Zhao, and A.M. Youssef. STDCS: A spatio-temporal data-centric storage scheme for real-time sensornet applications. *SECON*, pages 377–385, June 2008.

[2] Luigi Atzori, Antonio Iera, and Giacomo Morabito. The internet of things: A survey. *Computer Networks*, 54(15):2787 – 2805, 2010.

[3] Chen Avin, Yuval Lando, and Zvi Lotker. Simple random walks on radio networks (simple random walks on hyper-graphs). *Arxiv preprint arXiv:0907.1678*, 2009.

[4] Richard Becker, Ramón Cáceres, Karrie Hanson, Sibren Isaacman, Ji Meng Loh, Margaret Martonosi, James Rowland, Simon Urbanek, Alexander Varshavsky, and Chris Volinsky. Human mobility characterization from cellular network data. *Commun. ACM*, 56(1):74–82, 2013.

[5] E.J. Candes and T. Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Transactions on Information Theory*, 52(12):5406–5425, 2006.

[6] Yong-Jin Choi and Chin-Wan Chung. Selectivity estimation for spatio-temporal queries to moving objects. In *Proceedings of the ACM international conference on Management of data*, SIGMOD, pages 440–451, 2002.

[7] R.R. Choudhury. Brownian gossip: exploiting node mobility to diffuse information in ad hoc networks. In *IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing*, 2005.

[8] D. Ciullo, V. Martina, M. Garetto, and E. Leonardi. Impact of correlated mobility on delay-throughput performance in mobile ad-hoc networks. In *IEEE INFOCOM*, pages 1–9, 2010.

[9] M. Conti and M. Kumar. Opportunities in opportunistic computing. *Computer*, 43(1):42–50, 2010.

[10] Marco Conti, Sajal K. Das, Chatschik Bisdikian, Mohan Kumar, Lionel M. Ni, Andrea Passarella, George Roussos, Gerhard Trster, Gene Tsudik, and Franco Zambonelli. Looking ahead in pervasive computing: Challenges and opportunities in the era of cyberphysical convergence. *Pervasive and Mobile Computing*, 8(1):2 – 21, 2012.

[11] G. Corrente, R. Gaeta, M. Grangetto, and M. Sereno. Collaborative data retrieval using opportunistic networking. In *Wireless On-demand Network Systems and Services (WONS), 2012 9th Annual Conference on*, pages 139–142, 2012.

[12] Giuseppe Corrente, Rossano Gaeta, Marco Grangetto, and Matteo Sereno. Local data gathering using opportunistic networking in a urban scenario. In *Proceedings of the 8th ACM Symposium on Performance evaluation of wireless ad hoc, sensor, and ubiquitous networks*, PE-WASUN '11, pages 57–64. ACM, 2011.

[13] D.L. Donoho. Compressed sensing. *Information Theory, IEEE Transactions on*, 52(4):1289–1306, 2006.

[14] M. Fiore, F. Mininni, C. Casetti, and C.F. Chiasserini. To cache or not to cache? In *IEEE INFOCOM*, pages 235–243, 2009.

[15] Sabrina Gaito, Elena Pagani, and Gian Paolo Rossi. Strangers help friends to communicate in opportunistic networks. *Computer Networks*, 55(2):374 – 385, 2011.

[16] M. Garetto, P. Giaccone, and E. Leonardi. On the capacity of ad hoc wireless networks under general node mobility. In *INFOCOM . 26th IEEE International Conference on Computer Communications*, pages 357–365, 2007.

[17] E. Gelenbe. Energy packet networks: adaptive energy management for the cloud. In *Proceedings of the 2nd International Workshop on Cloud Computing Platforms*, page 1. ACM, 2012.

[18] Erol Gelenbe. Energy packet networks: smart electricity storage to meet surges in demand. In *Proceedings of the 5th International ICST Conference on Simulation Tools and Techniques*, pages 1–7. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2012.

