

N° d'ordre : 04/2022 - D/INF.

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique
Université des Sciences et de la Technologie Houari Boumediène
Faculté d'Informatique



THÈSE DE DOCTORAT EN SCIENCES

Présentée pour l'obtention du grade de Docteur

En : INFORMATIQUE
Spécialité : Programmation et Systèmes

par

Fairouz CHEHBOUR

Sujet

***La Collecte de Données dans les Réseaux de Capteurs
Vehiculaires***

Soutenue publiquement le **12/05/2022**, devant le Jury composé de :

M. A. AMROUCHE	Professeur	à l'USTHB	Président
Mme. S. MOUSSAOUI	Professeur	à l'USTHB	Directeur de thèse
Mme. H. LABIOD	Professeur	à Huawei-Paris	Examineur
M. S. BITAM	Professeur	à l'U. Biskra	Examineur
M. M. BACHIR	Professeur	à l'ENSIA	Examineur
M. F. BOUYAKOUB	Professeur	à l'USTHB	Examineur
Mme. Z. DOUKHA	Maitre de conf./B	à l'USTHB	Invitée

DEMOCRATIC REPUBLIC OF ALGERIA
Ministry of Higher Education and Scientific Research
University of Science and Technology Houari Boumediene
Faculty of Computer Science



THESIS OF DOCTORATE IN SCIENCES

Presented to obtain the

DOCTORAL DEGREE

In : COMPUTER SCIENCE

Speciality : Programming and Systems

by

Fairouz CHEHBOUR

Subject

Data Collection in Vehicular Sensor Networks

Publicly defended on 05/12/2022, in front of the Jury:

Mrs. A. AMROUCHE	Professor	à l'USTHB	President
Mme. S. MOUSSAOUI	Professor	à l'USTHB	Thesis Supervisor
Mme. H. LABIOD	Professor	à Huawei-Paris	Reviewer
M. S. BITAM	Professor	à l'U. Biskra	Reviewer
M. M. BACHIR	Professor	à l'ENSIA	Reviewer
M. F. BOUYAKOUB	Professor	à l'USTHB	Reviewer
Mme. Z. DOUKHA	Lecturer	à l'USTHB	Guest

Résumé

Les progrès des technologies de l'information et de la communication (TIC) conjugués à la prolifération des dispositifs de détection sans fil et des objets IoT (Internet of Things) ouvrent progressivement la voie à l'émergence des domaines d'application variés, notamment, les villes intelligentes. Ainsi, un grand nombre d'objets IoT ont été déployés le long des routes des villes pour une pléthore d'applications. Les exemples incluent: l'éclairage intelligent, le stationnement intelligent, les notifications d'obstruction des égouts, etc. L'un des principaux défis dans la réalisation de ces applications est la conception des mécanismes efficaces permettant la collecte et la livraison des données des objets IoT aux centres de données pour traitement et/ou stockage.

L'approche de collecte de données basée sur les réseaux de véhicules est à la fois rentable et économe en énergie car elle repose sur une infrastructure préétablie. En effet, les véhicules peuvent incidemment traverser la zone de couverture des capteurs actifs, et peuvent donc échanger des données. Cependant, cette approche offre une latence élevée, qui peut atteindre 24 heures, en raison des restrictions de la mobilité urbaine: tracés des routes, conditions de circulation, situations de congestion, et le manque de contrôle de la trajectoire des véhicules en raison de leur obéissance à la volonté de leurs conducteurs. Cela limite cette approche à la classe d'applications tolérantes aux retards. Étant donné que dans certains scénarios de villes intelligentes les applications doivent répondre dans un délai relativement court, nous jugeons primordial de réduire le délais de collecte induit par cette approche. Par exemple, un égout pluvial bouché doit être signalé dans un délai court; sinon, il pourrait être trop tard pour réagir. Un ATMS (Advanced Traffic Management System) nécessite des mises à jour fréquentes des données des capteurs, pour annoncer aux conducteurs des informations utiles sur: les conditions de circulation, la pollution de l'air, et le bruit environnemental. A cet égard, nous nous sommes intéressés à la problématique de réduction de la latence du processus de collecte de données à base de véhicules dans les environnements urbains.

En fait, en analysant les travaux précédents, nous avons constaté qu'en plus des contraintes inévitables liées à la mobilité urbaine, la longueur des trajectoires assignées aux véhicules agissants en tant que mules de données ou puits mobiles, contribue de manière significative à l'augmentation de la latence du processus de collecte. En effet, dans ces travaux, des trajectoires de plusieurs segments de route sont attribuées à un

seul élément mobile. Ainsi, les données collectées par ce dernier ne seront transmises au centre de données qu'après avoir parcouru tous les tronçons de route de sa trajectoire, ce qui ne fait que ralentir le processus de collecte. Alors que les véhicules sont abondants en milieu urbain et peuvent communiquer entre eux et/ou avec des éléments de l'infrastructure, nous estimons que les capacités de mobilité et de connectivité étaient souvent sous-exploitées dans ces travaux, ce qui s'est traduit par des solutions sous-optimales entraînant une latence importante.

Pour remédier à ce problème, nous proposons dans cette thèse de tirer parti d'une caractéristique importante de l'environnement urbain, à savoir sa richesse en éléments mobiles (véhicules). Nous suggérons d'exploiter la présence opportuniste des véhicules par l'implication du plus grand nombre de ces derniers dans le processus de collecte. Cela permettra d'alléger les tâches des éléments mobiles en leur attribuant des trajectoires plus courtes, ainsi que d'augmenter le degré de la répartition du processus de collecte, par conséquent réduire sa latence. La réduction de la latence d'une telle approche, qui permet d'une part d'économiser les coûts de déploiement et l'énergie et de l'autre part, d'assurer une plus grande ouverture aux applications qui ne tolèrent pas les retards, ce qui s'aligne avec le concept de la ville intelligente.

Dans cette thèse, nous examinons d'abord l'approche de collecte de données basée sur les réseaux de véhicules à travers une étude détaillée des solutions proposées dans la littérature. Ensuite, nous proposons une taxonomie des solutions de collecte de données urbaines basées sur cette approche. Enfin, nous proposons deux principales contributions. Il s'agit de deux stratégies de collecte de données urbaines rentables et économiques.

La première stratégie, appelée OMS (Opportunistic Mules for Short Delay Applications in Smart City), offre une latence de collecte de données réduite en tirant parti des communications véhiculaires opportunistes pour augmenter la distribution du processus de collecte de données. Précisément, il allège la tâche de collecte de données en attribuant un seul segment de route aux véhicules désignés comme mules de données. Ceci est réalisé grâce à la conception de deux processus qui opèrent simultanément : i) la dissémination d'une demande de mules de données et ii) l'élection opportuniste des mules de données.

La seconde stratégie traite le problème de la collecte de données urbaines avec des puits mobiles en situations de congestion. Dans ce système, nous proposons de tirer profit de la forte densité de véhicules dans les situations de congestion pour soutenir les puits mobiles avec des puits temporaires, désignés à partir des véhicules de citoyens, afin de collecter les données des routes encombrées. Enfin, des simulations dans des conditions de mobilité réalistes démontrent que les deux mécanismes de collecte proposés parviennent à fournir de façon opportuniste à une grande partie de la zone d'intérêt, des véhicules pour collaborer au processus de collecte de données en un temps réduit.

ملخص

إن التقدم في تقنيات المعلومات والاتصالات الى جانب انتشار أجهزة الاستشعار اللاسلكية والأشياء الإنترنتية تمهد الطريق تدريجياً لانتشار المدن الذكية. في الواقع، قد تم نشر عدد كبير من الأشياء الإنترنتية على طول طرق المدن من اجل توفير مجموعة متنوعة من التطبيقات من بينها: الإضاءة الذكية ومواقف السيارات الذكية والإخطار بعرقلة تصريف مياه العواصف وما إلى ذلك. يتمثل أحد التحديات الحاسمة لنجاح تطبيقات المدينة الذكية في توفير آليات فعالة لتسليم البيانات من الأشياء الإنترنتية إلى مراكز البيانات لمعالجتها.

من ناحية أخرى ، تلعب الشبكات المؤسسية الخاصة بالمرکبات دورًا مهمًا في تصميم تطبيقات المدن الذكية. في الواقع، قد تم استخدام المركبات في العديد من الابحاث العلمية كبغال بيانات أو أحواض متنقلة لجمع البيانات من مستشعرات الأشياء الإنترنتية وتسليمها إلى السحابة أو إلى مراكز البيانات. من الممكن ان تعبر المركبات عرضياً منطقة تغطية أجهزة الاستشعار النشطة ومن ثم يتم تبادل البيانات بطريقة اتهازية.

يوفر هذا النهج حلاً بديلاً قليل التكلفة لجمع البيانات، حيث إنه يستفيد من البنية التحتية لشبكات المركبات المنتشرة مسبقاً. كما إنه موفر للطاقة لان تفريغ بيانات أجهزة الاستشعار يتم باستخدام اتصال قفزة واحدة فقط. غير ان صعوبة التحكم في حركة المركبات المنقادة لإرادة السائقين ، إضافة إلى الديناميكية العالية للبيئة الحضرية و قيود التنقل الحضري مثل قوانين المرور، وحالات الازدحام، تؤدي حتما إلى تأخيرات عالية في جمع البيانات القائمة على هذا النهج.

إن ارتفاع مدة الاستجابة لنهج جمع البيانات القائم على السيارات يجعله مقصراً على فئة واحدة فقط من التطبيقات ،الا وهي: التطبيقات المتحملة للتأخير بالنظر إلى أنه في بعض سيناريوهات المدن الذكية، يتعين على التطبيقات الاستجابة في وقت قصير نسبياً، فإن تقليل زمن وصول البيانات استناداً على مثل هذا النهج الواعد لتلبية متطلبات تطبيقات التأخير القصير أمر

بالغ الأهمية. على سبيل المثال ، يجب الإبلاغ عن انسداد تصريف العواصف في غضون فترة تأخير محدودة ؛ وإلا ، فقد يكون الوقت قد فات. في هذا الصدد ، نركز اهتمامنا في هذه الأطروحة على مشكلة تقليل زمن عملية جمع البيانات القائمة على المركبات في المناطق الحضرية. من خلال تحليل الأعمال السابقة ، وجدنا أنه بالإضافة إلى القيود المفروضة على التنقل الحضري، فإن طول المسارات المخصصة للمركبات التي تعمل كبغال البيانات أو كأحواض متنقلة، يساهم بشكل كبير في زيادة مدة عملية جمع البيانات. في الواقع ، في هذه الأعمال ، يتم تخصيص مسارات مكونة من عدة أجزاء لمركبة واحدة. وبالتالي، فإنه لن يتم إرسال البيانات التي تم جمعها بواسطة بغل البيانات إلى مركز البيانات إلا بعد تغطية جميع أجزاء المسار المخصص له ، مما يؤدي إلى إبطاء عملية التجميع. في حين أن المركبات متوفرة بكثرة في البيئة الحضرية ويمكنها التواصل مع بعضها البعض ومع عناصر البنية التحتية، نجد أن إمكانيات التنقل والاتصال لشبكات المركبات لم تُستغل بشكل كافٍ في هذه الأعمال، مما أدى إلى حلول دون المستوى الأمثل بزمن انتقال كبير للبيانات

للتغلب على مشكلة ارتفاع زمن الاستجابة لعملية جمع البيانات، نقترح في هذه الأطروحة الاستفادة من خاصية مهمة للبيئة الحضرية، وهي ثرائها بالعناصر المتنقلة (المركبات). نقترح الاستفادة من الوجود المتوازي للمركبات من خلال إشراك أكبر عدد ممكن من المركبات في عملية الجمع. كما أننا نرى بأن زيادة درجة التوزيع لعملية الجمع ستقلل من مدتها وبالتالي سيسمح بمزيد من الانفتاح على تطبيقات التأخيرات القصيرة والتي تتماشى مع مفهوم المدينة الذكية.

في هذه الأطروحة، نقوم أولاً بدراسة نهج جمع البيانات المستندة إلى شبكات المركبات واقتراح تصنيف للحلول القائمة على النهج. تُوجت هذه الدراسة باقتراح مخططين قليلي التكلفة من حيث متطلبات البنية التحتية و استهلاك الطاقة لجمع البيانات الحضرية.

المخطط الأول ، و المسمى: «البغال المتوازية لتطبيقات التأخير القصير في المدينة الذكية»، يوفر زمن انتقال منخفض لجمع البيانات من خلال الاستفادة من اتصالات المركبات المتوازية لتوزيع عملية جمع البيانات بشكل أكبر. على وجه التحديد ، فإنه يسهل مهمة بغل البيانات من خلال تعيينهم فقط لجمع البيانات في جزء واحد من الطريق. يتم تحقيق ذلك من خلال عمليتين متزامنتين: (١) توزيع طلب بغال البيانات و (٢) اختيار بغال البيانات المتوازية

أما المخطط الثاني فيتناول جمع البيانات الحضرية في حالات الازدحام باستخدام الأحواض المتنقلة. في هذا المخطط، نقترح الاستفادة من الكثافة العالية للمركبات في حالات الازدحام لدعم الأحواض المتنقلة بأحواض مؤقتة مخصصة من سيارات المواطنين، لجمع البيانات من الطرق المزدحمة. أخيراً ، تُظهر نتائج المحاكاة باستعمال معايير التنقل الواقعية أن المخططات المقترحة تمكنت من

تزويد جزء كبير من منطقة الاهتمام بالمركبات الانتهازية للتعاون في عملية جمع البيانات في فترة قصيرة نسبياً.

Abstract

Advances in information and communication technologies combined to the proliferation of wireless sensing devices and IoT (Internet of Things) objects are gradually paving path for the emergence of smart cities. Consequently, plethora of IoT objects have been deployed along cities' roads for wide variety of applications. Examples include smart lighting, smart parking, storm drain obstruction notification, etc. A critical challenge for the success of a smart city application is to provide efficient mechanisms for data delivery from IoT objects to data centers for processing.

On the other hand, Vehicular Adhoc Networks are playing an important role in the design of smart city applications. Actually, in many works, vehicles are suggested to be used as data mules or mobile sinks to gather data from IoT sensors and to deliver it to the cloud or to data centers. Indeed, vehicles may incidentally cross the coverage area of active sensors. Therefore, they can exchange data. The benefits of this approach are multiple. First, it provides an alternative cost effective solution for data collection, as it leverages an already deployed cost-free infrastructure. Second, it is energy-efficient as data offloading of IoT sensors is performed using only one hop communication. However, it is difficult to control the movement of vehicles, since the latter only move according to the the drivers' will and according to trajectories that are not necessarily adapted to data transport. Moreover, the high dynamism of urban environment, in addition to the urban mobility constraints: road layouts, traffic conditions, congestion situations, lead to high delays for vehicular networks-based data gathering.

The high latency of the vehicle-based approach limits it to only one class of applications, namely: delay-tolerant applications. Given that, in some smart city scenarios, applications are required to respond in a relatively short delay, reducing the latency of such a promising approach to meet the requirements of short delay applications is paramount. For instance, a clogged storm drain should be reported within a limited delay; otherwise, it might be too late to react. An ATMS (Advanced Traffic Management System) requires frequent updates of sensor data to advertise useful information about traffic conditions, air pollution, and environmental noise to drivers. In this respect, we are interested in the problem of reducing the latency of the vehicles'-based data collection process in urban areas.

In fact, by analyzing previous works, we found that, in addition to the unavoidable

constraints related to urban mobility, the length of the trajectories assigned to vehicles acting as data mules or mobile sinks, significantly contributes to the increase in latency of the collection process. Indeed, in these works, trajectories of several road segments are assigned to a single vehicle. Thus, the data collected by the latter will be transmitted to the data center only after having covered all the road segments of the assigned trajectory, which slows down the collection process. While vehicles are abundant in the urban environment and can communicate with each other and with infrastructure elements, we find that mobility and connectivity capabilities were underutilized in these works, resulting in sub-optimal solutions with significant latency.

To overcome the high latency issue, we propose in this thesis to take benefit from an important characteristic of the urban environment, namely its richness in mobile elements (vehicles). We suggest leveraging the opportunistic presence of vehicles by involving as many vehicles as possible in the collection process. We argue that increasing the distribution degree of the collection process will reduce its latency. Reducing the latency of such a promising approach, will allow greater openness to short delay applications which is in line with the smart city concept.

In this thesis, we first, investigate the vehicular networks-based data collection approach and propose a taxonomy for vehicles-based urban data collection solutions. This investigation culminated in proposing two cost-effective and energy-efficient urban data collection schemes.

The first scheme, called OMS (Opportunistic Mules for Short Delay Applications in Smart City) provides a reduced data collection latency by leveraging opportunistic vehicular communications to increase the distribution of the data collection process. Precisely, it alleviates the data collection task by charging the vehicles that are designated as data mules, to collect data only on one road segment. This is achieved through two simultaneous processes of: i) a request for data mules propagation and ii) an opportunistic data mules election.

The second scheme deals with mobile sinks-based urban data collection in congestion situations. In this scheme, we propose to take benefit from the high vehicular density of congestion situations to support mobile sinks with temporary sinks, designated from citizen cars, to collect data from congested roads.

Finally, Simulations in realistic mobility settings demonstrate that the proposed schemes manage to provide a large portion of the area of interest with opportunistic vehicles to collaborate in the data collection process in a relatively short delay.

Acknowledgement

First and foremost, praise and thanks to ALLAH, the Almighty, for his blessings throughout my research work to bring it to fruition.

I have had the pleasure and honor of receiving the support of many wonderful people to whom I would like to express my deepest gratitude.

Foremost, I would like to express my sincere gratitude to my research supervisor, Prof. Samira MOUSSAOUI for her support, guidance, and encouragement throughout these years. During these years, Prof. MOUSSAOUI has not only revealed her scientific qualities but also her human qualities. She is willing to share her knowledge and career experience and give emotional and moral encouragement.

A very special thanks to Dr. Zouina DOUKHA, for her valuable help, for reviewing my papers and my thesis. I am extremely grateful for what she has offered me. I would also like to thank her for her friendship and empathy.

I am very much honored by and grateful to Prof. Abderrahmane AMROUCHE , Prof. Houda LABIOD , Prof. Salim BITAM , Prof. Malik BACHIR , and Prof. BOUYAKOUB Fayçal, my committee members, for their kind willingness to judge this work.

I am extremely grateful to my family and my husband for their tremendous love, understanding, support, and encouragement throughout all my thesis. These have been the driving force behind my accomplishments. Special thanks to my dear mother, my reason for living, for her love, prayers, and care.

Special thanks go to special friends Bahia ZEBANE, Manel CHENAIT, and Chafika BENZAID for their keen interest to complete this thesis successfully.

Contents

Résumé	i
ملخص	iii
Abstract	vii
Acknowledgement	ix
Abbreviations	xix
1 Introduction	1
1.1 Problem Statement and Motivations	1
1.2 Research methodology	3
1.3 Contributions	4
1.4 Thesis organization	5
2 Vehicular Sensor Networks and Smart Cities: Background	7
2.1 Introduction	7
2.2 Smart city	8
2.2.1 Domains of a smart city	8
2.2.2 Generic smart city architecture	10
2.2.3 Wireless sensor networks for smart city	10
2.2.4 WSN technologies for smart cities	11
2.2.5 Data collection in traditional WSNs	13
2.2.6 Data collection in smart city WSNs	16
2.3 Vehicular Sensor Networks	17
2.3.1 Architectures and Applications of VSNs in a Smart City	18

2.3.2	Key issues in VSNs	20
2.3.3	Communication technologies in VSN	21
2.3.4	City information sharing in VSNs	22
2.4	Conclusion	23
3	Vehicle-Based Data Collection in Smart Cities: Related Work	25
3.1	Introduction	25
3.2	WSN Data Collection with Mobile Elements (MEs)	26
3.2.1	Mobile elements (MEs)	26
3.2.2	Overview of data collection in WSN-MEs	28
3.3	Vehicle-based urban data collection schemes: overview	30
3.3.1	MDH: Multi-hop Data Harvesting	33
3.3.2	VDM: Virtual Data Mules for data collection in road-side sensor networks	34
3.3.3	MobiClutser: A Rendezvous-Based Approach for Data Collection with Mobile Sinks	36
3.3.4	DATAxis	38
3.3.5	WVSN for smart city applications: a Relevance-based approach for multiple sinks mobility	39
3.3.6	FDO: Fog-based Data Offloading in Urban IoT Scenarios	41
3.4	Discussion	42
3.5	Conclusion	49
4	OMS: Opportunistic Mules for Short Latency Data Collection in Smart Cities.	51
4.1	Introduction	51
4.2	Motivation	52
4.3	Overview	52
4.4	Opportunistic Mules for Short Delay Applications (OMS)	53
4.4.1	Query Dissemination	54
4.4.2	Opportunistic Election of Data Mules	58
4.4.3	Data Harvesting	65
4.4.4	Data Delivery	65
4.5	Simulations and Measurements	65
4.5.1	Performance Metrics	67
4.5.2	Simulation Results	67
4.6	Conclusion	74
5	Congestion Aware Data Collection With Mobile Sinks in Smart City	77
5.1	Introduction	77

5.2	Motivation	78
5.3	Assumptions	78
5.4	The scheme design	79
5.4.1	Slave sinks designation	79
5.4.2	Data collection	82
5.4.3	Data Delivery	82
5.4.4	Flowcharts	82
5.5	Simulation Study and Measurements	84
5.6	Conclusion	88
6	Conclusion	89
6.1	Summary of contributions	89
6.2	Perspectives	92

List of Figures

2.1	WSN data collection infrastructure [41]	9
2.2	Smart city system building blocks [64].	10
2.3	Generic smart city architecture [40].	11
2.4	Wireless technologies classification according to range and throughput characteristics [80].	13
2.5	An example of a Wireless Sensor Network.	14
2.6	Architectures and applications of VSNs in a smart city [109].	20
3.1	Architectures of WSN-MEs with MDCs.	28
3.2	Architecture of a WSN-ME with mobile peers.	28
3.3	Reference scenario for data collection in WSN-MEs.	31
3.4	Architecture of a WSN-ME with mobile peers.	31
3.5	Multi-hop data harvesting scheme with replicas.	33
3.6	Multi-hop data harvesting scheme with replicas and aggregation.	35
3.7	Data collection in an urban environment along a desired path with VDM.	36
3.8	Data collection and forwarding phases in MobiCluster.	37
3.9	Dataxis proposed scenario: Opportunistic data exchange between taxi cabs and IoT nodes.	38
3.10	Movement path of a sink.	40
3.11	Interactions in the reference three-tier fog network.	41
4.1	(a) Considered scenario. (b) Scheme overview.	53
4.2	Opportunistic Mules for Short Latency Data Collection: Scheme Architecture.	54
4.3	Opportunistic election of data mules: inputs parameters: (a)- DMEA's parameters, (b)- angular error, (c)- positional error.	59
4.4	Opportunistic mules election: examples of candidates with determined destination and candidates with undetermined destination.	61
4.5	OMS global diagram.	64
4.6	Road network topology in the Bab ezzouar, Algiers, Algeria, urban area.	66
4.7	Impact of weighting coefficients of the fitness function parameters in low, medium and high density networks.	68