[19] Marta C. González, Cesar A. Hidalgo, and Albert-Laszlo Barabási. Understanding individual human mobility patterns. *Nature 2008*, 453(7196):779–782, 2008.

[20] Jos Hernndez-Muoz, Jess Vercher, Luis Muoz, Jos Galache, Mirko Presser, Luis Hernndez Gmez, and Jan Pettersson. Smart cities at the forefront of the future internet. In *The Future Internet*, volume 6656 of *Lecture Notes in Computer Science*, pages 447–462. Springer Berlin / Heidelberg, 2011.

[21] Ting-Kai Huang, Chia-Keng Lee, and Ling-Jyh Chen. Prophet+: An adaptive prophet-based routing protocol for opportunistic network. In *Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference on*, pages 112 –119, 2010.

[22] Pan Hui, Jon Crowcroft, and Eiko Yoneki. BUBBLE rap: Social-based forwarding in delay tolerant networks. *IEEE Transactions on Mobile Computing*, 99(PrePrints), 2010.

[23] P. Jacquet, Bernard Mans, and Georgios Rodolakis. Information propagation speed in mobile and delay tolerant networks. *Information Theory, IEEE Transactions on*, 56(10):5001–5015, 2010.

[24] Sushant Jain, Rahul C. Shah, Waylon Brunette, Gaetano Borriello, and Sumit Roy. Exploiting mobility for energy efficient data collection in wireless sensor networks. *Mobile Network and Applications*, 11(3):327–339, 2006.

[25] Philo Juang, Hidekazu Oki, Yong Wang, Margaret Martonosi, Li Shiuan Peh, and Daniel Rubenstein. Energy-efficient computing for wildlife tracking: design tradeoffs and early experiences with zebranet. *SIGOPS Oper. Syst. Rev.*, 36(5):96–107, 2002.

[26] A. Keranen, M. Pitkanen, M. Vuori, and J. Ott. Effect of non-cooperative nodes in mobile dtns. In *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2011 IEEE International Symposium on a*, pages 1 –7, 2011.

[27] Jonghiun Kim, Vinay Sridhara, and Stephan Bohacek. Realistic mobility simulation of urban mesh network. *Ad Hoc Networks*, 7(2):411–430, 2009.

[28] Amir Krifa, Chadi Barakat, and Thrasyvoulos Spyropoulos. Optimal buffer management policies for delay tolerant networks. In *IEEE SECON*, pages 16–20, 2008.

[29] Fabian Kuhn and Rotem Oshman. Dynamic networks: models and algorithms. *ACM SIGACT News*, 42(1):82–96, 2011.

[30] Kyunghan Lee, Seongik Hong, Seong Joon Kim, Injong Rhee, and Song Chong. Slaw: A new mobility model for human walks. In *INFOCOM 2009, IEEE*, pages 855 –863, 2009.

[31] Suk-Bok Lee, Starsky H. Y. Wong, Kang-Won Lee, and Songwu Lu. Content management in a mobile ad hoc network: Beyond opportunistic strategy. In *IEEE INFOCOM mini-conf*, 2011.

[32] Uichin Lee, E. Magistretti, M. Gerla, P. Bellavista, and A. Corradi. Dissemination and harvesting of urban data using vehicular sensing platforms. *Vehicular Technology, IEEE Transactions on*, 58(2):882 –901, 2009.

[33] Uichin Lee, Biao Zhou, M. Gerla, E. Magistretti, P. Bellavista, and A. Corradi. Mobeyes: smart mobs for urban monitoring with a vehicular sensor network. *Wireless Communications, IEEE*, 13(5):52 –57, 2006.

[34] J. Leguay, A. Lindgren, J. Scott, T. Friedman, and J. Crowcroft. Opportunistic content distribution in an urban setting. In *Proceedings of the 2006 SIGCOMM workshop on Challenged networks*, pages 205–212, 2006.

[35] Anders Lindgren, Avri Doria, and Olov Schelen. Probabilistic routing in intermittently connected networks. In *SIGMOBILE Mobile Computing and Communication Review*, 2004.