4.8	Percentage of road segments covered with DMs, varying the vehicular density.	69
4.9	CDF of road segment ratio with DMs in low, medium and high density networks.	70
4.10	Propagation delay of the request for DMs, varying the vehicular density.	71
4.11	The network load induced by query propagation and DMs election. . . .	71
4.12	Impact of DMEA size on scheme efficiency.	72
4.13	OMS efficiency in three AoIs with different sizes.	73
4.14	Impact of AoI size on query propagation delay.	74
4.15	Impact of AoI size on network load.	75
5.1	Data collection architecture.	79
5.2	Slaves designation process.	81
5.3	The master sink operation diagram.	83
5.4	Slave sinks operation diagram.	84
5.5	The evolution of slaves designation over time.	85
5.6	The evolution of slaves designation in space.	86
5.7	Percentage of accessibility to IoT nodes with different traffic volumes. .	87
5.8	Data collection latency with different traffic volumes.	88

List of Tables

3.1	Key characteristics of the studied vehicular-based data collection schemes for urban environments	44
3.2	Comparative study of the studied vehicle-based data collection schemes for urban environments (continued)	46
3.2	Comparative study of the studied vehicle-based data collection schemes for urban environments (continued)	47
3.2	Comparative study of the studied vehicle-based data collection schemes for urban environments (continued)	48
4.1	List of abbreviations.	53
4.2	List of symbols.	56
4.3	Simulation Setup.	66
4.4	The values of weighting coefficients of the fitness function parameters.	67
5.1	Simulation Settings	85

Abbreviations

6LoWPAN	Low power Wireless Personal Area Networks
API	Application Programming Interface
ATMS	Advanced Traffic Management System
BLE	Bluetooth Low Energy
BWA	Broadband wireless access
CALM	Continuous Air-interface conceived for Long-and Medium-range
C2CNet	Car-to-Car NETwork architecture
C2C-CC	Car-to-Car Communication Consortium
CEN	Comite Européen de Normalisation
CH	Cluster Head
CMAC	Cooperative Media Access Control
DSRC	Dedicated Short-Range Communication
ESMT	Euclidean Steiner Minimum Tree
FSO	Free Space Optics
GPS	Global Positioning System
ICT	Information Communication Technology
IEC	International Electrotechnical Commission
IoT	Internet of things
ISO	International Organization Standardization
ITS	Intelligent Transportation Systems
VANETs	Vehicular Ad hoc NETworks
VSNs	Vehicular Sensor Networks
LMDS	Local Multipoint Distributed Service
LPWAN	Low Power Wide Area Networks
LoRaWAN	Longe Range Wide Area Networks
MAN	Metropolitan Area Network
MDCs	Mobile Data Collectors
MEs	Mobile Elements
MMDS	Multi-Channel Multi-point Distributed Service
MRs	Mobile Relays
MSs	Mobile Sinks

MULE	Mobile Ubiquitous LAN Extensions
OMS	Opportunistic Mules for Short latency data collection in smart cities
QoL	Quality of Life
RN	Rendezvous Node
RSU	Road Side Unit
RFID	Radio-Frequency Identification
SA-SWS	Self Adaptive Sleep/Wake up Scheduling
SN	Sensor Nodes
TTL	Time-To-Live
UNB	Ultra Narrow Band
V2I	Vehicle-to-Infrastructure
V2V	Vehicle-to-Vehicle
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network
WMAN	Wireless Metropolitan Area Network
WPAN	Wireless Personal Area Network
WSAN	Wireless Sensor and Actuator Network
WSNs	Wireless Sensor Networks
WVSN	Wireless Visual Sensor Network

Chapter 1

Introduction

1.1 Problem Statement and Motivations

Nowadays, cities are facing new challenges that are characteristics of modern society, including an aging population, reduced energy consumption and carbon emissions, the struggle for greater sustainability, and economic growth. In addition, the size of cities is increasing rapidly due to migratory movements. Indeed, it is expected that by 2050, about 70% of the world's population will live in cities [85].

To meet these challenges, smart city initiatives have emerged, suggesting a new approach of thinking about the development and management of cities. Actually, several definitions of smart cities exist. However, since late 2014, a definition of "smart cities" has been approved by international standards bodies. In fact, on January 2015, The ISO/TMB SAG (International Standards Organization Technical Management Board, Strategy Advisory Group) has launched the following definition of "smart city" which is being automatically adopted by the other international, more specific standardization bodies (International Electrotechnical Commission IEC, International Telecommunications Union ITU, Comité Européen de Normalisation CEN)[78]: *A smart city is one that*

*... dramatically increases the pace at which it improves its **social economic and environmental (sustainability) outcomes**, responding to challenges such as climate change, rapid population growth, and political and economic instability... by fundamentally improving how it **engages society**, how it applies collaborative leadership methods, how it works across disciplines and city systems, and how it **uses data information and modern technologies**...*

Accordingly, smart cities are heterogeneous systems that merge technological and social aspects while fostering the citizen participation. Nevertheless, efficiently running these systems at an urban metropolitan scale raises new challenges that range from urban-scale social and physical sensing to cloud-based urban data management, going through energy-efficient wireless communication protocols, privacy by design, heterogeneous distributed data collection, visual analysis, and urban system software engi-

neering [83]. Actually, infrastructure management is transformed towards a data-driven approach following four basic building blocks: data, analysis, feedback and adaptability [87]. This is done by deploying information systems which are fuelled using elements of the Internet of things (IoT) as data sources, such as Smartphones, radio-frequency identification (RFID) cards and wireless sensor networks (WSNs).

WSNs are core components for data acquisition in smart city applications. Indeed, due to their low cost and small size, sensing devices can be easily deployed at a large variety of locations allowing coupling the physical infrastructure to information and communication technologies. This has caused a significant and rapid increase in the number of IoT objects deployed on streets, which has resulted in massive amount of data to be delivered to data centers. In fact, the issue of collecting the data from the huge number of sensors and actuators spread across streets is still considered as the main challenge for smart city realization. As a matter of fact, the complexity of the urban environment can cause network fragmentation due to forcefully sparse node deployments or unpredictable dynamics in wireless link availability [19]. Even if cellular communications have been suggested as an alternative solution, it could be very expensive because of the need of a permanent infrastructure to be built to cover all the nodes. Moreover, it is difficult from economic and administrative points of view to equip each sensor device with a SIM card [18]. Furthermore, the energy consumed for the data collection process could be critical and highly dependent on the quality of the required communication.

One effective, economical and suitable way to deliver sensor data to the central controller is to exploit Mobile Elements (MEs) that may be robots or vehicles. In fact, while moving, MEs can download sensors data once they are within their communication range [19, 18, 105, 43, 91, 47, 97]. This approach, well known as MEs-based data collection approach, is proved to be effective in energy economy as well as energy consumption balancing which are considered as the main issues in traditional WSN multi-hop forwarding strategies.

On the other hand, vehicles could contribute to efficient operation of the city by serving as IoT mobile nodes. Indeed, the proliferation of vehicles and the maturity of ad hoc network technologies (VANETs), make today vehicles capable of retrieving data and transmitting them to collection gateways [23, 5]. Several works deal with urban data collection using vehicles as MEs. In [19], the data mule concept is adapted to urban mobility by handing over the mule role to a better candidate, when the trajectory of the vehicle acting as data mule does not fit the planned trajectory. The data mule concept is then extended to a more realistic entity called “oblivious data mules” [18] where taxicabs are used to deliver the data of a smart city to the urban monitoring center. Vehicles are also proposed to act as mobile sinks to collect the data from stationary sensors in smart cities [91, 60].

However, these works employ vehicles to collect data for long trajectories that correspond to vehicles' itinerary. Therefore, data delivery latency is obviously high. Given that some smart city applications are required to respond in a relatively smaller delay, reducing the latency of such a promising approach is crucial. For instance, a clogged storm drain should be reported within a limited delay; otherwise, it might be too late to react. An ATMS (Advanced Traffic Management System) requires frequent updates of sensor data to advertise useful information about traffic conditions, air pollution, and environmental noise to drivers. The high latency of the above mentioned vehicle-based data collection works makes them unsuitable for this kind of applications. This motivates us to deal with the problem of decreasing latency as much as possible.

Therefore, this thesis aims to reduce the data collection latency of the vehicle-based approach to meet the requirements of short-delay smart city applications. The thesis outlines the principal factors that lead to high latency of vehicle-based data collection in urban environment, and it proposes two reduced-latency data collection schemes for smart city applications. Below, are the specific objectives of our work.

1. Leveraging opportunistic vehicles to design a short latency data collection schemes for roadside sensors of urban areas,
2. Designing an optimal request propagation strategy in order to provide a distributed data collection process,
3. Identification of suitable vehicular routing algorithms that can be used to deliver data collected by vehicular collectors to data centers while maintaining a reduced data collection latency,
4. Performance evaluation of the proposed schemes.

Hence, this thesis aims to contribute with novel cost effective and energy efficient data collection solutions for a specif class of smart city applications, that is short delay applications.

1.2 Research methodology

In order to reach the thesis goals, we have applied a design and creation [20] research methodology. For each contribution, the five stages of the applied research methodology have been followed: awareness, suggestion, development, evaluation and conclusion. The first step in the research process was to identify the research issues. Essentially, the literature on vehicular sensor networks (VSNs) and smart cities has been reviewed and data collection solutions related to VSNs and smart cities have been critically assessed. This has allowed us to highlight issues that few research studies have addressed. Then, we sketched and proposed solutions that were implemented and evaluated using network

simulators. To this end, a set of metrics were used to evaluate the results and real data were used in the simulations where possible. Finally, for each solution, the process concluded by identifying the main contributions to the state of the art and outlining future works.

1.3 Contributions

In this section we summarize our contributions and the publications derived from [them](#). The main contributions of this thesis are [three folds](#):

- A classification of vehicle-based data collection schemes for smart city is proposed in this thesis. We classified the reviewed schemes according to two criteria:
 - i) *The type of vehicles used for data gathering*: whether they are specific vehicles initially dedicated for data collection such as municipality vehicles, or opportunistically elected from citizen cars.
 - ii) *Application response-time*: whether the data collection scheme is proposed for delay tolerant applications or short delay applications.
- The design of a cost effective short latency urban data collection scheme by leveraging opportunistic inter-vehicular communications. The scheme opportunistically designates to each road segment a vehicle that will act as a data mule. To that end, an optimized strategy to disseminate a request for data mules is proposed. The proposed scheme was described in [27] in 2018 and it was presented at the seventh IEEE International Conference on Smart Communications in Network Technologies¹.

The results of this contribution were extended with the enhancement of the opportunistic data mules election procedure. In this version of the proposed scheme called, OMS: Opportunistic Mules for Short latency data collection in smart cities, weighting coefficients of election parameters were considered. Deeper analysis and extensive simulation scenarios were provided. OMS was published in [26], in a special issue of International Journal of Communication Systems² on Smart Communications for Autonomous Systems in Network Technologies. However, OMS relies only on opportunistic mules whose presence is likely to be reduced in sparse networks, such as in nighttime scenarios. It was therefore necessary to deploy a set of dedicated vehicles to collect data from roads with low vehicle density.

¹International Conference on Smart Communications in Network Technologies (SaCoNeT). <https://ieeexplore.ieee.org/xpl/conhome/8552598/proceeding>

²<https://onlinelibrary.wiley.com/doi/full/10.1002/dac.4445>

- As a consequence, a new reduced latency urban data collection scheme is designed by integrating specific vehicles that are initially dedicated for urban data collection along with citizen cars. By involving dedicated vehicles, we aim at increasing vehicles' availability in low dense roads. In this new scheme, opportunistic vehicular communications are leveraged to create a collaborative data collection process that allows dedicated vehicles embedded with mobile sinks to act as masters by delegating slave sinks among citizen cars to collect data from congested roads. Master sinks will then, be responsible only for collecting data from low-density roads. By doing so, we guarantee a uniform distribution of data collection process. The proposed scheme was described in [25] on 2020 and it was presented at The 2020 International Symposium on Networks, Computers and Communications³.

1.4 Thesis organization

The thesis is organized as follows:

Chapter 2 provides a background to frame the following chapters of this thesis. It describes common features in smart city architectures focusing on the data collection issue. The chapter also gives an overview of Vehicular Sensor Networks (VSNs), their various communication architectures, their promising applications, and the main challenges faced by the VSN paradigm.

In Chapter 3, we investigate the schemes that utilize vehicles to collect data from roadside sensors in urban areas. We first introduce the mobile element-based data collection approach. Then we review existing vehicle-based data collection schemes for urban areas and provide a taxonomy.

Chapter 4, presents the design and the performance of OMS: Opportunistic Mules for Short delay smart city applications, our reduced-latency vehicle-based data collection scheme for smart city applications.

In Chapter 5, a new reduced-latency vehicle-based data collection scheme is designed. The new scheme employs specific vehicles devoted for urban data collection to take in charge low dense roads and to delegate opportunistic sinks for congested roads.

To conclude, Chapter 6 summarizes the main contributions of this thesis and discusses potential future work directions.

³IEEE International Symposium on Networks, Computers and Communications (ISNCC 2020). <https://ieeexplore.ieee.org/document/9297299>

Chapter 2

Vehicular Sensor Networks and Smart Cities: Background

2.1 Introduction

Cities are confronted with new challenges that are raised by the characteristic of contemporary societies: economic recovery, greater sustainability, reduction of energy consumption and carbon emissions, etc. To meet these challenges, smart city initiatives have emerged, offering new manners of looking at city development and management. To that end, investments in human and social capital as well as in traditional (transport) and modern (ICT) communication infrastructure should fuel economic growth and improve quality of life while judiciously managing natural resources. Accordingly, infrastructure management is evolving into a data-driven approach following four basic elements: data, analysis, feedback, and adaptability.

The main source of data in smart cities is the IoT elements which may be: Smartphones, radio-frequency identification (RFID) cards, wireless sensor networks (WSNs), and vehicular sensor networks (VSNs). The data collected by these elements is used in a wide range of applications.

One of the main goals of building smart cities is to improve the efficiency of traffic management. To this purpose, intelligent transportation systems (ITS) are being implemented by fitting contemporary vehicles with various sensing devices such as actuators, GPS devices, and micro-embedded computers, providing them with data collection and processing functionalities. In addition, vehicles can communicate with each other or with roadside infrastructures to form what is called a vehicular sensor network (VSN).

This chapter includes the background to frame the following chapters of this thesis. Section 2.2 describes common features in smart city architectures focusing on the data collection issue. Secondly, Section 2.3 gives an overview of VSNs, showing their various communication architectures, their promising applications, and the main challenges faced by the VSN paradigm. Finally, Section 2.4 concludes the chapter.

2.2 Smart city

The concept of “*smart city*” has been proposed both to resolve bottlenecks in urbanization and to support sustainable urban development. A city is smart if it is able to capture, analyze and integrate critical information in public safety, industrial activities, social public services, etc. With the rapid evolution of the IoT and cloud computing, the benefits of a smart city are growing and include improving the quality and performance of urban services, reducing costs and wasting resources, as well as improving the management capacity of governments. According to Caravela IoT solutions engineering company¹: “A Smart City is an urban area that uses different types of IoT sensors to collect data then use insights gained from that data to serve centralized (governmental) applications such as efficient management of assets, resources, and services and decentralized (citizen) applications such as insights into rainfall, micro-climates and air quality. Cellular technology is used for data collection where there is a ready source of power, where large data throughput is necessary and where there are relatively few sensor endpoints. Low-power wide area networks are used where there is no ready source of power, minimal data throughput is sufficient, and where thousands of endpoints must be supported.” The infrastructure shown in figure 2.1 is designed to collect and deliver information captured by urban sensors to the city servers. The information flow in this schema begins from the sensors, which gather data about their environment and transfer it to a gateway. The latter delivers the collected data to the data center.

2.2.1 Domains of a smart city

The motivation behind smart cities is to make citizens life easier by improving their quality of life (QoL). To this purpose, smart city should affect several aspects of the good operating of the city. This section describes the most common smart city domains referred in the literature.

People: Since the purpose of a smart city is to help people, it is then necessary to pay special attention to the human factor. The following are examples of possible solutions that could develop this factor: improving the services provided by the healthcare system [59], Bringing new ways of using technology to improve the education system, maintaining public spaces, enhancing public safety as well as cultural and entertainment activities [86][84].

Economy: Smart economy aims at improving the economy and business opportunities by incorporating innovative solutions based on technology. This would increase the flexibility of the labor market, offer citizens various economic opportunities, overcome the challenges of economic globalization, promote sharing economy, and provide a balanced and sustainable economic growth [64].

¹<https://www.caravelaiot.com/2019/11/1294/>

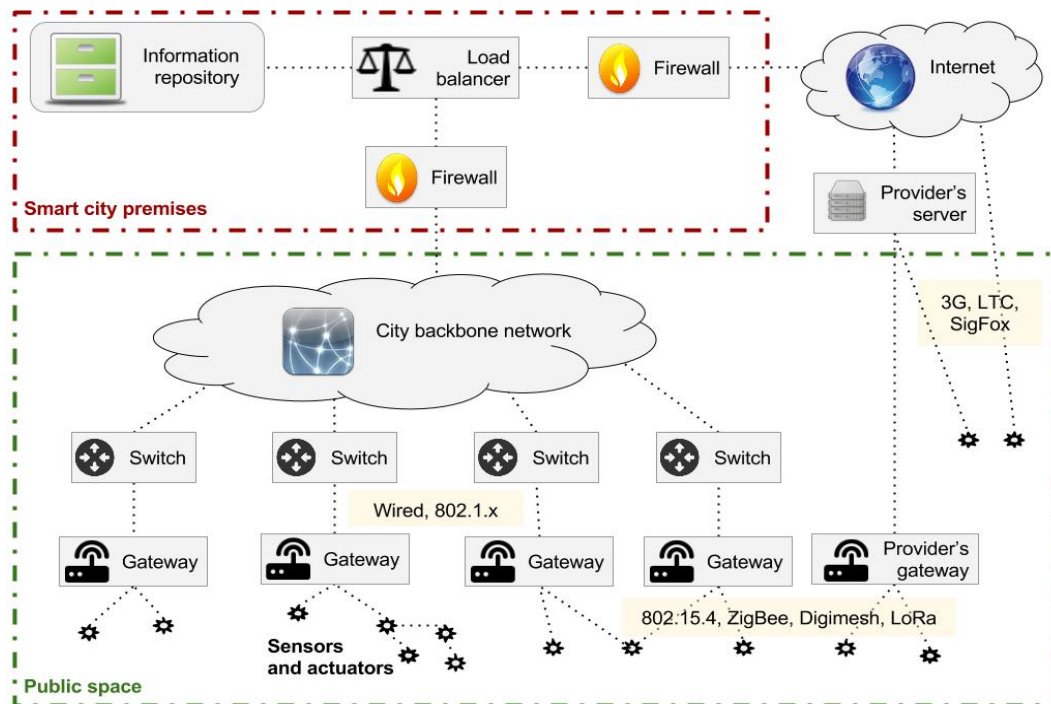


Figure 2.1: WSN data collection infrastructure [41]

Mobility: Smart Mobility focuses on the mobility of people as well as on that of vehicles. Among its objectives we cite: supporting walk-ability and cycling, effectively managing vehicles' traffic and pedestrian movement, minimizing traffic jams, providing mass rapid transit systems, such as metro rail, light metro, monorail, or "skytrain" for high-speed mobility, and providing seamless mobility for differently-abled people [64].

Management: This domain is one of the most developed aspect among smart cities. It concerns transport, energy, and natural resources management [7]. Examples of initiatives being implemented at present are: improving public transport and city logistics, control of public lighting, energy demand and renewable energies, management of crop cultivation and feed dosing [79], production development for farms, allocation and measurement of water productivity for agriculture [89], municipal and domestic purposes.

Governance: The smart government consists on *digitalizing* the administration to improve government systems efficiency. For example, the voting system can be improved by incorporating the e-democracy which aims at providing citizens with an easy access to public documents, hence improving transparency [86].

Infrastructures: This aspect encompasses facility management such as property

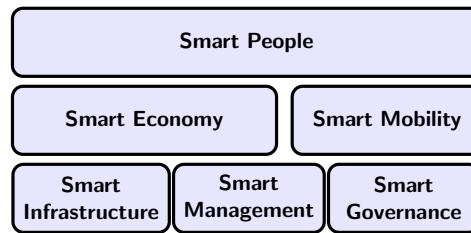


Figure 2.2: Smart city system building blocks [64].

and leasing, and services such as cleaning and maintenance. Quality of Living (QoL) aspects such as lighting, heating, and ventilation are also included. Figure 2.2 summarizes the six building blocks of a smart city.

2.2.2 Generic smart city architecture

The architecture of smart city information systems follows certain common patterns that can be easily found in most smart cities. In fact, smart city frameworks have shoot up with the aim of facilitating the development of applications involving multiple shareholders by providing a platform of common services.

The design of smart city systems typically follows a service-oriented, three-layer architecture as it is illustrated by figure 2.3. The first layer includes the required elements to collect information from the city (e.g. sensor devices, surveillance cameras, social networks, supervisory control and data acquisition systems). The second layer provides the city with services such as relational and non-relational storage, geographic information systems, data analysis, cloud computing, natural language processing, business intelligence, etc. In fact, this layer acts as a middleware, which provides the city with an API to connect the elements of the first layer to the services offered in this layer. The elements of the infrastructure shown in figure 2.1 are part of the elements from the first and second layers. Finally, the third layer is an application layer, in which applications based on data and services provided by the middle layer are implemented.

2.2.3 Wireless sensor networks for smart city

Sensor networks are essential components of smart cities as they are the main technology for collecting data. Wireless sensor networks (WSNs) are self-organized networks that can be built by relatively cheap and small sensors with low power consumption and maintenance cost whose ability to transmit data remotely allows their deployment at a large variety of locations (on the ground, in the air, in vehicles, on bodies, inside buildings, etc). A sensor node, also known as mote, consists of three basic units: a sensing unit for data acquisition from the physical environment, a processing unit for local data processing and storage, and a wireless communication unit for data transmission. Motes may be also equipped with actuators that enable them to execute

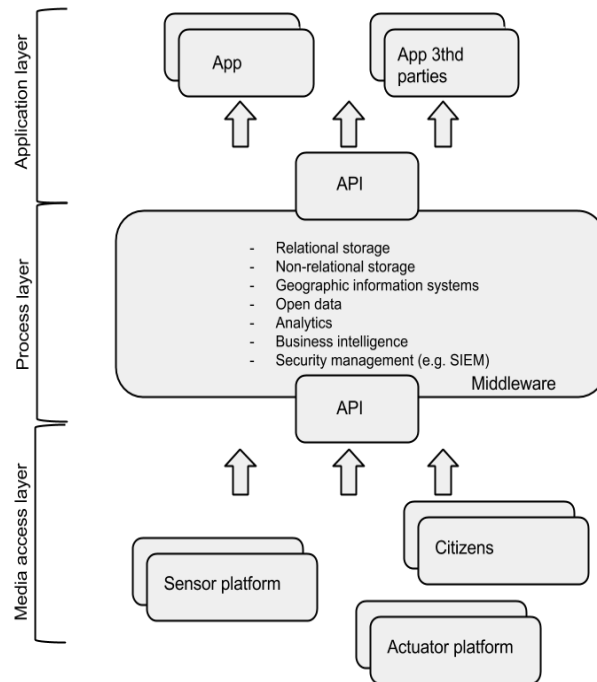


Figure 2.3: Generic smart city architecture [40].

actions. In this case, the networks are known as wireless sensor and actuator networks (WSANs). WSNs are based on autonomous cooperative communications to transmit sensors data to the smart city data centers. This data will first be sent to a device on the edge of the WSN called a gateway or base station. Gateways are equipped with multiple communication interfaces. Different types of wireless networks can be built based on the various existing wireless communication technologies. In the following, we present the most appropriate WSN technologies for smart cities.

2.2.4 WSN technologies for smart cities

The most important technologies according to communication range and throughput are classified in figure 2.4.

WPAN (Wireless personal area networks) is developed by the IEEE P802.15 Working Group for short-range wireless networks. This technology meets the communication and interoperability requirements of ubiquitous computing devices such as PCs, peripherals, cell phones, consumer electronic devices, etc. Primary considerations are maintaining low power consumption, small size and low cost devices. Relevant technologies are: Bluetooth, ZigBee, and 6LoWPAN. Bluetooth is used for streaming data applications and control and monitoring functions. A typical use case of Bluetooth in smart city is *Smart Ticketing* [39] which focus on

using Bluetooth Low Energy (BLE) beacons for mobile ticketing in urban passenger transport. Zigbee is designed for small scale projects which need wireless connection such as medical device data collection, home automation, and other low-power low-bandwidth needs (e.g. street lighting control [68]), where as 6LoWPAN is designed to enable low-power devices with limited processing capabilities to participate in the Internet of Things.

WLAN (Wireless local area networks) provide low-range but broad throughput wireless networks for a limited area such as a home, school, laboratory, or office building. Wi-Fi is the most widely used WLAN in the world. For instance, in Montreal, Wi-Fi access points are deployed on the streets not only for providing citizens with the Internet, but also to connect IoT objects that require a wide bandwidth such as IP-based traffic cameras [65].

WMAN (Wireless metropolitan area networks) has been developed by the IEEE 802.16 Working Group to respond to the market demand for Broadband wireless access (BWA). This technology is intended for cellular service providers and leaves many of the implementation decisions to the implementer, particularly at the base station side [21]. WMANs can be implemented using a variety of wireless technologies: LMDS (Local Multipoint Distributed Service), MMDS (Multi-Channel Multipoint Distributed Service), FSO (Free Space Optics), WiMAX (Worldwide Interoperability for Microwave Access), etc. For instance, in [68], WiMax is used in conjunction with the ZigBee sensor network for complete control of street lighting.

LPWAN (Low-power wide-area networks) has emerged as an answer to IoT applications that require long transmission range, low power consumption, and low-cost deployment. LPWAN provide a communication range up to 10 km in urban areas [72] with a bit-rate ranging from 0.3 kbit/s to 50 kbit/s per channel [4]. This makes them suitable for IoT applications that require transmitting a small amount of data multiple times per day. LoRaWAN, SigFox, , and NB-IoT are arguably the most widely adopted technologies. For instance, LoRaWAN promises pervasive connectivity in outdoor IoT applications while keeping network structures simple and manageable. In addition, LoRa uses unlicensed frequencies which are freely accessible. It is perfect for battery-operated sensors and low-power applications in the Internet of Things [8]. Smart agriculture [82], smart water management [66], and smart parking [51] are among the projects that have solved real problems with LoRaWAN. A technical overview of the three competing technologies is presented in [29].

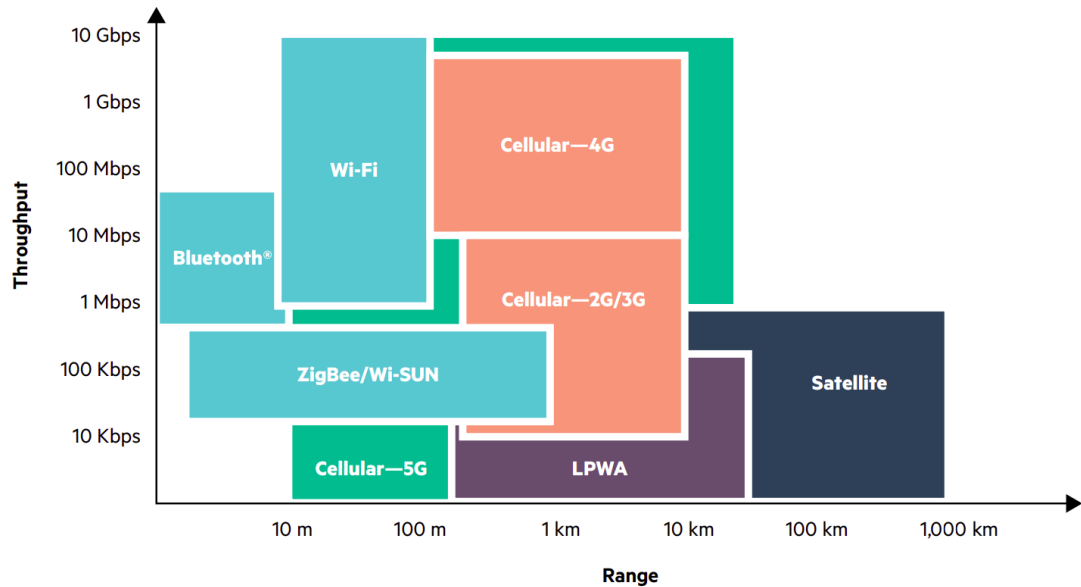


Figure 2.4: Wireless technologies classification according to range and throughput characteristics [80].

2.2.5 Data collection in traditional WSNs

Data collection is one of the main operations performed in WSNs, where a central unit or base station receives all of the data generated by each sensor through wireless communications. Data collection is necessary for estimating network size, determining the system load, processing user queries, etc. Several approaches have been designed to efficiently collect WSN data with different objectives such as energy efficiency, delay, network throughput, energy-delay trade-off, etc. Aggregation functions such as MAX, MIN, AVG can also be used when collecting WSN data to meet specific demands like user queries or network characteristics estimation. In this case, the data collection process is referred to as data aggregation. Figure 2.5 depicts an example of a WSN including one sink node denoted by S and 9 sensor nodes denoted by S_i ($0 \leq i \leq 8$). S is connected to the Internet through a wired or wireless link. Each bidirectional edge between two nodes in Figure 2.5 corresponds to a one-hop communication link. For two nodes without a direct link, multi-hop communications are used for data transfer.

For instance, if the WSN in Figure 2.5 is used to monitor the temperature of an area of interest, then, if the purpose is to obtain the average temperature value of a monitored area at time t , aggregation is applied as an AVG function. On the other hand, if individual temperature values are required at time t , sensors transmit their sensed temperature to the sink.

When dealing with data collection for WSNs, the limitations of WSNs on power

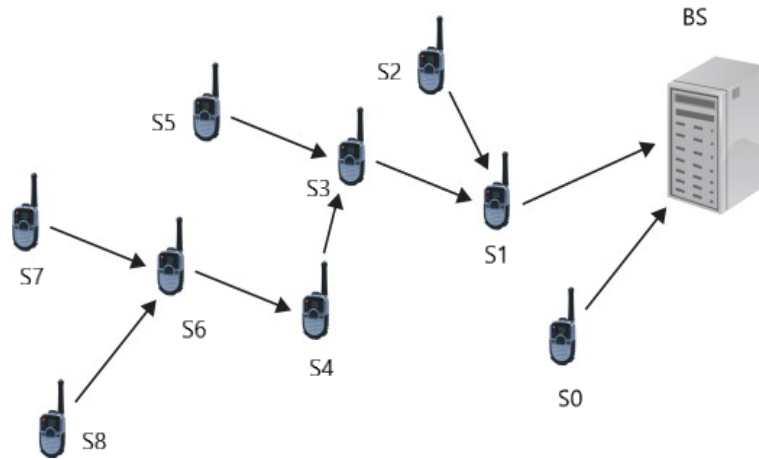


Figure 2.5: An example of a Wireless Sensor Network.

supply, computation, storage, and communication capacities bring several challenges which can be summarized as follows [50].

Energy Efficiency. Most sensor nodes are battery-powered, which makes them energy-limited. In addition, it is usually undesirable or even impossible to recharge sensors deployed in hostile areas. This requires taking into consideration the energy efficiency criterion when designing a data collection method.

Timeliness and Real-time. The interference caused by wireless communication together with the limited communication capacity of the sensor nodes makes designing efficient data collection algorithms for real-time applications a challenge.

Scalability and Robustness. WSNs are large-scale networks with a dynamic topology. At any time, new nodes may join a network and other existing nodes may disappear. Therefore, scalable data collection algorithms are highly desired. On the other hand, it is challenging to design optimal distributed algorithms due to the difficulty of obtaining the overall real-time network information. Consequently, designing optimal distributed data collection algorithms is also challenging.

Time Synchronization and Distributed Solutions. It is difficult to obtain a strict time synchronization for WSNs due to the unstable deployment environment, clock drifts, and technical limits. Therefore, to understand the performance of data collection for practical WSNs, it is important to investigate distributed data collection algorithms for asynchronous WSNs.

Other Issues. There exist many other challenges for data collection in WSNs induced by node mobility, security issues, etc. To address these challenges, application-aware data collection schemes are to be developed.

Much research has been done on data collection in the WSNs. In the following, we will summarize some of them according to their objective.

Energy-Efficient Data Collection. In [107] optimal solutions are developed for the traffic-aware relay node deployment problem, with the aim of maximizing the network lifetime. The problem is transformed into a generalized Euclidean Steiner Minimum Tree (ESMT) problem and a hybrid algorithm that returns optimal results with all test cases is developed. However, the solution may return fractional numbers of relay nodes as it was designed for the continuous domain. The authors hence, developed algorithms for discrete relay node assignment, together with local adjustments that yield practical solutions. In [106], the WSN is partitioned into clusters using an Approximate Data Collection (ADC) algorithm. In addition, local data correlations on each cluster head is discovered, and a global approximate data collection is performed at the sink based on the model parameters provided by cluster heads. To this end, the authors designed a local estimation model to approximate the readings of the clusters. Simulation results show that ADC can reduce the communication cost with given error bounds. A near-lifetime-optimal and scalable data collection solution is presented in [69]. Instead of using a fixed network topology, the authors construct a set of communication topologies and apply each topology to different data collection cycles. The solution provides a spatial load balancing by flexibly distributing the traffic load across different routes in the network. Moreover, it balances the energy consumption in the time domain. The solution is further extended, for the aim of scalability, in order to be applied to networks with multiple base stations which can cooperate with each other to provide a distributed data collection solution.

Delay-Aware Data Collection. Besides energy saving, there are strong needs for developing WSNs algorithms with other optimization priorities such delay reduction. A delay-aware data collection network structure for WSNs is proposed in [28] with the aim of minimizing the data collection delay. To this end, two network formation algorithms are designed to construct the proposed network structure in a centralized and a decentralized approach. To minimize the data collection delay of an event-driven WSNs while maximizing the network lifetime, an anycast-based packet forwarding scheme is developed in [58] which allow to each sensor node to opportunistically forward the data packet to the first neighbor that wakes up among several candidates. To this end, anycast forwarding schemes optimization has been investigated in order to minimize the expected packet-delivery delays from the sensor nodes to the sink. Then, an optimal solution is provided to jointly control the system parameters of the sleep-wake scheduling protocol and the any-cast packet-forwarding protocol to maximize the network lifetime, while taking into account the constraints on the expected end-to-end packet-delivery

delay. In [94], a delay and energy efficient data collection protocol for WSNs is proposed, where time-slots for data forwarding are allocated with the aid of a Self Adaptive-Sleep/Wake-up Scheduling (SA-SWS) protocol using Cooperative Media Access Control (CMAC). Forwarding nodes are chosen from nodes having the highest residual energy. Using the reinforcement learning technique, each node is able to autonomously decide, in a decentralized manner, its own operation mode (sleep, listen, transmit).

Mobility-Aware Data Collection. Actually, data collection solutions for static WSNs have suffered from the unbalanced energy exhaustion issue. Indeed, despite the considerable efforts made to design energy conservation techniques, the multi-hop data transmission mechanism followed by the WSNs for data collection, leads to a faster depletion of the battery of the nodes closest to the sink because they should relay more packets than other nodes. Therefore, the network is subject to disconnect due to the nodes' premature energy depletion. Moreover, the inefficiency and unreliability of long-distance imperil the convenient delivery of data to the sink. To face these challenges, an interesting approach is to designate mobile elements such robots, animal, or human beings as responsible for data collection. Mobile nodes with adequate mobility pattern may visit sensors individually and collect their measurements. This provides implicit resolution of problems related to the network connectivity and multi-hop communication [19]. Furthermore, mobile nodes trajectories can be optimized to minimize the probability that sensors experience buffer overflow [37]. Data collection with mobile elements will be investigated in detail in the first section of the next chapter.

2.2.6 Data collection in smart city WSNs

Realizing smart city projects requires deploying billions of sensors and actuators for collecting information from city's infrastructures resulting in many self-organized WSNs. Data generated by the WSNs is to be efficiently delivered to data centers for processing, analysis, and storage in the cloud, which notifies municipal services for possible infrastructure maintenance. However, the issue of collecting data from WSNs deployed in urban environments differs from that in traditional WSNs. In traditional WSNs, most data collection schemes have provided mechanisms for routing data from each sensor node to the sink through multi-hop communications. However, the traditional WSN architectures are based on the assumption that the network is dense, so that any two nodes can communicate with each other through multi-hop paths [74] which is not the case of urban scenarios. In fact, the complexity of the urban environment can cause network fragmentation due to forcefully sparse node deployments or unpredictable dynamics in wireless link availability [19]. Even if cellular communications have been suggested as an alternative solution, it could be very expensive because of the need of

a permanent infrastructure to be built in order to cover all the nodes. Moreover, it is difficult from economic and administrative points of view to equip each sensor device with a SIM card [18]. Furthermore, the energy consumed for the data collection process could be critical and highly dependent on the quality of the required communication. Thereby, data collection with mobile elements is the most suitable way to collect the WSNs data in smart cities, especially with the ubiquitous and opportunistic availability of mobile elements in urban environments, namely vehicles and pedestrians.

On the other hand, Vehicular Ad hoc Networks (VANETs) play an important role in the efficient operation of the city. Actually, contemporary vehicles are equipped with many sensors, actuators, and communication interfaces that enable them ubiquitously sensing environmental phenomena, and communicating with the surrounding devices such as other vehicles, roadside infrastructure, and IoT objects. This vehicular sensor networking (VSN) paradigm provides an alternative cost effective solution for data collection in smart city, as it leverages an already deployed cost-free infrastructure [98], because vehicles may incidentally cross the coverage area of active sensors, therefore, they can exchange data. Moreover, data offloading from IoT sensors to the vehicle can be performed using only one-hop communication, which makes the vehicle-based data collection energy-efficient. In the literature, many works have suggested using vehicles as mobile elements [19, 18, 98, 71, 61, 91, 62] for data delivery from IoT sensors to the cloud. An overview of existing vehicle-based data collection schemes is presented in section 3.3 of the next chapter, while the next section of the present chapter is devoted to VSNs.

2.3 Vehicular Sensor Networks

The concept of smart cities is leading to fundamental changes in our urban lifestyle, as well as in the way cities are governed. The proliferation of vehicles along with the constant expansion of road networks make the intelligent transportation system (ITS) the most crucial function in building smart cities[108]. ITS not only increases efficiency of traffic management, but also provides smart cities with compelling applications. Thanks to the maturation of technologies in the vehicular ad hoc networks (VANETs), vehicular sensor networks (VSNs) may be invoked both for vehicle-to-vehicle (V2V) communications and vehicle-to-infrastructure (V2I) interactions, as well as for sensing, transmitting, and integrating important information related to a city's operation[109]. By deploying sensors and IoT objects along roads, as well as equipping vehicles with sensors, various sources of information can be sensed and transmitted relying on cooperation and coordination among vehicles. However, due to traffic conditions and sudden events such as accidents, VSNs exhibit irregularly fluctuating dynamic topological structures, which increases communication challenges. In this section, we present the VSNs' architectures, key technologies and applications in a smart city.

2.3.1 Architectures and Applications of VSNs in a Smart City

VSNs provide connected sensors to collect data for providing safer and smoother traffic on the roads [8]. Indeed, wireless sensors are embedded on vehicles [69] or on road-side structures [71] to provide drivers and city authorities with relevant information. Contemporary vehicles are already equipped with GPS devices, actuators, various sensing devices, and built-in microcomputers to collect and process data. In addition, vehicles can communicate with each other or with a road-side infrastructure which allows developing services to enhance road user's safety and convenience. Figure 2.6 illustrates the architecture of VSNs for the smart city. As can be observed from the figure, vehicles collect information from city operations, and provide broadcast services with the aid of V2V or V2I communications. Furthermore, vehicles can share significant information on the Internet relying on access networks or hot spots. On the basis of this information, efficient traffic management and long-term urban planning become possible[35]. Relying on VSNs, along with smart sensing and intelligent wireless communication infrastructure, data delivery, processing, and sharing become feasible, giving rise to a wide variety of applications. In the following, we highlight a compelling set of VSNs applications for the smart city.

Traffic Management: Road congestion is an acute problem for metropolitan areas administrators. Traditional traffic management strategies fail to solve the congestion problem because they do not take into account the conditions of the road network, the distribution of the urban population, and the peak/off-peak factor. VSNs provide an effective traffic management solution that relies on information interaction and coordination among vehicles to plan and recommend optimal routes for drivers, to maintain acceptable traffic flow, and to provide constructive feedback for the design of future urban road networks. In addition, by relying on smart sensors (roadside or vehicular sensors) as well as efficient information processing algorithms, intelligent mechanisms for transportation monitoring and management can be designed [36]. Indeed, based on historical monitoring data and real-time GPS information, it is possible to predict traffic flow, thus avoiding heavy traffic routes for vehicles by recommending optimal routes for them. For instance, an intelligent traffic flow control system periodically collects information about traffic to control traffic lights in real-time. Free flow tolling systems are used on roads that require tolls. This system communicates with the driver to electronically collect tolls. Hence, vehicles do not need to stop, resulting in free-flow traffic[8]. In addition, traffic safety can be ensured by helping drivers to comply with traffic conditions, for example by broadcasting speed limit notifications.

Emergency Management: VSNs also help reduce the loss of peoples' lives and

properties by providing early warnings in emergencies[73]. In fact, relying on the flock of vehicles and on their widely distributed locations in a city, and given the information-sharing mechanism, VSNs can reduce the risk of an accident or limit its severity if it happens. For instance, collisions may be avoided using an in-vehicle system that keeps the vehicle at a safe distance from others. Additionally, drivers can be alerted about a speed limit when cornering by deploying sensors near the road bends. In the event of an unfortunate accident, preventing secondary accidents is possible by promptly spreading the time and the location of the accident. Additionally, the timely and effective accident-warning allows assisting the drivers' decision-making in order to avoid traffic jams. Furthermore, to manage fast-track rescue operations, emergency vehicle warning systems are proposed to alert the driver about an oncoming ambulance. Moreover, a potential fire or water/gas leak can be detected by a smart roadside/vehicular sensor, which immediately advertise the city's central controller, to take adequate actions.

Urban Planning: Effective urban planning improves the residents' quality of life as it supports the comfortable mobility of the population, as well as of goods. VSNs are able to collect relevant urban information and to transmit it to data centers to be used by decision makers for road-network layout planning and infrastructure construction.

Infotainment: VSNs play a bridging role in the construction of a metropolitan area network (MAN). By serving as wireless relays and access points for the MAN, the widely distributed vehicles of a VSN participate in providing stable multimedia materials and high-rate Internet services [99]. This will provide to passengers entertainment services such as distributed games where a commuter in a vehicle can play with passengers on other vehicles. In addition, VSNs provide contextual applications that enable drivers to learn about the route. For instance, assisting the diver to find nearby restaurants or available parking spaces and their prices.

Reliability of smart services: VSNs can be used for sensing and handling the genuine " water, electricity and gas billing information" by cars passing near each property [109]. Indeed, Based on the relevant road-network facilities, vehicles are capable of reading the water, gas and electricity meters and reporting them automatically which allows avoiding overdue payment. This is possible by assigning to each vehicle a unique identification number that corresponds to specific services bills. VSNs can also be used to recommend information to drivers reliably connected to smart homes and personal information systems. In fact, based on advanced data mining and machine-learning aided analysis, VSNs can send recommendations to the driver according to his personal will.