[36] Chenyang Lu, Guoliang Xing, Octav Chipara, Chien liang Fok, and Sangeeta Bhattacharya. A spatiotemporal query service for mobile users in sensor networks. In *IEEE International Conference on Distributed Computing Systems*, pages 381–390, 2005.

[37] M.G. Luby, M. Mitzenmacher, M.A. Shokrollahi, and D.A. Spielman. Efficient erasure correcting codes. *Information Theory, IEEE Transactions on*, 47(2):569–584, 2001.

[38] A. Mei and J. Stefa. Swim: A simple model to generate small mobile worlds. In *INFOCOM 2009, IEEE*, pages 2106–2113, 2009.

[39] H. Morcos, G. Atia, A. Bestavros, and I. Matta. An information-theoretic framework for field monitoring using autonomously mobile sensors. *Ad Hoc Networks*, 9(6):1049 – 1058, 2011.

[40] H. Morcos, A. Bestavros, and I. Matta. Amorphous placement and informed diffusion for timely field monitoring by autonomous, resource-constrained, mobile sensors. In *5th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, pages 469–477, 2008.

[41] A. Pettarin, A. Pietracaprina, G. Pucci, and Eli Upfal. Tight bounds on information dissemination in sparse mobile networks. In *30th Annual ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing*, 2011.

[42] Anna Kaisa Pietilinen and Christophe Diot. Dissemination in opportunistic social networks: the role of temporal communities. In *MobiHoc*, pages 165–174. ACM, 2012.

[43] Giovanni Resta and Paolo Santi. The effects of node cooperation level on routing performance in delay tolerant networks. In *IEEE SECON*, pages 413–421, 2009.

[44] I. Rhee, M. Shin, S. Hong, K. Lee, and S. Chong. On the levy-walk nature of human mobility. In *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, pages 924–932. IEEE, 2008.

[45] N. Ristanovic, G. Theodorakopoulos, and J. Le Boudec. Traps and pitfalls of using contact traces in performance studies of opportunistic networks. In *INFOCOM, 2012 Proceedings IEEE*, pages 1377 –1385, 2012.

[46] A.D. Sarwate and A.G. Dimakis. The impact of mobility on gossip algorithms. In *IEEE INFOCOM*, pages 2088–2096, 2009.

[47] Hans Schaffers, Nicos Komninos, Marc Pallot, Brigitte Trousse, Michael Nilsson, and Alvaro Oliveira. Smart cities and the future internet: Towards cooperation frameworks for open innovation. In *The Future Inter-*

net, volume 6656 of *Lecture Notes in Computer Science*, pages 431–446. Springer Berlin / Heidelberg, 2011.

[48] Tara Small and Zygmunt J. Haas. The shared wireless infostation model: a new ad hoc networking paradigm (or where there is a whale, there is a way). In *Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, MobiHoc '03, pages 233–244. ACM, 2003.

[49] Vinay Sridhara and Stephan Bohacek. Realistic propagation simulation of urban mesh networks. *Computer Networks: The International Journal of Computer and Telecommunications Networking*, 51(12):3392–3412, 2007.

[50] Yu Wang and Hongyi Wu. Delay/fault-tolerant mobile sensor network (dft-msn): A new paradigm for pervasive information gathering. *Mobile Computing, IEEE Transactions on*, 6(9):1021 –1034, 2007.

[51] Mark Weiser. The computer for the 21st century. *SIGMOBILE Mob. Comput. Commun. Rev.*, 3(3):3–11, July 1999.

[52] Mark Weiser and John Seely Brown. The coming age of calm technology, 1996.

[53] Gjergji Zyba, Stratis Ioannidis, Christophe Diot, and Geoffrey M. Voelke. Dissemination in opportunistic mobile ad-hoc networks: The power of the crowd. Technical report, Technicolor, Tech. Rep. CR-PRL-2010-07-0001, 2010.