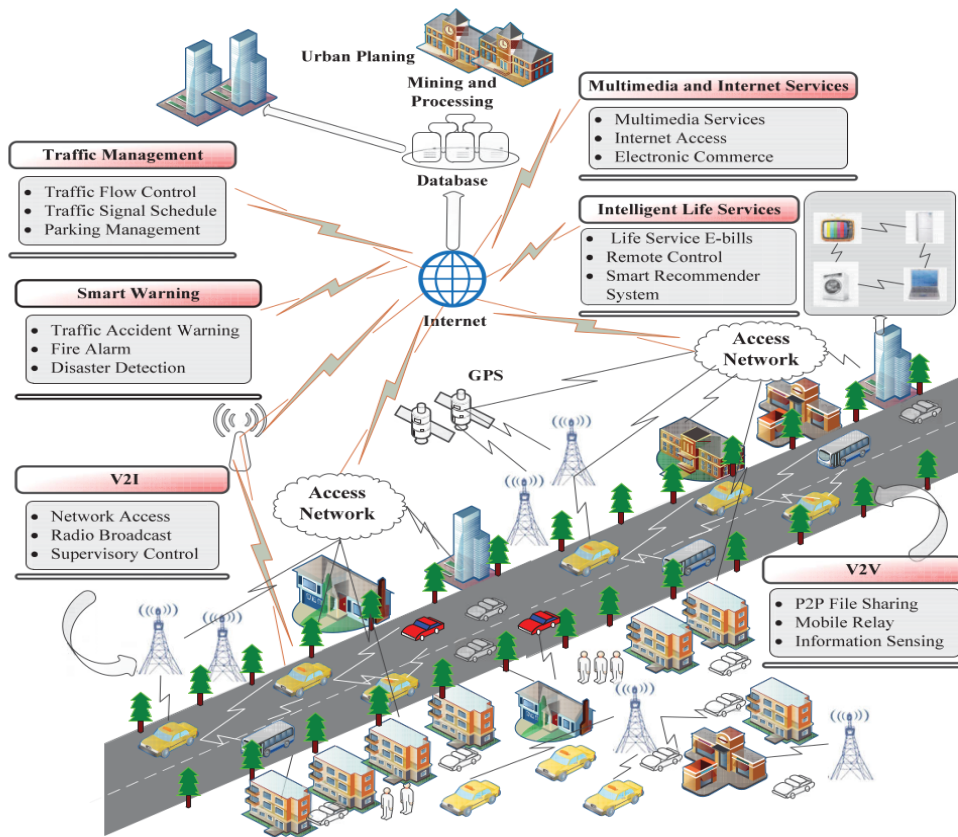


Figure 2.6: Architectures and applications of VSNs in a smart city [109].

2.3.2 Key issues in VSNs

VSNs face many challenges in implementing smart city applications. The key issues of VSNs are summarized in [109] as follows.

- **Wireless Channel Characteristics:** Limited bandwidth, high Doppler shift, and building occlusions are the main challenges facing the VSNs of a smart city.
- **Mobility and Dynamic Topology:** The unpredictable mobility of vehicles leads to a regionally heterogeneous but globally stable network topology.
- **Efficient Routing Strategy:** Efficient vehicular routing protocols are mandatory for optimizing bandwidth allocation as well as for reducing network congestion risks.
- **Congestion Control:** The transmission congestion control is necessary for reducing the probability of communication blocking, especially on roads of high-traffic flow. An appropriate transmission congestion control scheme is capable of improving the network throughput.

- **Security Mechanism:** VSNs applications in smart city require the acquisition and sharing of a large amount of personal information. Therefore, designing reliable security mechanisms is crucial to prevent the interception of malicious information that could potentially lead to loss of life due to road accidents and privacy invasion.

2.3.3 Communication technologies in VSN

Several key technologies and protocols were proposed for constructing a large-scale ad-hoc vehicular network. Examples include the continuous air-interface conceived for long-and medium-range (CALM) enacted by the international organization for standardization (ISO), the car-to-car network architecture (C2CNet) designed by the car-to-car communication consortium (C2C-CC) and the IEEE 1609 protocol. In this section, we discuss some of the vehicular communication technologies. There are two main communication technologies in VSNs: short-range and long-range communication systems.

- **Short-range communication**

Short range communication systems allow communicating over a short distance. An example of such systems includes dedicated short-range communication (DSRC) technology that provides V2V and V2I communications. DSRC is largely used for preventing collisions by stopping the vehicle or producing warning signals when encountering a blind spot [90]. The millimeter-wave (mmWave) system is also proposed for V2V and V2I communications in a smart city. A study presented in [17] claims that mmWave is the only feasible candidate for vehicles that require high bandwidth systems as it allows a data rate delivery of gigabit-per-second. To connect sensors and actuators in an IoT environment, Bluetooth 4.0 is designed. This technology consumes low power and provides a high data rate of about 24 Mbps. IEEE 802.11p presents positive aspects for V2V and V2I communications[77]. Indeed, it provides several non-overlapping channels each of them is broadcasted in a DSRC band of about 5.9 GHz. One of these channels is reserved for control purposes while the others are utilized for different services.

- **Long-range communication**

With the growth of connected objects, the demand for long-range communication networks is increasing. In the foreground is LP-WAN (Low Power-Wide Area Network) which enables Long Range (LoRa) communications of about 10 km of coverage and low power consumption. LoRa wireless communication paradigm allows transmitting minute volumes of data packets with a high resistance to noise. Another LPWAN communication technology used in vehicular communications is sigfox. The latter is based on software-defined cognitive radios. Sigfox utilizes

Ultra Narrow Band (UNB) modulation in the physical layer and promises that each gateway can connect up to a million objects, with a communication range of between 3 to 10 km in an urban environment. Moreover, Sigfox provides the advantages of low noise level and ultra-low power consumption.

5G/6G/7G Cellular network technologies are emerging platforms that provide broadband access. 5G offers tremendous advantages including a high data rate, higher bandwidth, low latency, higher spectral efficiency, and higher network capacity. A detailed study on vehicular communications using the building blocks provided by 5G is provided in [102]. The study provides an overview of the adopted 5G architecture and its main characteristics for vehicular communications. Specifically, cellular communications are proposed for automatic vehicle location. Indeed, radio signals are used to activate a cellular device within a vehicle to compute and send back the vehicle's location.

2.3.4 City information sharing in VSNs

VSNs play an important role in providing comprehensive information services for the smart city. Information dissemination/sharing is one of the main benefits of vehicular networks that allows supporting a wide range of smart city applications by relying on the cooperation between vehicles as well as on that between the vehicles and the infrastructure. Hence, efficient cooperation mechanisms for city-wide information sharing are required before the widespread introduction of VSNs. For instance, to exchange information about surrounding events, the concept of encounter probability is proposed in [22] to decide on the feasibility of information re-diffusion. Relying on short interactions among vehicles, the protocol proposed in [22] provides information-sharing to numerous vehicular network services such as requesting parking slots, diffusing alert messages, etc. The performance of information sharing in VSNs was assessed through an analytical model that was presented in [55]. Precisely, the distribution of the concurrent transmissions was discussed, as well as numerical and simulation results were provided to confirm the precision of the model. In addition, other performance metrics have been proposed to characterize the performance of information sharing in VSNs namely, packet loss rate, packet transmission distance, end-to-end reliability, throughput, and delay. However, during information sharing, malignant vehicles may provide falsified information for attaining an unfair priority or to harm innocent users. In order to secure the information sharing process, a trust-management toolkit for identifying and mitigating the vehicles' compromised actions was conceived in [38]. The toolkit includes four modules for constructing a genuine information sharing mechanism, namely i) a monitoring module which detects misbehaviour and provides responses to cooperating vehicles, ii) The reputation-management module which evaluates the vehicle's integrity according to historical records, iii) The path-management module which is

in charge of redirecting the information dissemination relying on information provided by the reputation-management module, and iv) The trust-management module which includes a list of both genuine and malicious nodes. Relying on reinforcement learning, a mechanism guarantying sharing genuine information for all vehicles is designed in [109]. The authors invoke the outage probability to model the vehicles' future utility trend for predicting their future information sharing decisions. Additionally, each car has a reputation value based on its past behavior, which is also utilised to govern its future behavior, when sharing information.

2.4 Conclusion

In this chapter, we have covered some aspects of smart cities, wireless sensor networks, and vehicular sensor networks. After giving the definition and architecture of a smart city, we presented the WSN technology that is used to fuel smart city data centers with information. We then focused on vehicular sensor networks as an alternative way to gather environmental data from urban environment, and presented their architectures, challenges, application domains, and communication standards.

WSNs are crucial for collecting the data required by smart city services. Moreover, WSN architectures based on mobile elements (MEs) provide efficient data collection solutions that address the problems of traditional WSNs. In particular, in urban scenarios, vehicle mobility has been leveraged to design efficient and cost-effective data collection solutions. The following chapter (chapter 3) reviews the major existing schemes that are based on vehicular networks to collect sensors data in urban environments.

Chapter 3

Vehicle-Based Data Collection in Smart Cities: Related Work

3.1 Introduction

Smart cities are one of the important applications of IoT, which aim at improving the quality of life in big cities by providing efficient and reliable services. Indeed, many IoT devices are deployed in cities for traffic monitoring, resources management, security surveillance, etc. Moreover, future vehicles are being designed as mobile IoT devices that sense their local environment and communicate with cloud-based platforms. In such scenarios, data captured by IoT sensors have to be conveyed to data centers or central controllers for processing and storage. Therefore, designing efficient data gathering algorithms to successfully collect IoT sensors' data is paramount for realizing smart city applications.

Various schemes have been proposed for data gathering in VSNs with the aim of achieving efficient data gathering with low overhead. These schemes can be broadly classified into two categories based on the targeted sensing platform. According to Uichin *et al.* [69], vehicular sensing platforms are divided into two categories, namely i) V2V based platforms where only on-board vehicular sensors are employed and the vehicles are seen as distributed mobile servers, and ii) infrastructure based platforms where roadside sensors are also employed and the data captured by both vehicular and stationary sensors is to be conveyed to infrastructure-based servers such as Internet, data centers or central controllers.

In the sequel of this chapter, we focus on the data collection of infrastructure-based sensing platforms. Precisely, we investigate the schemes that utilize vehicles as mobile elements to collect data from roadside sensors. We first introduce the mobile element-based data collection approach that have proven its effectiveness in wireless sensor networks in terms of energy conservation and network connectivity relaxation [19, 47, 105]. Then we review existing vehicle-based data collection schemes for urban areas and provide a taxonomy.

Hence, the chapter is organized as follows: Section 3.2 illustrates the concept of data collection with mobile elements. Section 3.3 reviews related work about urban data collection with vehicular networks, followed by a discussion in Section 3.4. Finally, Section 3.5 concludes the chapter.

3.2 WSN Data Collection with Mobile Elements (MEs)

Several works in the field of WSN have proved the efficiency of mobile elements (MEs) applicability for sensors' data retrieval compared to multi-hop transfers to central controllers. Indeed, WSN mobility brings many advantages especially in terms of: (i) connectivity because MEs can cope with isolated regions so that network connectivity constraints can be relaxed and nodes redeployment can be reduced, (ii) cost since few nodes need to be deployed and mobile nodes can be obtained simply by exploiting existing MEs in the area of interest, such as buses, cars, motorcycles or pedestrians, (iii) reliability because MEs can visit sensor nodes individually and download their data directly through single-hop communications which allows avoiding collisions, (iv) energy efficiency as MEs can visit different regions in the network, the energy consumed is uniformly spread allowing to reduce the effect of premature energy depletion that can affect nodes closer to the sink. Nevertheless, mobility introduces serious challenges that do not exist in static WSNs and are outlined as follows.

- *Contact detection.* It is necessary to effectively detect the presence of MEs because communication between them and the sensor nodes is only possible when they are within the transmission range of each other.
- *Reliable data transfer.* Since MEs move during data transfer, MEs-sensors communications should be mobility aware and the number of messages correctly transmitted needs to be maximized.
- *Mobility control.* In some cases, the movement of MEs is either controllable (eg robots) or predictable (eg buses). In such cases, it is necessary to put in place adequate policies for visiting the sensor nodes in which the path, speed, and sojourn time of MEs have to be defined in order to improve the data collection performance.

Since the network architecture strongly depends on the role of the MEs, we will devote the following subsection to their description.

3.2.1 Mobile elements (MEs)

MEs are special support nodes that leverage mobility to support network operation or data collection by acting as intermediate data collectors or mobile gateways. In the

following we present the different types of MEs with increasing level of mobility as it was introduced in [32].

- *Relocatable nodes*. These are mobile nodes that change their location for connectivity or coverage purposes. Relocatable nodes do not carry data as they move in the network, but rather change network topology to enhance sensing coverage or network connectivity. They usually remain stationary after location change and forward data to the sink through multi-hop paths.
- *Mobile Data Collectors (MDCs)*. These MEs throughout the WSN and collect data from nearby stationary sensors. There are two types of MDCs depending on how the collected data is managed.
 - i. *Mobile Sinks (MSs)*. They represent the endpoints of data collection in WSN-MEs. The data collected is either consumed autonomously by MSs for their own purposes or it is made available to end-users by forwarding it to internet servers. Figure 3.1(a) depicts the MS-based WSN-ME architecture. The data transmission path between the sensors and the MSs is either a multi-hop that changes over time due to changes in the location of MSs such as in [110][93], or it is one hop such as in [10], where MSs are in direct contact with sensors.
 - ii. *Mobile Relays (MRs)*. These are support nodes that act as mobile forwarders. Their role is to collect data from sensors and to deliver it to base stations. For this purpose, the collected data is carried by MRs until they come into contact with the sink or the base station. Figure 3.1(b) depicts the MR-based WSN-ME architecture. A three-tier architecture has been proposed in [101][49] to collect data from WSNs where the middle layer consists of mobile relays called *Mobile Ubiquitous LAN Extensions (MULE)*. In a data-MULE system, a MULE collects data as it passes near the sensor nodes and delivers the collected data to the base station to be stored and made available to remote users.
- *Mobile peers*. These are ordinary mobile sensors that act as data generators or/and relays. During its movement in the sensing area, a mobile peer may receive data from neighboring peers and forwards it to others. When getting in the communication range of the base station, it uploads all the data it holds. A mobile peers-based WSN-ME architecture is depicted in Figure 3.2. Mobile peers have been widely used in wildlife monitoring, such as tracking zebras in the ZebraNet project [52] or whales in the SWIM system [46].

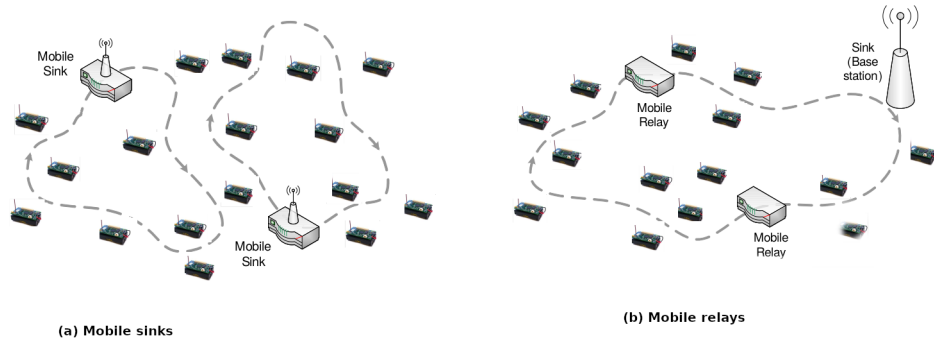


Figure 3.1: Architectures of WSN-MEs with MDCs.

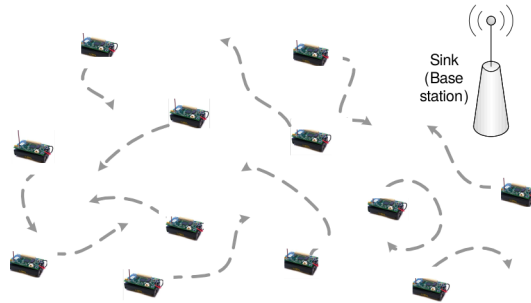


Figure 3.2: Architecture of a WSN-ME with mobile peers.

3.2.2 Overview of data collection in WSN-MEs

In WSN-MEs, the data collection process takes place in three main phases: *discovery*, *data transfer* and *routing to MEs* [32]. In the following, we will briefly describe each phase and highlight its main requirements and issues.

- **Discovery** which is defined as the process that allows a sensor node to detect the presence of an ME in its communication range. The aim of this phase is to detect sensor-ME contacts as soon as they occur, and with low energy expenditure. To this purpose, two complementary approaches are used: i) Mobility-independent discovery where MEs can be detected regardless of their mobility pattern, and ii) Knowledge-based power management where knowledge about the mobility model of the ME is leveraged to allow the sensor node to perform discovery only when the ME is likely to be in contact and sleep for the rest time. In this approach, sensors obtain the knowledge about the mobility model by observing the arrivals of the ME [53] [42] [33].

The mobility-independent discovery approach includes three different schemes: *Scheduled rendez-vous* schemes in which the sensors know exactly when the ME will enter the contact area¹, and can thus wake up at preset times. This happens

¹The contact area of a node is the region where that node can possibly be in contact with other

when MEs follow a strict schedule as in [24] where public transportation shuttles with a strict schedule have been proposed to perform the MEs role. Although scheduled rendez-vous schemes are energy efficient and simple to implement, they require sufficient accuracy in the mobility of the ME to meet schedules, which is difficult to hold in practice.

On-demand schemes assume that the ME can trigger the wake-up of the static sensors. Two main approaches have been proposed to achieve this. In the first approach, the nodes use two radios [100][112]: i) a long-range, high-power data radio, and ii) a low-range, low-power wake-up radio. The static sensor continuously monitors the related channel for activity. When entering the contact area, the ME sends a message to the related channel. Once an activity is detected on the channel, the static sensor powers up the data radio to initiate communication with the ME. In the second approach [45][14], the ME sends wake-up messages with sufficient energy in order to be used by the static sensors to enable the radio transceiver.

asynchronous schemes assume that the nodes can communicate without specifying the activation instant. To this end, periodic listening is commonly used [49][11][12]. Specifically, the ME periodically sends discovery messages, and the static node cyclically wakes up and listens for advertisements.

- **Data transfer** which is the process of exchanging data between the ME and neighboring static sensors. Because of the communication process is affected by the distance between the source and the receiver, data transfer protocols have to be aware of the issues resulted from mobility. For instance, Experiments carried out in [54] have shown that the amount of data collected by a slow moving robotized ME is in fact independent of the speed. However, the situation is much different when MEs move at higher speeds. Specifically, in [13] an urban scenario is considered with an ME that moves along a street with different speeds (such as 1 m/s for pedestrians, 20 and 40 km/h for vehicles and buses). This work shows that the message loss has a sort of parabolic trend while the ME is in the transmission range of the static sensor.
- **Routing to MEs** which is the process of data forwarding toward an ME. According to the ME type, routing protocols are divided into two main classes:
 - (i) routing protocols for uncontrollable ME, which have to adapt to the motion of the ME, and
 - (ii) routing protocols for controllable ME, which exploit motion control.

Routing techniques for uncontrollable MEs include two main classes, namely, *flat routing* and *proxy-based routing*. In *flat routing*, all nodes behave the same way,

while in *proxy-based routing* a number of sensor nodes are elected to take the role of proxies or gateways to bridge communications between the static sensors and the ME.

Several flat routing protocols have originated from solutions initially conceived for traditional WSNs. For instance, a modification of *Directed Diffusion (DD)* [48] addressing MEs is proposed in [54] where a higher priority is given to the interest propagation coming from MEs, in addition to using acknowledgment-based communication to deal with the time constraints of communication between static nodes and ME. Another example of extending a WSN routing protocol to WSN-ME is the *MobiRoute* [76] which is a modification of the *MintRoute* protocol [111]. As examples of routing protocols specifically devised for WSN-MEs we cite: EARM (*Energy-Aware Routing to Mobile gateway*) [6] which targets an ME that moves along a piece-wise linear path, and WEDAS (*Weighted Entropy Data dissemination*) protocol [9] which is based on an information-theoretic approach.

Concerning proxy-based routing, the proposed solutions are divided into two categories: solutions which are limited to dense WSN-MEs, and solutions which are (also) suitable to sparse WSN-MEs. TTDD (Two-Tier Data Dissemination) [75] is a proactive routing that builds an overlay grid-based forwarding structure to route data to MEs in a dense WSN, where nodes are assumed to be aware of their locations. SEAD (Scalable Energy-efficient Asynchronous Dissemination protocol) [56] follows a similar approach to build and maintain a dissemination tree for routing, and for data caching. SEAD is proved to perform better than Directed Diffusion and TTDD. An enhanced version of SEAD, called DEED (DElay-constrained minimum-Energy Dissemination), which aims to reduce the data collection latency is proposed in [57]. For relatively dense WSNs, HLETDR (Hybrid Learning-Enforced Time Domain Routing) [16] was introduced based on reinforcement learning where proxies keep track of ME arrivals and calculate the probability of the ME being in the vicinity, assuming that its mobility model distribution is Gaussian. A solution for moderately sparse WSNs has been investigated in [104]. The solution considers a set of disconnected clusters to be visited by an ME. shortest paths are established between a cluster head and cluster members by propagating beacon messages by the cluster head. Cluster heads are nodes which are one-hop neighbors of the ME.

3.3 Vehicle-based urban data collection schemes: overview

With the rise of IoT devices on city roads, vehicles are highly recommended to be used as MEs to deliver data from IoT sensors to data centers. Indeed, vehicles may incidentally cross the coverage area of active sensors, therefore, they can exchange data.

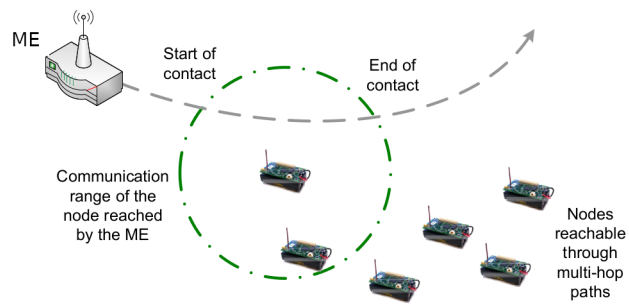


Figure 3.3: Reference scenario for data collection in WSN-MEs.

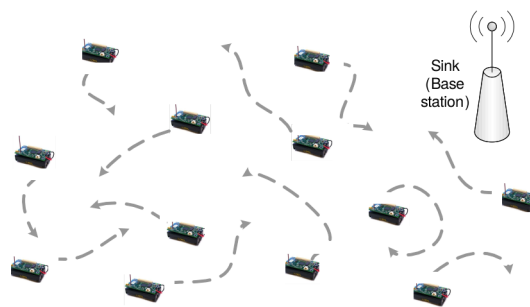


Figure 3.4: Architecture of a WSN-ME with mobile peers.

The benefits of this approach are multiple. First, it provides an alternative cost effective solution for data collection, as it leverages an already deployed cost-free infrastructure [98]. Second, it is energy-efficient as data offloading of IoT sensors can be performed using only one hop communication.

Data collection from urban areas using vehicles as MES has been firstly investigated by Chakrabarti et al. in [24], where the predictable mobility of buses was exploited to pull the data of sensors deployed on the buildings. In this work, the authors model the data collection process as a queuing system and the random deployment of sensors by random arrivals. The authors demonstrate through this modeling that relying on only single hop communications, allows significant savings in sensor network energy. To monitor isolated urban area, the authors in [60] propose a rendezvous based solution. The authors use a hierarchical WSN in which sensor nodes disseminate their data to a specific nodes (rendezvous nodes) through a cluster head. Vehicles collect the data from these specific nodes in near-periodic schedules. The rendezvous nodes are chosen so as to be close to the vehicles path and having sufficient residual energy.

In [67], the authors studied the feasibility of using public transport as data mules to collect information about road traffic collected by stationary sensors. The public transport vehicles then transmit the collected data to the traffic management center via static or mobile gateways. In [19], Borsetti et al. introduce the notion of *Virtual Data Mule* or VDM to deal with the mobility constraints of urban environments. VDM opportunistically moves the mule role between vehicles following a predefined trajectory. The mule role, recognized by a set of geographical coordinates, is assigned to a vehicle by a road side unit (RSU). If the vehicle is unable to ensure the mule role, it hands over the role to the most appropriate neighbor. A multi-hop data dissemination scheme using data replication is proposed in [71] to collect roadside sensors data. The data harvesting process relies on multi-hop communications between roadside sensors to respond to a request issued by a vehicle. The work in [18] suggests using taxicabs to connect the sensors deployed in smart cities to data centers. The authors use the concept of *oblivious data mules* as taxis can opportunistically exchange data with sensor nodes without having relation with smart city service provider or be required to follow any pre-established path. In [91], mobile sinks are used to collect data from sensor cameras. Mobiles sinks move towards relevant sensors according to dynamically computed optimal paths. To that end, relevance level of each sensor as well the urban mobility constraints are taken into consideration. A multi-tier data offloading protocol for urban IoT scenarios is devised in [62] by leveraging the fog networking paradigm. Some of the methods mentioned above are investigated in details in the following subsections.

3.3.1 MDH: Multi-hop Data Harvesting

MDH [71] is based on the V2I architecture where the sensors are deployed on the road-side with uniform distance from each other as illustrated by Figure 3.5. The data harvesting scheme relies on multi-hop communications between road-side sensors to respond to a request transmitted by a vehicle to a roadside sensor. Both roadside sensors and vehicles periodically broadcast HELLO messages to maintain updated location information of neighboring nodes. GPS (Global Positioning System) is used for acquiring current location coordinates. Two schemes have been proposed for collecting data namely, MDH-R (for Multi-hop Data Harvesting with Replicas) which is designed to be used when a vehicle transmits a request to a single sensor node and MDH-RA (for Multi-hop Data Harvesting with Replicas and Aggregation) which is designed for the case where a vehicle requires data from multiple sensors in a certain region.

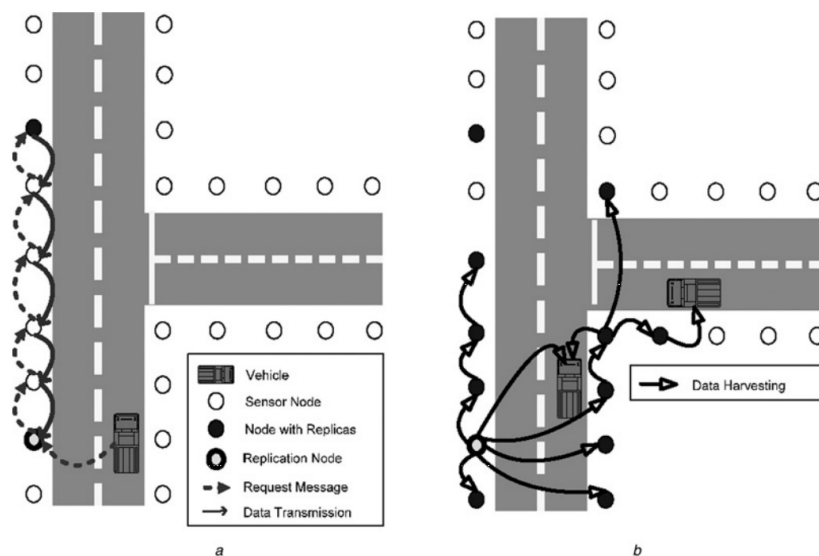


Figure 3.5: Multi-hop data harvesting scheme with replicas.

In MDH-R, a requesting vehicle transmits a uni-cast control packet to the closest sensor node. The request control packet includes the vehicle's velocity which is required later for creating replicas in the network. The sensor that receives the request message declares itself as "*replication node*" and inserts its ID as well as its coordinates in the request message then transmits the modified request to the closest sensor to the destination. The packet will be forwarded until it will be received by the destination node. The latter will transmit its data packets towards the "*replication node*" using multi-hop transmissions as it is illustrated by Figure 3.5 (a). When the "*replication node*" receives the required data, it will flood the data packets to its neighbours (Figure 3.5 (b)), creating replicas of the data. The data flooding process is controlled by the

time-to-live (TTL) value of the broadcast-ed replica packet, which is calculated using the following formula (3.1).

$$TTL = \frac{v * t}{d} + \frac{v * \frac{p}{r}}{d} \quad (3.1)$$

Where,

v is the speed of the vehicle requesting the data. t is the overall time taken by the data to reach the replication node after being requested. p is the size of the data, r is the transmission rate of the sensor, and d is the distance separating two consecutive sensor nodes.

With the TTL formula above, more replications will be performed if the vehicle becomes faster and the transmission rate of the sensor becomes slower. Each sensor node maintains an ordered list of the received packets. In case of packet collisions, sensors can identify lost packets from this list. The vehicle may harvest data packets from any of the sensor nodes holding a replica of the data.

As earlier mentioned, a second scheme named MDH-RA is designed in the same work [71] to harvest data from multiple sensors in a certain region. In this scheme, a vehicle transmits a geo-cast request to a group of sensors using geographical information. The request message is handled and forwarded by "replication node" and intermediate nodes as performed in the MDH-R's request transmission process until it reaches the first sensor of the geo-cast region, which will announce itself as the "dynamic cluster head". The latter forwards the request to all the sensors of the geo-cast region as it is depicted in Figure 3.6 (a), while announcing to them that it is the cluster head. Upon receiving the request message, the sensor nodes transmit their data to the cluster head (Figure 3.6 (b)) which will aggregate them and transmit the aggregated data back to the "replication node", which will proceed to their replication as it was described in MDH-R.

3.3.2 VDM: Virtual Data Mules for data collection in road-side sensor networks

In traditional wireless sensor networks, the data mule role is permanently assigned to one or more specific mobile nodes that move according to predefined or controllable trajectories. This assumption often does not hold in urban areas, because the most suitable elements for collecting data from such an environment are vehicles or pedestrians, who move only according to their will and along trajectories that are not necessarily suitable for the data ferry. To overcome this issue, the concept of *virtual data mule* (VDM) [19] has been introduced in which the data mule function is not strictly associated to a specific mobile node, but can be handed from one node to another. In VDM, vehicles are used as data mules for collecting data from urban environments. To collect data from road-side sensors, a VDM role, recognized by a set of geographical

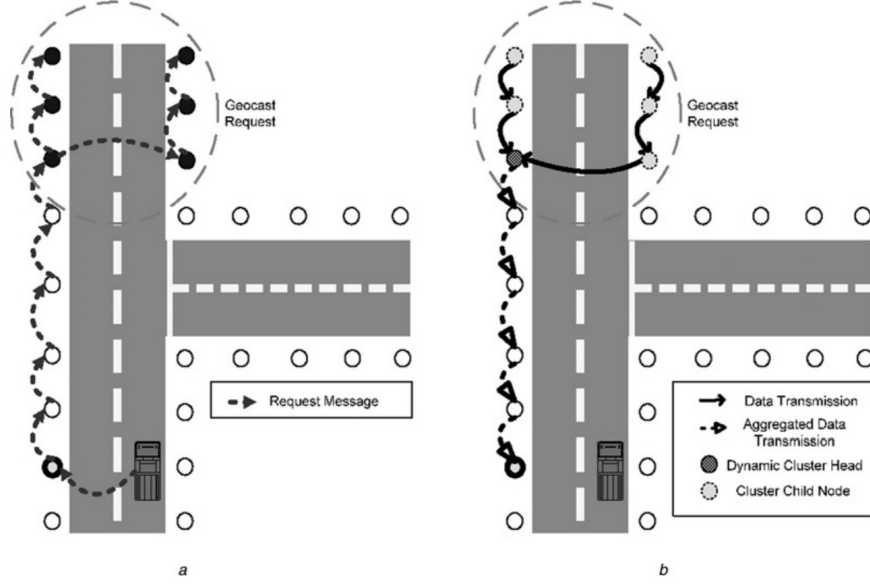


Figure 3.6: Multi-hop data harvesting scheme with replicas and aggregation.

coordinates, is assigned to a vehicle by a road side unit (RSU), as it is depicted in Figure 3.7. If the vehicle is unable to ensure this role, it hands over the role to the most appropriate candidate among its neighbors as follows. The vehicle acting as data mule periodically checks its fitness by computing an index value that measures how close the vehicle’s route approaches the planned VDM trajectory. If the index value is below a minimum threshold, the current data mule initiates an election process by advertising the mule role, the current target point, and the VDM desired speed. Neighboring vehicles calculate their indices and send them back to the current data mule, which hands over the VDM role to the vehicle with the highest index. The index is formulated so that the candidate follows as closely as possible the desired path while keeping low the control overhead (number of handovers). Two mechanisms are proposed to compute such an index. The first one is based on motion geometry and the second on fuzzy logic. Particularly, the fuzzy logic-based index is given by the following formula (3.2).

$$i_{FL} = DF + 1_{[DF=0]}(2 - |SF| + |TF|) \quad (3.2)$$

Where,

DF is the **distance factor** which measures the distance between the vehicle orthogonal projection on the ideal trajectory and the current desired VDM position. It takes values in $[-2, 0]$. SF is the **steer factor** which takes values in $[-1, 1]$. More positive values of SF are representative of car headings closed to that of the ideal VDM trajectory. TF is the **throttle factor** which is determined by employing the vehicle speed and acceleration. Its values range in $[-1, 1]$, where positive values map onto speed/acceleration pairs that are closer to those needed to follow the movement of the VDM over time.

$1_{[DF=0]}$ is the indicator function that takes the unitary value when $DF = 0$, and it is zero otherwise. When the VDM reaches the current target, the next target in the

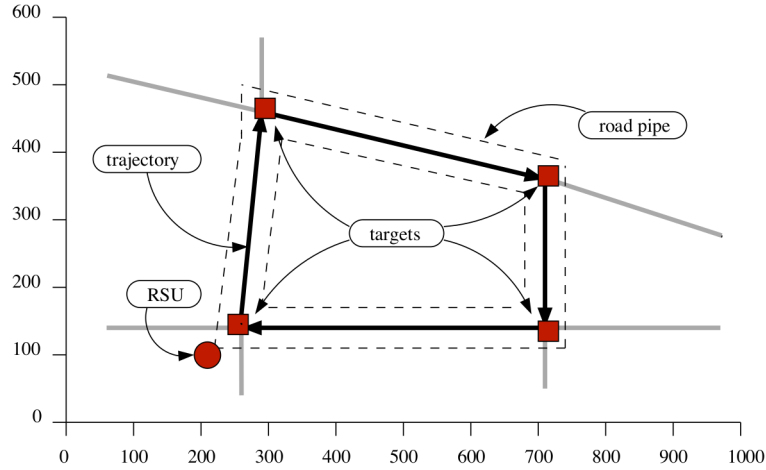


Figure 3.7: Data collection in an urban environment along a desired path with VDM.

geographical coordinates sequence is picked as the new VDM target.

Although VDM adapts the data mule approach to the characteristics of urban environment, the fact that the collection task has to be performed according to a strict sequence of geographical coordinates leads to important data delivery latency.

3.3.3 MobiClutser: A Rendezvous-Based Approach for Data Collection with Mobile Sinks

Mobile sinks-based solutions have been excessively explored for the energy hole problem in WSNs. The use of Mobile-sinks (MSs) provides considerable energy saving and extends the network lifetime. However, it introduces different degrees of data acquisition latency depending on the path followed by MSs. To mitigate this issue, rendezvous-based data acquisition methods, which aim to provide a trade-off between energy consumption and data acquisition latency, are proposed. The methods are grouped into two categories: rendezvous point (RP) based and rendezvous area (RA) based [81].

MobiCluster [61] is an RP-based method that targets monitoring of isolated urban areas with respect to environmental parameters, surveillance, fire detection, etc. In MobiCluster, a number of Sensor Nodes (SNs) located at the periphery of the sensor field are designated as ‘rendezvous’ points to collect sensory data from neighboring SNs and to deliver them to a MS when the latter approaches within radio range. The authors argue that city buses are the adequate infrastructure for sensory data collection since such vehicles repeatedly follow a predefined trajectory with a periodic schedule and are likely to approach the perimeter of the isolated sensor fields.

Aiming at minimizing the network overhead and energy expenditure of data retrieval, MobiCluster builds clusters of member SNs that route their data to their as-

signed Cluster Head (CH). The CHs filter the received data using potential spatial-temporal data redundancy and forward the filtered information to their assigned RNs (Rendezvous Nodes), typically located in proximity to the MS's trajectory. Data of source SNs that are located far from the MS trajectories are delivered through multi-hop transfers to RNs. To that end, hierarchical cluster structures comprising neighbor SNs are built to increase the performance of intra-cluster data filtering and to minimize the data relaying overhead as illustrated by Figure 3.8. In Mobi-Cluster, data collec-

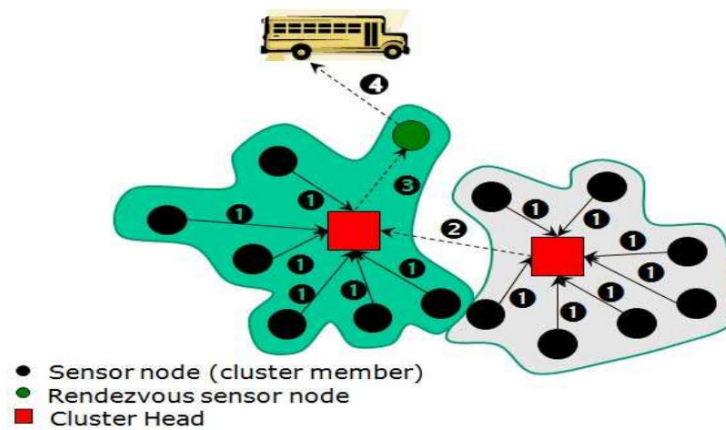


Figure 3.8: Data collection and forwarding phases in MobiCluster.

tion takes place in two main phases: the setup phase and the steady phase. The setup phase includes three phases summarized as follows.

1. **Clustering:** where cluster structures are formed within a single iteration and CHs are elected on the basis of their residual energy supply so that to be reachable in a single hop from their cluster members. CHs collect data from their attached cluster members and forward them towards the RNs following an inter-cluster overlay graph.
2. **RNs selection:** RNs are SNs that lie within the range of travelling sinks to guarantee connectivity of sensor islands with MSs. The best candidates for acting as RNs are the SNs with sufficient residual energy and remain long within the sink's range.
3. **CHs attachment to RNs:** The aim of this phase is to attach each CH located far from the MS's trajectory to an RN so as to address to it the clusters' data.

The steady-state of MobiCluster passes by two phases namely:

1. **Data aggregation and forwarding to the RNs:**

In this phase, each SN sends its data to the local CH which will process them

to remove redundancy and forwards the filtered data towards the cluster where their attached RN belongs to.

2. Communication between RNs and mobile sinks:

The last phase of Mobicluster is to deliver the data buffered in RNs to the MSs. To that end, an acknowledgment-based protocol is used to allow RNs to transmit data only when the connection with MSs is available. RNs detect an MS due to a POLL message that is periodically broadcasted by MSs to announce their presence to RNs.

3.3.4 DATAxis

DATAxis [18] addresses the issue of collecting data from sensors and actuators of IoT objects spread across smart cities' streets. Unlike MobiClutser which relies on public transportation vehicles with predefined trajectories, DATAxis leverages transport fleets that are not required to follow pre-established or optimized paths, namely taxi cabs. In fact, the taxi service covers an extensive area of the city with fairly random routes and operates 24/7. The authors use the concept of *oblivious data mules* as taxis can opportunistically exchange data with sensor nodes without having relation with smart city service providers or be required to follow any planned path. In addition, taxis provide better coverage for densely populated city areas which are usually the places where more data needs to be collected and transmitted. Moreover, many taxi fleets have already almost all the hardware on-board to ensure the data mule function. As depicted in Figure 3.9, while the taxi cab moves around in the city according to customers' request, it incidentally crosses the coverage area of active nodes with which it can exchange data.



Figure 3.9: Dataxis proposed scenario: Opportunistic data exchange between taxi cabs and IoT nodes.

To evaluate the feasibility and the performance of this idea, the authors consider a scenario where a small fleet of taxi cabs moves inside a city doing their job. The taxi

cabs are equipped with an on-board terminal which could be an ordinary Tablet or a Smartphone that integrates different wireless technologies to communicate with the surrounding environment. Moreover, the authors assume that special areas in the city (for example taxis' parking lots) act as hot-spots that provide internet access. Sensors and actuators are assumed to be severely energy constrained i.e. they are powered via energy harvesting solutions (photo-voltaic, piezoelectric, etc) and are plugged on top of other existing solutions to enhance existing services. Dataxis considers services that require communicating small amounts of data and tolerate significant delays such as smart waste collection and street lighting control. The performance of this scenario is assessed through extensive simulations on data gathering and data dissemination from IoT nodes to oblivious data mules and vice-versa. The simulation results show that with this scenario, a delay of one day is required to collect more than half of the data and two days to gather more than 90%.

3.3.5 WWSN for smart city applications: a Relevance-based approach for multiple sinks mobility

In this work[91], an adaptive mobile sinks-based data collection solution is designed for surveillance applications in a smart city. The work provides a solution for crimes detection. To that end, the authors assumes a wireless visual sensor network deployed along the roads and streets of a smart city and provides an adaptive sinks position algorithm that takes into consideration the dynamism of sensing relevance of each visual sensor. To that end, a Relevance Level (RL, expressed as a priority index) is assigned to every visual sensor node. According to the authors, the RL of each sensor can be determined by the application, the users, or by a specific trigger and can vary along the time from 0 (irrelevant) to 15 (highly relevant). Indeed, The algorithm takes into account the particularities of every part of a city because depending on the characteristics of the area to be monitored, some of the retrieved information may be more relevant for the monitoring system. For example, the system should consider cameras of neighborhoods with high crime rates during the day as more important, while lowering their relevance level during the night in favor of the more dangerous neighborhoods.

The authors' objective in this work is to position sinks closer to sensors with higher sensing relevance since they are expected to transmit more data packets. To that end, an algorithm is designed to dynamically compute the optimal positions for sinks and the optimal route toward the computed position for each sink. The algorithm creates sectors by associating each sensor node to the nearest sink. To prioritize important sensors, the RL of sensors inside each sector are used to calculate the position of sinks. The authors divide the multiple sink positioning problem into different steps which resulted

in three algorithms. Initially, sinks are assumed to be positioned in an uniform grid-like topology within the network area. In the first step, each sink “virtually” moves to the nearest sensor node, while avoiding that several sinks will occupy the same position. The second step consists in creating logical clusters by associating each sensor to the nearest sink. Finally, the relevance of sensors inside each cluster ($RL(k_i)$) are used to balance sinks positions. This is done by virtually moving each sink to more relevant clusters as described by the following equation (3.3).

$$k_{i(x,y)} = k_{i(x,y)} + (l_{j(x,y)} - k_{i(x,y)}) \cdot \frac{RL(l_j)}{RL(S)}, \forall k_i \in K, \forall l_j \in K - k_i \quad (3.3)$$

Where,

$S = \{s_0, s_1, \dots, s_n\}$ is the set of sensors and $K = \{k_0, k_1, \dots, k_n\}$ is the set of mobile sinks. $k_{i(x,y)}$ and $l_{j(x,y)}$ are the (x, y) coordinates of sinks k_i and l_j respectively. $RL(l_j)$ is the sum of relevance level of sensors of l_j and $RL(S)$ is the sum of relevance levels of sensors of the whole network.

After the balancing step, an update step is performed to re-associate the sensor nodes to the nearest sink, resulting in cluster changes. The update step is repeated until convergence (when sinks’ locations are no longer changed). Finally, sinks are notified to move to the their new expected locations (using any application-specific protocol). The last step of this solution is to define a path for each sink to reach its new position. For that, a graph is created and Dijkstra’s algorithm [34] is executed to determine the best route for each sink. The vertices of the created graph represent the set of sinks’ current and final positions, as well as road intersections, while edges represent the segments of the road network as depicted by Figure 3.10.

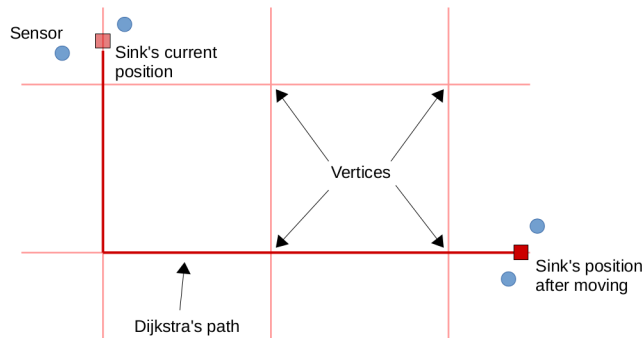


Figure 3.10: Movement path of a sink.

3.3.6 FDO: Fog-based Data Offloading in Urban IoT Scenarios

FDO[62] is a multi-tier data offloading protocol that is devised for urban IoT scenarios by leveraging the fog networking paradigm². Specifically, it takes advantage of heterogeneity in the network so that sensors can collaboratively offload data to each other or to mobile gateways (cars or people carrying smartphones). Indeed, the authors argue that the uncontrolled and unpredictable mobility of gateways in urban IoT scenarios induces significant dynamic heterogeneity between different sensor environments, requiring sensors to adapt to changes in gateway mobility. The work aims to optimize communications in fog networks where devices have limited buffering capability. Figure 3.11 illustrates the considered fog network architecture.

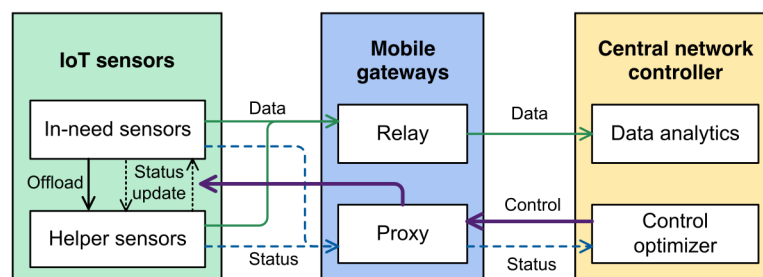


Figure 3.11: Interactions in the reference three-tier fog network.

IoT sensors periodically collect data from the environment and store it in a local buffer. Based on their buffer occupancy and communication opportunities, IoT sensors forward the data to other sensors or to mobile gateways in the urban environment which may be vehicles. The mobile gateways are assumed to have two radio transceivers, one for exchanging data with the IoT sensors and another for Internet connection in order to forward the data to the central controller (cloud) where data will be analyzed and stored. The controller also performs network management by dynamically re-configuring the devices in the lower tiers. Based on the buffer availability and the probability of contacting a mobile gateway, IoT sensors are logically divided into two categories, namely:

- *In-need sensors* which are rarely in reach of a mobile gateway, resulting in a high data drop-off rate due to unavailable buffer space and,
- *Helper sensors* which frequently encounter mobile gateways and can temporarily store the data of *in-need sensors* in their own buffers.

This partitioning is dynamically performed by the central controller on the basis of status information (including contact opportunities) received from the mobile gateways.

²Fog computing describes a decentralized computing structure located between the cloud and end-user devices. Fog responds to the specific challenges of IoT, where multiple edge and terminal devices collaboratively perform a large amount of computing, storage, communication and management.

When an *in-need sensor*'s buffer is full, the sensor broadcasts the least recent amount δd of data to its neighboring helpers and removes the data from its buffer. Since all *in-need sensors* are independent, the probability that an *in-need sensor* j successfully offloads its data to a *helper sensor* k is given by:

$$Pr(s_j \rightarrow s_k) = (1 - p_k) Pr(b_k \leq (L - \delta)d) \sum_{n=0}^{|I_k \setminus s_j| - 1} \sum_{[I_k \setminus s_j]^n \subset I_k \setminus s_j} \frac{1}{n + 1} \prod_{s_m \in [I_k \setminus s_j]^n} Pr(b_m = Ld) \prod_{s_i \in I_k \setminus s_j - [I_k \setminus s_j]^n} (1 - Pr(b_i = Ld)) \quad (3.4)$$

Where,

p_k is the probability that the sensor K meets any of the N gateways. b_k is the buffer occupancy which is defined as the amount of data stored by the sensor K . L is the maximum number of data items that fit in the buffer. d is the size of individual data items sampled in a time slot. $|I_k \setminus s_j|$ and $[I_k \setminus s_j]^n$ are the cardinality and n -subset of $I_k \setminus s_j$, respectively. I_k is the subset of in-need sensors that are neighbors of the helper sensor k .

When a gateway is in the range of a sensor m , the sensor first updates its observed probability p_m of encountering a gateway then transfers it to the gateway with the buffered data. The gateway replies with the updated values of the operating parameters, ϵ and δ . With, ϵ is the threshold of data drop-off rate to distinguish between helper and in-need sensors and δ is the amount of data items offloaded from an in-need sensor to a neighboring helper sensor at once. If no gateway is in range, the sensor similarly updates its observed probability p_m and derives its class (i.e., in-need or helper) based on its last known value of the parameter.

3.4 Discussion

As mentioned earlier, due to its effectiveness in energy conservation and network connectivity relaxation, the ME-based data collection approach has been widely recommended for urban environments. Therefore, several solutions have been proposed to leverage the vehicular networks to collect data from roadside IoT objects in a smart city. Table 3.1 summarizes the main characteristics of the vehicle-based data collection schemes studied in this chapter. We classified the schemes according to the type of the targeted applications, the data collection type, and the type of mobile collectors. Throughout our investigation, we also attempted to answer the following two questions:

- i) Does the scheme employ vehicles exclusively dedicated for data collection?
- ii) Is the reduction of data collection latency considered by the scheme?

Apparently, all the studied data collection schemes are designed for delay tolerant applications except for MDH-R which is dedicated for short delay applications as it

provides an on demand data collection. The remaining schemes provide either: i) a periodic data collection such as VDM[19] and RL-based [91] where the period is set by a defined entity such as an RSU or a data center, or ii) a near periodic collection where the schemes rely on vehicles with near periodic predefined trajectory such as buses which is the case of MobiCluster [61]. The case of DATAxis is special because taxis are available 24 hours / 24 hours and 7 days/7 days leading to an opportunistically continuous data collection. Another important criterion related to the cost of implementation is the dependence of the scheme on specific vehicles exclusively devoted to data collection. According to our study only RL-based [91] requires specific vehicles to perform data collection which will increase the scheme implementation cost. Finally, only MobiCluster [61] and RL-based [91] have evoked the point of reducing the data collection latency in their design.

We then classifies the studied schemes according to: the mobile collector trajectory type (planned or not), the number of road segments covered by the mobile collector during its collection function (coverage), the number of hops for data transfer to the mobile collector, and the latency of the overall data collection process. The means of measuring the performance of each scheme are considered, and the pros and cons of each are also included. Table 3.2 summarizes our classification based on the aforementioned criteria.

The trajectory of the mobile collector (MC) is either planned or unplanned. According to the studied schemes, the type of the vehicles used as MCs as well as the type of data collection controls the type of the MCs' trajectory. For instance, the trajectory of the MC in MDH-R and DATAxis is not planned because in MDH-R the trajectory depends on the location of the vehicle that requests the data collection, and in DATAxis it depends on the taxi (client) trip. The other methods use planned trajectories as they either rely on special vehicles exclusively dedicated for data collection such as RL-based, or on buses that travel according to a near periodic trajectory such as MobiCluster. Also, a data collection scheme may at anytime elect vehicles from the urban area to follow a predefined trajectory which is the case of VDM. With regard to FDO, it focuses on the data drop-off problem by considering only the moment of the MC passage without giving any indication of its trajectory.

Most of the studied schemes utilize trajectories of several road segments where the MC has to follow in a strict sequence, which leads to high data collection latency except for MDH-R who's MC coverage depends on the location of the requested data. In VDM and DATAxis, sensors data are directly offloaded to the MC using one-hop communications whereas, in the remaining schemes, sensors data traverse a number of intermediate sensor nodes to reach the MC which makes these schemes energy costly compared to one-hop-based schemes.

In summary, we can notice that the energy consumption is reduced in single-hop-

Table 3.1: Key characteristics of the studied vehicular-based data collection schemes for urban environments

Scheme	Application type	Collection type	Vehicles type	Requires exclusive vehicles	Considers latency reduction
MDH [71]	Short delay	On demand	Any	No	No
VDM [19]	DTN ¹	periodic	Any	No	No
MobiCluster [61]	DTN	Periodic	Public busses	No	Yes
DATAxis [18]	DTN	Opportunistic	Taxi cabs	No	No
RL-based [91]	DTN	Periodic	Any	Yes	Yes
FDO [62]	DTN	N/A	Any	No	No

¹ Delay Tolerant

based schemes, however, at the expense of high data delivery delay, as MCs are constrained to visit sensors individually, increasing MC’s trajectory, hence the delivery delay of data. In multi-hop schemes, the delay is reduced but the power consumption is quite high due to the additional communications between the sensors. Moreover, relying on specific vehicles to collect roadside sensors data increases the scheme’s deployment cost, whereas leveraging the ubiquitous vehicular network provides a cost-effective alternative, because it relies on an already deployed cost-free infrastructure. It would therefore be interesting to combine the advantages of the studied schemes to design energy and cost-efficient data collection schemes with a reduced data delivery delay. This can be achieved by taking advantage of the vehicular network to designate opportunistic collectors to collect data from the sensors via a single hop by acting only on the road segment they are traveling on.

Table 3.2: Comparative study of the studied vehicle-based data collection schemes for urban environments.

Scheme	ME trajectory	ME coverage	Number of hops to ME	Latency	Pros	Cons	Performance measurement
MDH [71]	Not planned	Depends on the location of the requested data	Multi-hop	Short	Data thanks to replication	-The neighbors of the replication node are prone to energy depletion. -The TTL depends on the vehicle speed which is variable	Simulation with QualNet 4.0
VDM [19]	Planned	Several road segments per a DM	One-hop	High	-No additional deployment cost (opportunistic ME) -Well adapted to urban mobility constraints	-Additional communication cost due to multiple handovers -Not suitable for short delay applications	Simulation using NS2 and Matlab fuzzy logic

Table 3.2: Comparative study of the studied vehicle-based data collection schemes for urban environments (continued)

Scheme	ME trajectory	ME coverage	Number of hops to ME	Latency	Pros	Cons	Performance measurement
MobiCluster [61]	Planned	Several road segments per a DM	Multi-hop	Reduced	-Reduced data delivery delay due to the reduction in the number of nodes that the ME has to pass through -No additional deployment cost (use of buses)	-Specific sensor (isolated areas) and cannot be adapted for uniformly deployed roadside IoT objects -Energy consumption relatively high compared to one-hop-based schemes	Simulation (simulator N/A)
DATAxis [18]	Not planned	Depends on cabs trip	one hop	High (> 24 hours)	-No additional deployment cost (oblivious data mules) -DMs Availability (7/7, 24h/ 24h)	Not suitable for short delay	Simulation using real traces (simulator N/A)

Table 3.2: Comparative study of the studied vehicle-based data collection schemes for urban environments (continued)

Scheme	ME trajectory	ME coverage	Number of hops to ME	Latency	Pros	Cons	Performance measurement
RL-based [91]	Planned (adaptive planning)	Several roads (Dynamic optimized route)	Multi-hop	High	Data of relevant zones are collected first	-Additional deployment cost (relies on dedicated vehicles) -The need to travel through multiple segments to reach the desired sensors	Simulation (simulator N/A)
						Not suitable for short delay applications	

Table 3.2: Comparative study of the studied vehicle-based data collection schemes for urban environments (continued)

Scheme	ME trajectory	ME coverage	Number of hops to ME	Latency	Pros	Cons	Performance measurement
FDO [62]	N/A	N/A	Multi-hop	N/A	<ul style="list-style-type: none"> -Collaborative data offloading -Decreased data drop-off rate -Handling sensors heterogeneity (Fog networking based) 	<ul style="list-style-type: none"> -May increase power consumption due to increased communication between sensors -Helper sensors are prone to energy depletion since they are elected without taking care of their residual energy 	Analytical model and simulation with a custom Python simulator.

3.5 Conclusion

In this chapter, we have provided a brief overview of mobile element-based data collection while highlighting the main challenges and issues at each stage of the data collection process. Then we conducted a comprehensive study on solutions using vehicles as MEs to collect IoT data from urban environments. We also provided a comparative study to investigate and evaluate existing vehicle-based data collection schemes.

Finally, we can conclude that vehicular networks are suitable for collecting data from IoT sensors in smart cities. However, the high dynamism of urban environments, combined with the vehicle's itinerary length, result in high data collection latency which limits vehicle-based data collection schemes to only DTN applications. To take advantage of this cost-effective and energy-efficient approach on behalf of short-delay applications, this thesis proposes two schemes for low-latency data collection. The next chapter is devoted to the description of the first scheme which we called, OMS (Opportunistic Mules for Short Latency Data Collection in Smart Cities).

Chapter 4

OMS: Opportunistic Mules for Short Latency Data Collection in Smart Cities.

4.1 Introduction

As shown in the previous chapter, introducing mobile elements (MEs) has proven to be efficient and economical for WSNs data collection. However, this mobile data gathering unavoidably introduces longer data collection latency which has amplified in urban settings due to urban mobility constraints. Moreover, the length of the trajectories assigned to the vehicles used as MEs contributes in increasing the data collection process latency which reaches 24 hours such as in [18]. Given that, some smart city applications are required to respond in a relatively lower latency. So reducing the latency of such a promising approach becomes crucial. To that end, we propose a new scheme that leverages opportunistic inter-vehicular communications to provide each road segment with a Data Mule (DM) in order to limit the DM task to only one segment. By doing so, multiple DMs will operate simultaneously within different segments of an area of interest leading to collection latency reduction. Following that idea, an opportunistic DMs election strategy was proposed in which the ability of a vehicle to serve as a DM on a given road is determined by a fitness function that is based on the vehicle's movement characteristics. The proposed scheme is called, OMS (Opportunistic Mules for Short Latency Data Collection in Smart Cities). A preliminary version of OMS was presented in [?] and has been revised and improved in [26].

In this chapter, first, we give an overview of our proposal in Section 4.3. Then, the detailed description and analysis of the proposed scheme are provided in Section 4.4. The performance evaluation of our scheme is provided in section 4.5. Finally, Section 4.6 concludes the chapter.

4.2 Motivation

In a smart city application, it is required that the sensed data be conveyed to Central Controllers (CCs) for processing and sharing. As mentioned before, data mulling is an efficient solution to deal with issues related to communication in WSNs such as connectivity, reliability, and energy efficiency. However, in data mule-based works targeting urban environments, mobility and communication capabilities are often underutilized, resulting in sub-optimal solutions that incur large latency. Motivated by the need to meet short delay applications while taking advantage of the MEs-based data collection approach, we focused on reducing the latency of the mule-based data collection in urban environments. To that end, we take benefit from an important characteristic of the urban environment, namely its richness in opportunistic mobile elements (vehicles). We argue that the distribution of the collection process will reduce its latency. To do so, we propose to involve all kinds of vehicles to elect a data mule for each road segment to take the responsibility of collecting its sensor data. Therefore, the mule role will be held along only one road segment instead of long itinerary such as in [19, 18, 91]. Furthermore, unlike [18] that rely only on taxicabs, our scheme leverages all vehicle kinds to increase chances of getting DMs. Doing so; we fully exploit the opportunistic presence of vehicles to distribute at best the collection process and therefore improve the collection latency.

4.3 Overview

We focus on data gathering from sensors embedded in objects which can be spread across smart city roads. We address the case where the collection process is periodically initiated by a fixed base station (BS) or central controller (CC) within a specific region called Area of Interest (AoI). We consider short delay applications in which data have to be delivered to the CC in relatively small delays. The considered scenario is depicted in Figure 4.1 (a), where sensors are deployed on roads or integrated in roadside IoT objects such as road signs, garbage cans, etc. The idea of this proposal is to leverage the characteristic of the high presence of vehicles in urban areas to increase the distribution degree of the data collection process hence, reducing its latency. To that end, we propose to assign, as much as possible, to each road segment of an AoI a DM responsible of collecting its sensory data. Hence, two important issues should be solved, namely: data mule assignment and the collected data delivery from each segment to the CC all while having short latency. To deal with these issues, OMS covers the AoI with a request for data mules election providing each road segment with a DM. To achieve this, intersections are preferred data mule election zones (DMEAs) since they constitute incoming/outgoing points for each road segment. After being elected, each DM downloads the sensors' data of the segment for which it is responsible, then

forwards the collected data to the BS or CC using a delay optimal vehicular routing protocol. Figure 4.1 (b) gives an overview of the proposed scheme, while Table 4.1 lists the main abbreviations used in this chapter.

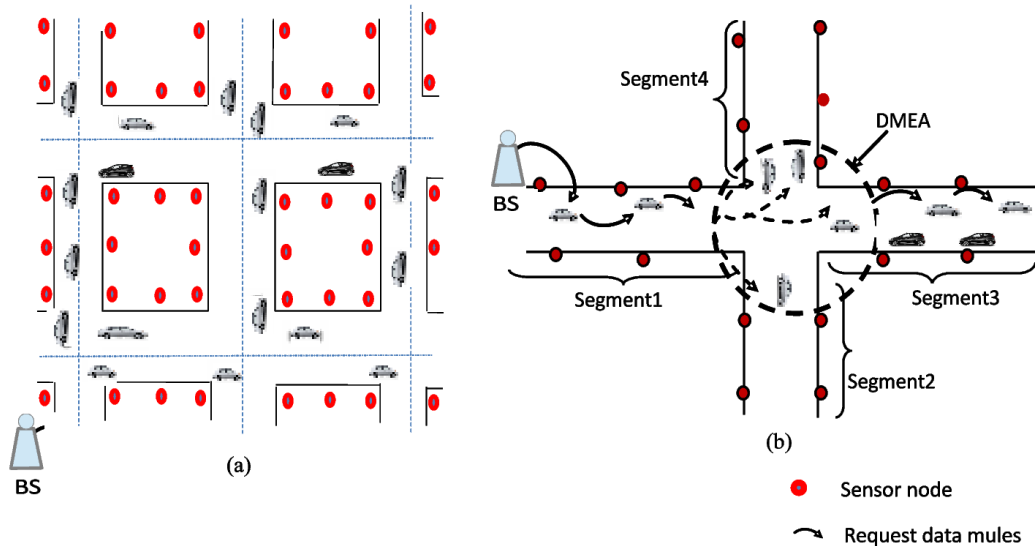


Figure 4.1: (a) Considered scenario. (b) Scheme overview.

Table 4.1: List of abbreviations.

Abbreviation	Description
ME	Mobile Element
DM	Data Mule
AoI	Area of Interest
DMEA	Data Mule Election Area
BS	Base Station
CC	Central Controller
REQ_DM	Request for Data Mules

4.4 Opportunistic Mules for Short Delay Applications (OMS)

This section describes OMS. Our data mule-based data collection scheme, whose main goal is to reduce the latency of data collection processes in urban environments. OMS relies on multiple data mules opportunistically elected from moving vehicles to provide simultaneous collection of different road segments of an urban area. The basic idea of OMS is to abolish the main cause of high delays in DM-based collection, namely the length of trajectories assigned to DMs. In fact, OMS aims to limit the mule role of a

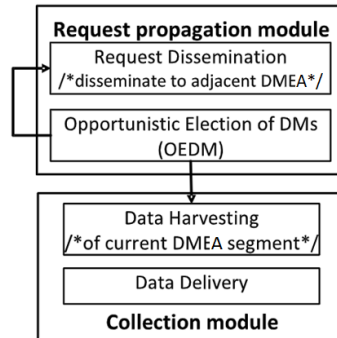


Figure 4.2: Opportunistic Mules for Short Latency Data Collection: Scheme Architecture.

vehicle to one road segment. To do so, it takes advantage of the large population of moving vehicles to opportunistically designate a DM to each road segment by propagating a request for DMs to DMEAs within an optimum delay. The proposed scheme includes two modules as it is shown in Figure 4.2:

1. **The request propagation:** in this module, the goal is to disseminate a request for DMs to all DMEAs of the AoI to opportunistically elect a DM for each outgoing segment. Thus, two components are designed. The first consists of the query dissemination and the second is the opportunistic DMs election.
2. **The data collection module:** in this module, each elected DM collects the sensors data of its road segment and forwards them to the BS or CC through a delay optimal vehicular routing algorithm. This requires the design of two sub-modules namely: the data harvesting and the data delivery.

The election process, as well as the data harvesting and data delivery processes take place in parallel within different road segments according to the propagation progress of the request for DMs. However, they perform sequentially within the same segment. Note that in this thesis we have focused on the design of query dissemination and opportunistic election processes. Data harvesting and data delivery processes are planned as perspectives.

4.4.1 Query Dissemination

Initially, data retrieval of an AoI is periodically launched by a CC which sends a request for data mules (REQ_DM) message that contains a unique ID, the position of the CC, and other necessary information to be detailed below. The REQ_DM should then be propagated to DMEAs of the AoI where each outgoing segment has to be assigned a DM according to the speed and direction of existing vehicles. To preserve sensors energy, our scheme leverages the opportunistic inter-vehicular communications to propagate the

REQ_DMs to DMEAs. However, planning the REQ_DMs propagation is mandatory to avoid buckles and failures to reach some DMEAs which are likely to happen in case of unplanned propagation. Moreover, the propagation process has to be time constrained in order to reduce collection latency. In fact, reaching the entire DMEAs in a minimum delay induces rapidly assigning DMs to all the roads in the AoI. Hence, this allows data collection simultaneously over all roads. We consider the issue of reaching all DMEAs of an AoI in a minimum delay as a variant of Minimum Spanning Tree (MST) problem [44]. MST problem is an optimization operation that aims to span over all the vertices of a weighted graph with a minimum cost and without any cycle. MST has direct applications in the design of networks of different domains such as telecommunication and transportation. In WSN, MST has been widely used to construct energy efficient routing and topology control algorithms [70, 15]. In our case, we use the MST concept to propagate the REQ_DMs message to all DMEAs within a minimum delay. To that end, the AoI has to be modeled using an undirected graph $G = (V, E)$, where vertices V correspond to DMEAs (intersections) and edges E correspond to AoI road segments. As our aim is to reach all DMEAs within a minimum delay, the weight of an edge e_{ij} can only be the delivery delay of REQ_DMs between two successive intersections (i, j) . This delay is generally a function of a succession of carry and forward phases as it depends on the traffic state of the road segment e_{ij} that connects the two intersections (i, j) . The average transmission delay of a message along a road segment e_{ij} is given in [96] by Equation (4.1):

$$delay_{T_{ij}} = (\pi_F * delay_F + \pi_C * delay_C) * E[R_d]_{min} \quad (4.1)$$

Where,

$E[R_d]_{min}$ is the average number of forward and carry cycles. π_F and π_C are the probabilities of forwarding phases and carry phases, respectively. They are both functions of the vehicle's transmission range and the inter-vehicular distance. $delay_F$ and $delay_C$ represent, respectively, the delays of forwarding and carrying phases. $delay_C$ depends on the speed of the vehicle carrying the message, whereas $delay_F$ is governed by the forwarding policy which will be discussed in the next subsection. Table 4.2 lists the main symbols used to describe OMS.

The fitness $fit_{veh}(seg)$ of veh for the mule role in seg We assume that the road layouts and traffic statistics are available. Indeed, traffic statistics can help to find roads with better connectivity to DMEAs. Such information is becoming more accessible thanks to GPS traces.

The decision variable for the MST formulation and weight matrix construction is given by:

$$X_{ij} = \begin{cases} W_{ij} = delay_{T_{ij}} & \text{if edge } e_{ij} \in E_T \\ 0 & \text{otherwise} \end{cases} \quad (4.2)$$

Table 4.2: List of symbols.

Symbols	Description
e_{ij}	The edge between two vertices (i, j) , which corresponds to the road segment connecting intersections i and j .
$delay_{T_{ij}}$	The average transmission delay of a message along the road segment e_{ij} .
X_{ij}	The weight matrix.
ST, MST, E_T	The spanning tree, the minimum spanning tree, and the set of MST edges resp.
in-DMEA, out-DMEA	In DMEA vehicles and out DMEA vehicles resp.
veh	A candidate for the DM role.
$er_{dir}(seg), er_{pos}(seg), er_{vel}(seg)$	Directional error, Positional error, and Velocity error, resp.
$E_{pos}(seg), E_{dir}(seg), E_{vel}(seg)$	The maximum admissible errors in positioning, heading and velocity, resp.
$fit_{veh}(seg)$	The fitness of a candidate veh for the DM role in the segment seg .

where E_T is the set of MST edges.

The total cost corresponding to the query propagation delay according to a spanning tree ST is written as:

$$C_{Total}(ST) = delay_{Total}(ST) = \sum_{\substack{i,j=1 \\ i,j \in V}}^n delay_{T_{ij}} \quad (4.3)$$

with n is the number of intersections of the AoI.

Thus, having the weight matrix X_{ij} , the CC computes the spanning tree of minimum REQ_DMs propagation delay using one of the most used algorithms such as Prim [92] or Kruskal [63]. Algorithm 1 presents the Prim's algorithm performed by CC. Given the connected weighted graph representing the AoI (line 1), Algorithm 1 provides a spanning tree of minimum delay (E_T) to propagate the REQ_DMs (line 2). The tree is initialized with the CC'S location (line 3). At each iteration, the tree is extended by one edge (road segment) of minimum cost (lines 5-8) among the edges that connect the tree E_T to the vertices that are not yet added to the tree (line 6). The process

continues until adding all the vertices to the tree (line 8).

Algorithm 1 : Prim's algorithm performed by the Central Controller (CC).

- 1: **Input:** the connected weighted graph of the AoI with vertices V and edges E ;
 - 2: **Output:** V_T and E_T which describe a minimal spanning tree;
 - 3: $V_T \leftarrow x$; // x is the node representing the CC's location (starting point) from V
 - 4: $E_T \leftarrow \{\}$;
 - 5: **Repeat**
 - 6: select a road (x, y) with minimum delay such that x is in V_T and y is not;
 - 7: // In case of multiple roads with the same delay, select any of them.
 - 8: Add y to V_T , and (x, y) to E_T ;
 - 9: **Until** ($V_T = V$)
-

In fact, before sending the REQ_DMs, the CC should first, computes the MST of the AoI; i.e., determines the spanning tree to allow the propagation of a given message to all DMEAs of the AoI in a minimum delay. Once determined, the MST is joined to the REQ_DMs message to pilot its propagation. Accordingly, vehicles located in road segments that do not belong to E_T , are not involved in the propagation process. Therefore, upon receiving the REQ_DMs, the vehicle checks its location on an MST segment in order to decide on its commitment to the propagation process. This helps reducing communication cost as REQ_DMs is delivered to the whole DMEAs by crossing only a part of AoI's roads. We assume that the road layouts and traffic statistics are available and are frequently updated by leveraging the data gathered from previous collections. Indeed, traffic statistics provide road density information which is a key parameter to compute probabilities of forwarding and carrying a message along a road segment, thus to estimate the delivery delay of a message along that segment. Moreover, this information is available thanks to GPS traces data.

Forwarding Policy

Note that it's not of our interest to design a new forwarding policy. We just adapt to our needs one of the existing policies that have proven their performance. We made use of the MIN (Middle Is Next) forwarding strategy [95], in which the next transmitter is selected such that to be the closet vehicle to the middle of the previous transmitter's communication range. Because of our scheme aims at minimizing the REQ_DMs propagation delay, we adapted MIN to pick as next transmitter the closest vehicle to the target DMEA so that to accelerate the query dissemination process. Upon receiving the REQ_DMs message, vehicles located on an MST road set a back-off counter (BC) proportional to its distance from the target DMEA using the following equation:

$$BC = \frac{e^{(\lambda * Dist)^2}}{2} \quad (4.4)$$

where λ is the vehicle density and $Dist$ is the distance separating the receiver from

the targeted DMEA. Thus, the closest receiver to the targeted DMEA will be the first to forward the REQ_DMs (because it will have the smallest BC) while the others will stop their timers. Such way, query dissemination delay is minimized while reducing collisions.

4.4.2 Opportunistic Election of Data Mules

This process is opportunistically activated when REQ_DMs reaches DMEAs. It aims to assign to each outgoing segment a DM according to vehicles' speed and heading. To that end, each REQ_DMs receiver should first check its belonging to a DMEA in order to decide about its candidature for the mule role. We delimit the geographical coordinates of a DMEA by the area of the center $C_{DMEA} = cc$ and the radius $R_{DMEA} = r + l$ as it is depicted in Figure 4.3 (a). Where cc and r are, respectively, the center and radius of the intersection, while l is the average length of a vehicle. In the following, vehicles belonging to a DMEA at the time the DM election process is triggered are referred to as 'in-DMEA' vehicles. An 'in-DMEA' vehicle veh is eligible to perform the mule role for an outgoing segment if it moves on or towards that segment with a velocity close to the DM's desired cruise speed. More precisely, based on simple geometrical considerations on the movement of 'in-DMEA' vehicles with respect to the geographical coordinates of intersection segments, each 'in-DMEA' vehicle can assess its ability (fitness) to the role of mule in the segment seg , by computing the following error vector:

$$EV_{seg} = (er_{dir}(seg), er_{pos}(seg), er_{vel}(seg)) \quad (4.5)$$

where $er_{dir}(seg)$, $er_{pos}(seg)$, and $er_{vel}(seg)$ are, respectively, errors in heading, position and velocity with regard to the segment seg , $seg = 1, k$, where k is the number of outgoing segments of the DMEA. Note that the geographical information of intersections and the outgoing segments layouts can be easily obtained from digital maps. However, for the sake of simplicity, we consider the Manhattan like geometry as depicted in Figure 4.3 to illustrate the three error metrics. We denote by $E_{pos}(seg)$, $E_{dir}(seg)$ and $E_{vel}(seg)$ the maximum admissible errors in positioning, heading and velocity, respectively. In case of Manhattan like geometry $E_{pos}(seg) = r$ and $E_{dir}(seg) = \frac{\pi}{2}$ for each outgoing segment. Otherwise, they are functions of intersection geometry (intersection size, road layouts, etc.). $er_{dir}(seg)$, $er_{pos}(seg)$, $er_{vel}(seg)$ are defined as follows:

- **Directional error** $er_{dir}(seg) \in [-\pi, +\pi]$, is the angular deviation with respect to the segment seg direction (see Figure 4.3 (c)). If the vehicle veh is moving on seg , $er_{dir}(seg) = 0$; otherwise $er_{dir}(seg) > 0$;
- **Positional error** $er_{pos}(seg)$, is the distance separating an 'in-DMEA' vehicle veh from the beginning of the segment seg (see Figure 4.3 (b)). Note that if veh is already in seg , $er_{pos}(seg) = 0$; Otherwise, $er_{pos}(seg) > 0$. In the remaining of this section, veh denotes an 'in-DMEA' vehicle.

- **Velocity error** $er_{vel}(seg) = |Vel_d(seg) - Vel_{veh}|$ where, $Vel_d(seg)$ is the desired cruise speed on segment seg and Vel_{veh} is the vehicle's speed. $er_{vel}(seg)$ gives a measure of vehicles moving at around the desired speed or away from it. In fact, the speed of the mule determines the mule-sensor contact time.

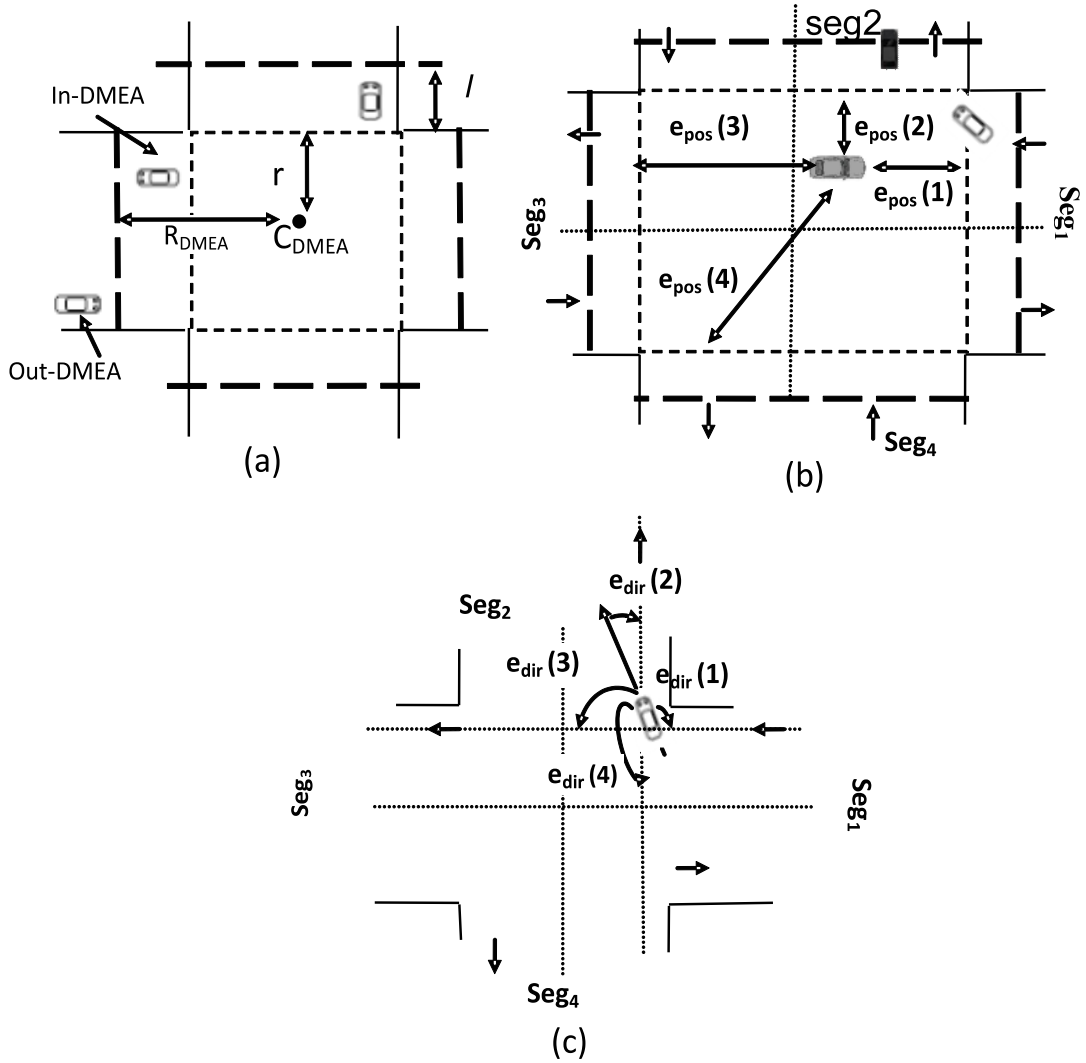


Figure 4.3: Opportunistic election of data mules: inputs parameters: (a)- DMEA's parameters, (b)- angular error, (c)- positional error.

The fitness $fit_{veh}(seg)$ of veh for the mule role in seg is then given by formula 4.6.

$$fit_{veh}(er_{dir}(seg), er_{pos}(seg), er_{vel}(seg)) = \frac{er_{dir}(seg)}{E_{dir}(seg)} + \frac{er_{pos}(seg)}{E_{pos}(seg)} + \frac{er_{vel}(seg)}{E_{vel}(seg)} \quad (4.6)$$

Note that, with this fitness function, the lower the fitness value is, the better the ability of veh for the mule role in seg will be. Hence, veh will be candidate for the segment of the lowest fitness value.

It is worth noting that, with formula (4.6), the fitness function is formulated without taking into account the contribution factor of each input parameter; i.e., input parameters were considered contributing with the same weight in determining a candidate's suitability for the DM function, which does not correspond to reality. Actually, the heading parameter is more representative than the distance and the velocity parameters, due to the fact that a car moving at the desired speed can be a DM only on the road segment it is heading towards. Therefore, the heading parameter should be assigned a high contribution factor when calculating the ability of vehicles for the mule role. We define the contribution factor of an input parameter of the fitness function as the percentage with which that parameter participates in determining the suitability of a candidate vehicle for the DM role. By taking into account those factors, the fitness function has been revised and modified from formula (4.6 [27]) to formula (4.7 [26]).

$$fit_{veh}((er_{dir}(seg), er_{pos}(seg), er_{vel}(seg))) = a * \frac{er_{dir}(seg)}{E_{dir}(seg)} + b * \frac{er_{pos}(seg)}{E_{pos}(seg)} + c * \frac{er_{vel}(seg)}{E_{vel}(seg)} \quad (4.7)$$

Where,

a , b and c are weighting coefficients which measure the contribution factor in the fitness function. The weighting coefficients are normalized such that: $a + b + c = 1$. Since with this fitness function, the lower the fitness value is, the better the ability of veh will be, weighting coefficients are to be chosen so that small weights are given to favorite parameters. To that end, we define preference between input parameters so that $er_{dir}(seg) < er_{pos}(seg) < er_{vel}(seg)$, because it seemed obvious that the direction is the best indicator for the next destination of a vehicle, then its position and finally its velocity. Therefore, weighting coefficients respond to the following constraint: $a < b < c$.

Actually, it is not always possible for a candidate to determine its target segment at the time of the request for opportunistic mules reception. Consequently, two cases can rise according to the ability of an "in-DMEA" vehicle to determine its destination during the calculation of its fitness value. Algorithm 2 describes how an "in-DMEA" vehicle is opportunistically assigned the role of DM for a given outgoing segment whose illustration is given as follows:

Case 1: The candidate is able to determine its targeted segment If the destination of an "in-DMEA" vehicle veh is determinable (i.e., $er_{dir}(seg) \leq E_{dir}(seg)$ and $er_{pos}(seg) \leq E_{pos}(seg)$) upon receipt of the request for opportunistic mules (such as for the white car in Figure 4.4), and the velocity of veh is close to the desired cruise speed (i.e., $er_{vel}(seg) \leq E_{vel}(seg)$), veh considers itself a candidate for the mule task in the targeted segment and thereby it cancels its assessment towards the remaining segments (lines 6 – 24). Then, it should assess its ability compared to its potential

competitors to take in charge the target segment (line 9). To do so, it calculates its fitness $fit_{veh}(seg)$ towards that segment using Equation (4.7) (line 12).

Case 2: The candidate is not able to determine its targeted segment Some candidates may not be able to determine the target segment due to their location and orientation when receiving the request for opportunistic mules (lines 26 – 30). The black car and the red one in Figure 4.4 are examples. In this case, the candidate continuously assesses its direction, by comparing the angular and positional errors to the corresponding acceptable maximum errors, until determining its direction (lines 31 – 39). Once its direction is determined, it advertises its commitment, if it is moving around the desired speed (line 40). In case the candidate does not meet the condition, it hands over the mule role to another vehicle moving at around the desired speed (line 42).

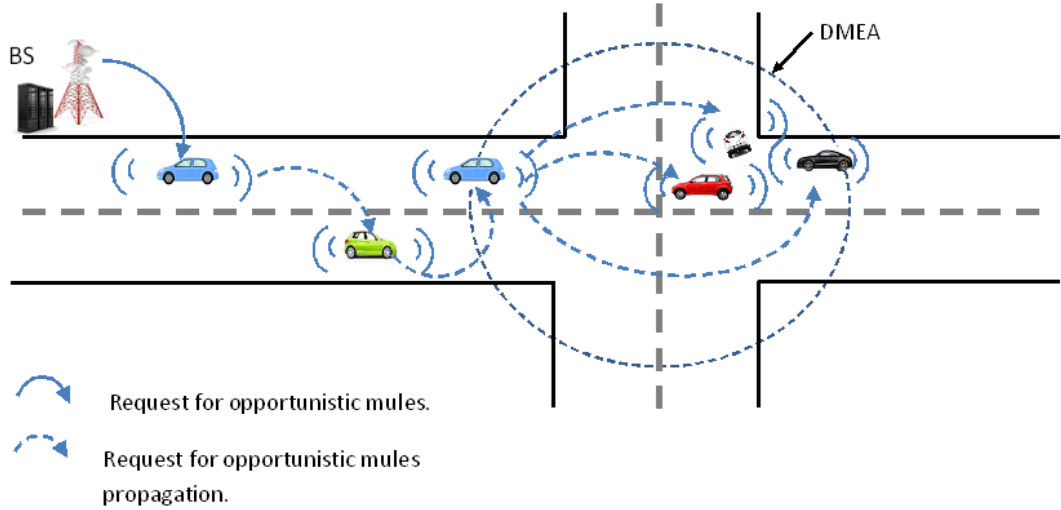


Figure 4.4: Opportunistic mules election: examples of candidates with determined destination and candidates with undetermined destination.

Each candidate broadcasts a `DM_ANNOUNCE_SEG` message to announce to its competitors its commitment for the mule task in the segment seg . To avoid candidature announcement collisions, candidates wait a back-off time (BCT_{seg}) proportional to their fitness (lines 14) before advertising their commitments. The BCT_{seg} calculation is given by formula (4.8) as follows:

$$BCT_{seg} = fit_{seg} * \Delta T \quad (4.8)$$

with $\Delta T = \frac{prio(seg)}{x*y} * T_s$, where $prio(seg)$ is the priority order of the segment seg . We defined a priority order between DMEA segments to avoid collisions between candidates from different segments in case they have the same fitness value. To avoid collisions between competitors of a given segment, the geographical coordinates x and

Algorithm 2 : Opportunistic data mules election. /*runs at each in-DMEA vehicle*/

```

1: Input:  $E_{pos}(seg), E_{dir}(seg), E_{vel}(seg)$ : maximum acceptable errors in position, direction and velocity;
2: Vector  $Vel_d[nb\_segments]$ : a vector of desired cruise velocities;
3:
4: Output: the segment to which the in-DMEA vehicle should guarantee the mule role
5: Begin
6: for each segment  $seg$  of the DMEA do {
7:
8:  $EV_{seg} = Calculate(e_{dir}(seg), e_{pos}(seg), e_{vel}(seg))$ ;
9: if ( $e_{dir}(seg) \leq E_{dir}(seg)$  and  $e_{pos}(seg) \leq E_{pos}(seg)$  and  $e_{vel}(seg) \leq E_{vel}(seg)$ ) then {
10:   //  $V$  is a candidate for  $seg$ .
11:
12:  $fit_{seg} = compute\_fitness\_seg(e_{dir}(seg), e_{pos}(seg), e_{vel}(seg))$ ;
13:
14:  $BCT_{seg} = fit_{seg} * \Delta T$ ; //initialize the back-off time
15:
16: On.Timeout( $BCT_{seg}$ ) {Broadcast (DM_ANNOUNCE_SEG);}
17:
18: On.Receive(DM_ANNOUNCE_MESSAGE_SEG, before_  $BCT_{seg}$ -expiration) {
19:   //there is another more fitted candidate for  $seg$ 
20:
21: Cancel.Broadcast (DM_ANNOUNCE_SEG);}
22:
23: exit;
24: }}
25:
26: if ( $(e_{dir}(seg) > E_{dir}(seg)$  or  $e_{pos}(seg) > E_{pos}(seg)$  ) for all outgoing segments) then{
27:   // the vehicle direction cannot be deduced at present
28:
29:  $Probable\_segments\_set = all\_outgoing\_segments$ ;
30:
31: repeat until ( $(\exists seg\_k \in Probable\_segments\_set$  such that  $e_{dir}(seg\_k) \leq E_{dir}(seg\_k)$  and  $e_{pos}(seg\_k) \leq E_{pos}(seg\_k)$ ) or ( $Probable\_segments\_set = \emptyset$ ))
32: {
33:
34: On.Receive(DM_ANNOUNCE_MESSAGE_SEG_l){  $Probable\_segments\_set = Probable\_segments\_set - seg\_l$ ;}
35:
36: for each segment  $seg\_k$  of the  $Probable\_segments\_set$  do {
37: calculate( $e_{dir}(seg\_k), e_{pos}(seg\_k)$ ); } }
38:
39: if ( $\exists seg\_k$  such that  $e_{dir}(seg\_k) \leq E_{dir}(seg\_k)$  and  $e_{pos}(seg\_k) \leq E_{pos}(seg\_k)$ ) then{ // the direction of the candidate is now determined
40: if ( $e_{vel}(seg\_k) \leq E_{vel}(seg\_k)$ ) then Broadcast(DM_ANNOUNCE_SEG_k);
41: else
42: Hand_Over_Mule_Role ( $Vel_d[seg\_k]$ , when_approaching_sensor_position); }
43: End

```

y of a candidate veh are used. T_s is the duration of one round transmission. According to the back-off time formulation, the candidate with the smallest fit_{seg} value will be the first to announce its commitment while the others will abort their announcements (lines 16–18). This provides an opportunistic distributed election of data mules with a low communication cost. Indeed, thanks to the back-off time formulation that depends on the candidate ability level, the communication cost of DMs election is reduced to only one message per a DMEA segment, namely the DM_ANNOUNCE_SEG message. DMs in charge of MST segments first pursue the propagation of the request while those who manage "not_MST" segments may start collecting data as outlined in the global diagram of OMS depicted by Figure 4.5.

Note that a segment with no DM has a second chance to be taken into account by the neighboring DMEA. To do so, OMS leverages the DM_ANNOUNCE messages. Indeed, these messages, likely to be received by vehicles located outside DMEAs called "out-DMEA" vehicles, may serve as an indicator of the election period of DMs. When an "out-DMEA" vehicle receives a DM_ANNOUNCE message, it should check the message source. If it is from neighboring segments, the "out-DMEA" vehicle starts a back-off timer $MBCT_{seg}$ (the Maximum Back-off Time) which corresponds to the time taken by the weakest candidate to announce its commitment. In such a case, $EV_{seg} = (E_{pos}(seg), E_{dir}(seg), E_{vel}(seg))$. Thus, $MBCT_{seg}$ can be given by:

$$MBCT_{seg} = Max_fit_{seg} * \Delta T \quad (4.9)$$

If $MBCT_{seg}$ expires before DM_ANNOUNCE_SEG reception, the "out-DMEA" vehicle renews the propagation process by sending a REQ_DMs message with dm flag set to 0 whether it belongs to an MST segment or not. Indeed, to ensure effective data mules' election, REQ_DMs message should contain additional information. Namely, for each DMEA, for each outgoing segment, the following information has to be specified.

- The desired cruise speed $Vel_d(seg)$,
- A dm bit to point out to the target DMEA whether the segment on which the REQ_DMs was propagated has already a DM or not. Indeed, this allows to segments with no DMs to be taken into account by the target DMEAs as each segment links two DMEAs. Moreover, the dm bit avoids assigning more than one DM for a given segment. This prevents sensors sending redundant data hence, avoiding wasting their energy.

To summarize, Figure 4.5 describes the global operation of OMS. At the reception of a request for opportunistic mules (REQ_DMs), a vehicle first checks whether it is an 'in -DMEA' (action 1 in figure 4.5) to decide about its candidacy to the DM role. 'in -DMEA' vehicles calculate the error vector EV and check their directions (actions

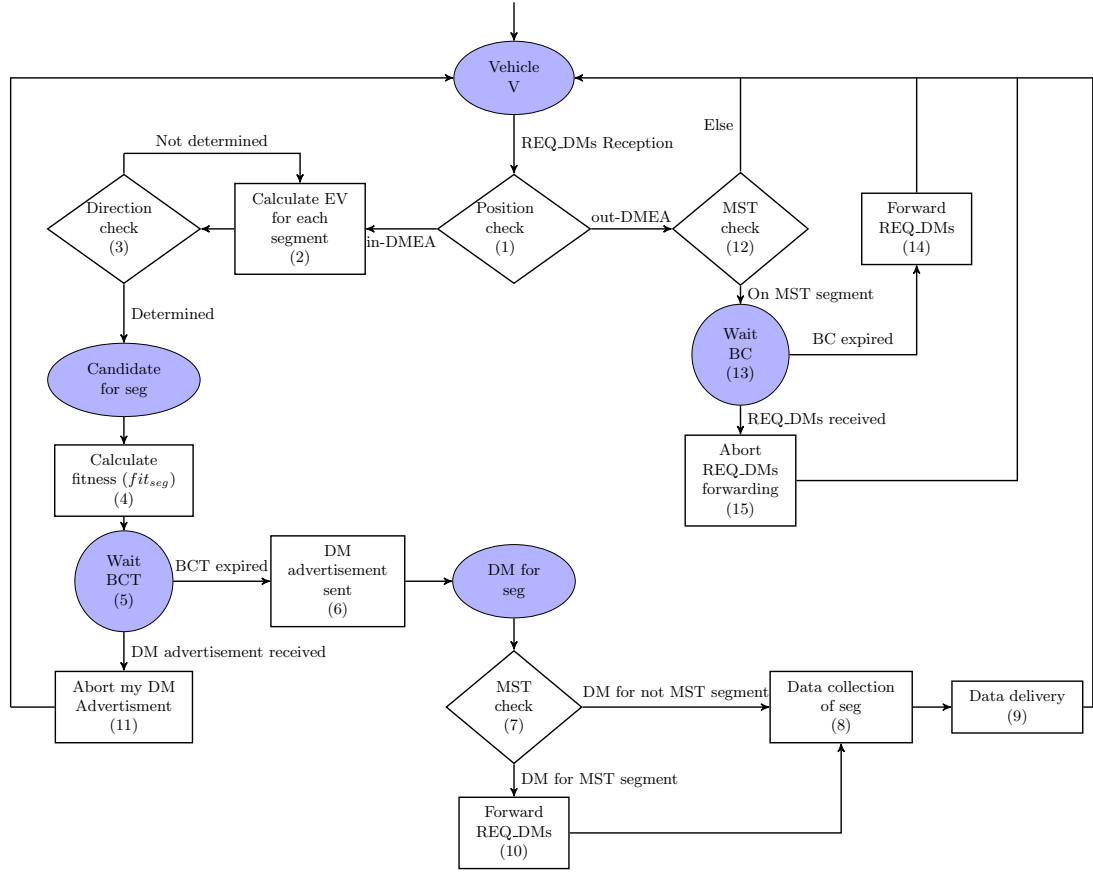


Figure 4.5: OMS global diagram.

2 and 3). Each ‘in -DMEA’ vehicle with determined direction declares itself candidate for the targeted segment, calculates its fitness toward that segment and waits for a back-off time BCT to advertise its commitment for the mule role (actions 4, 5 and 6). If, while waiting, the vehicle receives a ‘DM advertisement’, it cancels its candidacy (action 11) and returns to the initial state. DMs in charge of ‘not-MST’ segments proceed to the data collection (action 8) and to their delivery (action 9) at the end of the segment, while those in charge of ‘MST’ segments should first propagate the request to the successive DMEAs (action 10). If the REQ_DMs message receiver is not an ‘in-DEMA’ vehicle, it checks whether it is located on an ‘MST’ segment to contribute to the request propagation (actions 12, 13, 14 and 15).

The candidate with the smallest $fit_{veh}(seg)$ advertises first its candidature when the others cancel their advertisements. With this way, DMs are opportunistically elected in a distributed manner without additional communication cost. Indeed, thanks to the back-off time which depends on the candidate fitness, the communication cost of DMs election is reduced to only one message per outgoing segment. DMs of ‘MST’ segments should first pursue the propagation of REQ_DMs to next DMEAs before starting data collection, while those of ‘not-MST’ segments should start to collect data of theirs

segments.

4.4.3 Data Harvesting

After announcing its responsibility to perform the mule role in a specific segment, each DM should proceed to the download of sensory data of the encountered IoT nodes. Although, the mule discovery process and transmission protocol go beyond the scope of this thesis, it is worth noting that on demand mule discovery is more suitable to our scheme for two reasons:

1. Sensor nodes are deployed in known positions allowing the DM to initiate a connection when approaching their coverage area. The geographical coordinates of IoT nodes could be included in the REQ_DMs message.
2. Having the desired DM's cruise speed, the collection periodicity and the average time of query propagation of the AoI, we can predict when the DM will be in visibility of each sensor. This allows implementing sleep/wakeup scheduling schemes such as [113] which allow sensors to wake up only when the opportunistic DM is likely to pass.

4.4.4 Data Delivery

After collecting the data of the encountered IoT nodes, each DM should deliver its data to the CC. As reducing the collection latency is a primary goal of the proposed scheme, data delivery has to be performed within an opportune time. This requires a delay optimal routing strategy which is still considered as the main challenge in vehicular networking. In fact, data may be delivered to the CC in two ways as:

1. Relying on existing delay optimal vehicular routing protocols such as OVDF [30], which exploits vehicle traffic statistics and knowledge of future trajectories of vehicles to route the data within a minimum delay.
2. Making use of query propagation and data harvesting phases to estimate the current traffic state of the AoI in order to compute optimal routes towards CC. Designing such a strategy and comparing it with existing delay optimal routing protocols are planned as perspective of this thesis.

4.5 Simulations and Measurements

We investigate the OMS scheme performance using the network simulator NS-3 [1]. To the best of our knowledge, there is no DM-based work dealing with the high latency of data collection in urban environment. So, OMS couldn't be compared to other work because it is obvious that neither objectives nor the metrics could be fairly compared.



Figure 4.6: Road network topology in the Bab ezzouar, Algiers, Algeria, urban area.

However, we applied OMS to several simulation scenarios to evaluate its characteristics in detail. In our simulations, we used a map of an Algerian region, namely Bab ezzouar, which is a part of the metropolitan area of Algiers. The simulation scenarios are extracted from OpenStreetMap [2] as it is shown in figure 4.6. At first, we evaluated OMS in an area of interest of $1500m \times 800m$ consisting of 16 intersections and 21 road segments. To assess OMS in terms of scalability, we rerun the same experiments over two other AoI namely $2km^2$ and $3km^2$. Also, we performed OMS on various road vehicular densities (5veh/km – 50veh/km) to evaluate its performance within different traffic volumes. For that, we used SUMO [3](Simulation of Urban Mobility) to adjust OSM mobility traces in order to obtain the desired road densities.

We used IEEE802.11p (WAVE) MAC layer protocol and Nakagami propagation model physical layer protocol. The later is reported in [103] to be the most suitable propagation model for VANETs. Transmission ranges are set to 250 m for both vehicle-to-BS and vehicle-to-vehicle communications. The simulation parameters are summarized in Table 4.3.

Table 4.3: Simulation Setup.

Parameter	Value
Size of simulation area	1, $2km^2$, $2 km^2$, $3 km^2$
Vehicular density	5-50 veh/km
Transmission range	250 m
Simulation time	300 seconds
Vehicles' velocity	18-50 km per hour
MAC protocol	IEEE802.11p (wave)
Propagation model	Nakagami_m
Bandwidth	6 Mbp/s

4.5.1 Performance Metrics

We assessed OMS based on the following performance metrics.

- **Average query propagation delay**, which we define as the average delay required by the request for data mules to reach all DMEAs. Query propagation delay characterizes the required time to cover as maximum as possible the AoI by opportunistic mules.
- **Average coverage ratio**, which we define as the average number of road segments to which data mules have been successfully assigned to the total number of AoI's segments. The average coverage ratio shows the ability of the OMS election process to successfully assign to each road segment a data mule in an opportunistic manner, hence covering the maximum of the AoI.
- **Network load**, which is defined as the amount of data being carried by the network due to the query propagation and the DM election process. It measures the OMS cost in terms of bandwidth consumption. It is measured in number of bytes per second.

4.5.2 Simulation Results

The curves plotted in the following figures represent the average of the results obtained from 10 independent simulations. First, we examine the impact of weighting coefficients of the fitness function parameters on the OMS election process. For that, we vary the values of weighting coefficients and compute at each settings the percentage of ideal DM with respect to a supervised election. The aim of this experiment is to determine the best values of the weighting coefficients that allow the OMS election process to converge towards a robust election. Table 4.4 presents the values given to the weighting coefficients while figure 4.7 shows obtained results for different traffic volumes (low: 10 veh/km, medium: 20 veh/km and high: 30 veh/km).

Table 4.4: The values of weighting coefficients of the fitness function parameters.

Coefficient	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6
a	0.05	0.1	0.15	0.2	0.25	0.3
b	0.1	0.15	0.2	0.2	0.3	0.33
c	0.85	0.75	0.65	0.55	0.45	0.37

The results show that more than 80% of elected DMs are ideal in low and medium density whereas in dense area the percentage of ideal DMs can exceed 90% for all the values given to the weighting coefficients. We also show that all the variants of the weighting coefficients give the same result in low traffic volume and a slight variation

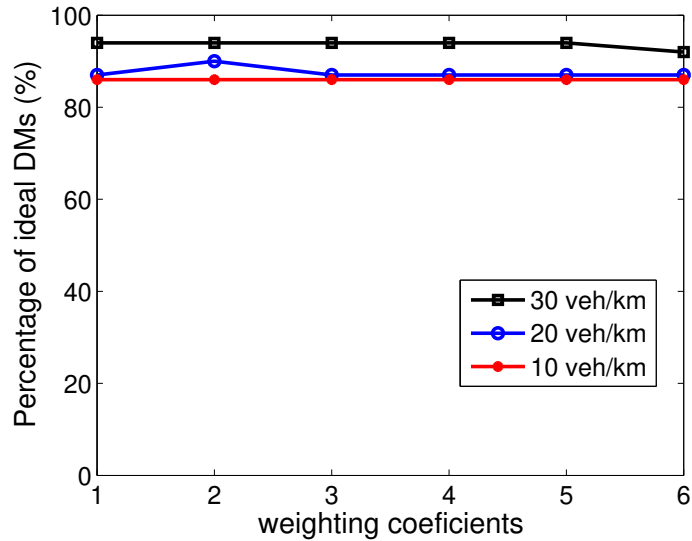


Figure 4.7: Impact of weighting coefficients of the fitness function parameters in low, medium and high density networks.

in high and medium traffic volumes. This demonstrates that the priority order established on those coefficients allows electing the best DMs in almost all scenarios. In fact, prioritizing the direction parameter helps assigning to segments only candidates moving towards those segments. Moreover, the independence of coefficients from network density presents an advantage of OMS from implementation point of view. Indeed, candidates are not constrained to obtain the current density to adjust their coefficients during election process. This allows preventing OMS from additional communication cost.

Next, we evaluate the capability of OMS to designate DMs to all road segments of the AoI. For that, we consider two simulation scenarios. In the first one, the election is performed using the fitness function stated in the primary version of this work [?] (formula 4.6). In this case, the coefficients of the fitness function parameters are identical and are all equal to 1 ($a = b = c = 1$). In the second scenario, the election is performed according to the fitness function of the enhanced OMS (formula 4.7). In this case, the direction parameter is preferred over the position which in turn is preferred over the speed ($a < b < c < 1$). Figure 4.8 plots the obtained results. As it can be seen, the results obtained by the second scenario are closer to the supervised scenario compared to the first one. This is due to the fact that in the first scenario, direction, position and speed parameters contribute with the same weight in the election of DMs. This may allow to candidates positioned near a given segment and moving with a velocity close to the desired speed to be elected for that segment even if their error in direction is very close to the maximum admissible error. This explains the higher percentages ob-

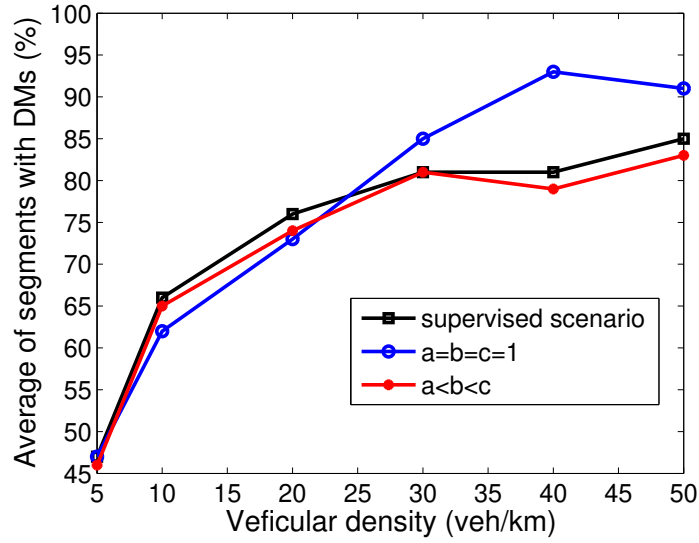


Figure 4.8: Percentage of road segments covered with DMs, varying the vehicular density.

tained by the first scenario compared to the supervised one. Note that, those cases are discarded from election in the second scenario thanks to the high preference assigned to the direction parameter.

The CDF (Cumulative Density Function)¹ of road segments ratio being assigned DMs is depicted in figure 4.9. As it is shown, the probabilities that OMS provides large parts of AoI with DMs is satisfactory, even in the case of low traffic volume. Indeed, there is a moderate probability to cover more than half of the AoI with DMs in low traffic volume. However, this probability increases with the increase in the volume of traffic to reach the values of 0.6 for a moderate volume of traffic and 0.8 for a high volume. This is confirmed by the plots of figure 4.8 where we can see that high ratios of covered segments are obtained by higher values of vehicle densities. This demonstrates the efficiency of the DM election algorithm in case of candidates availability. Note that in a sparse network (5veh/km), OMS can provide 46% of road segments with DMs which we consider as a satisfactory result as it represents approximately the half of the IoA. Indeed, OMS offers a second chance for segments not being assigned a DM, to be considered in an adjacent DMEA thanks to the control information included in the REQ_DM message.

Next, we evaluate the average delay of the query propagation. The aim of this experiment is to measure the required average time to designate DMs to the entirety of

¹The cumulative density function (CDF) of a real-valued random variable X evaluated at x , is the probability that X will take a value less than or equal to x [31]. It is given by [88]: $F(x) = P(X \leq x)$.

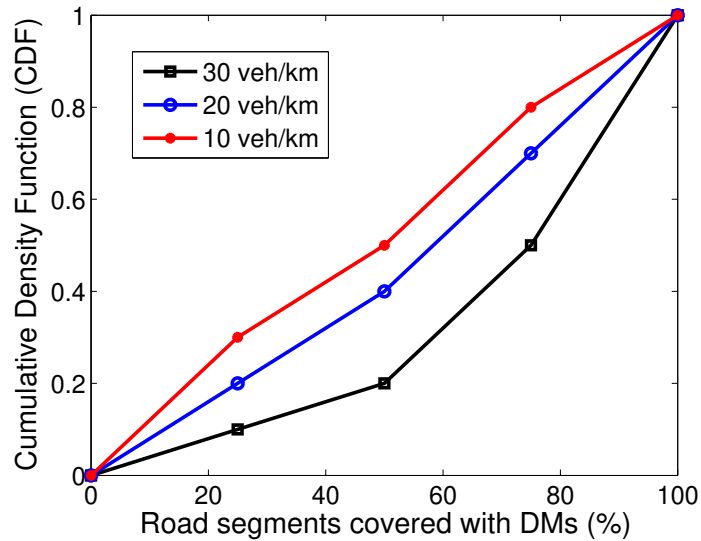


Figure 4.9: CDF of road segment ratio with DMs in low, medium and high density networks.

AoI. Looking at the figure 4.10, we observe that the scheme manages to propagate the REQ_DMs to the entirety of DMEAs within less than 1 minute (56 seconds) in a sparse network (5veh/km). This delay decreases then with the vehicular density increase to reach a value of less than 1 second in a highly dense network (50 veh/km). Indeed, it is obvious that in dense area the connectivity level of network increases, raising then the chance to multi-hop the REQ_DMs to neighboring DMEAs, unlike low dense area, where the REQ_DMs is more likely to be carried by the holder vehicle to the target DMEA. We agree that query propagation delays are too acceptable in low density and very interesting in medium and high densities. This is due to the fact that:

- (i) The REQ_DMs is disseminated according to an optimized strategy which selects roads with high connectivity to propagate the request to neighboring DMEAs.
- (ii) The next transmitter is selected to be the closest car to the targeted DMEA to minimize query propagation delay.

We point out that for medium vehicle density; almost all road segments of the AoI will be assigned a DM in less than 10 seconds.

Figure 4.11 shows the average network load. This figure gives a measure of the cost of query propagation and DMs election in term of bandwidth consumption. As we observe, the network load increases with the vehicular density increase. It takes small values (less than 3000 bytes/sec) in low and medium density values (≤ 20 veh/km). In addition, it takes acceptable values (< 4000 bytes/sec) in high densities (between 30 and 40 veh/km). Even with 50 veh/km, the load reaches 6000 bytes/sec. In fact,

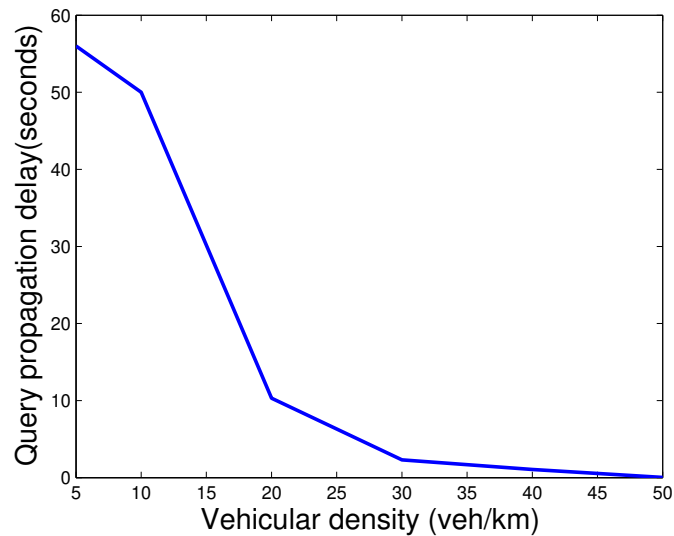


Figure 4.10: Propagation delay of the request for DMs, varying the vehicular density.

with high vehicular density, REQ_DM messages are more likely to be multi-hopped to DMEAs, increasing then the network load. However, it remains very low compared to the available bandwidth. Indeed, in the worst case (6000 bytes/sec), the protocol consumes 0,8% of the available bandwidth. The results demonstrate that the scheme optimizes the bandwidth. This was expected due to : i) the optimized query propagation strategy that broadcasts queries only over a portion of the road network, ii) in addition to the opportunistic DM election that costs only one message per road segment.

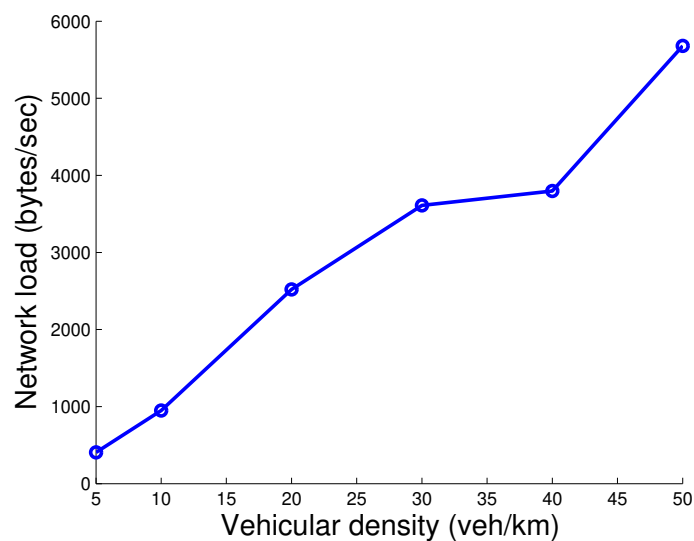


Figure 4.11: The network load induced by query propagation and DMs election.

Next, we assess the impact of DMEA size on the efficiency of the scheme. As the

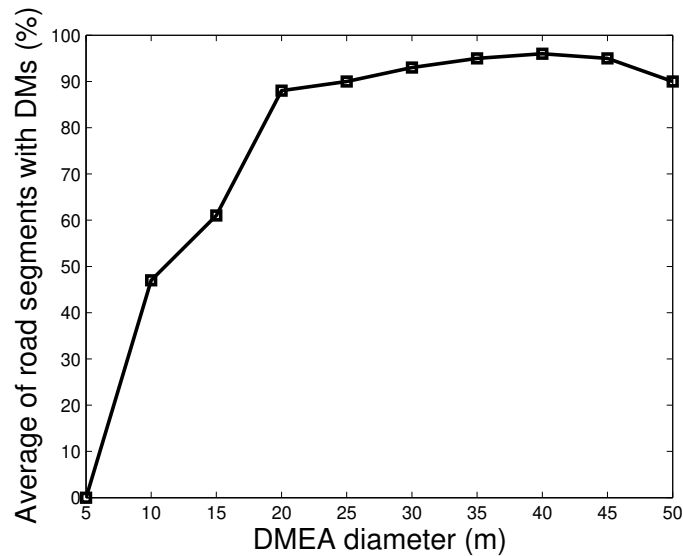


Figure 4.12: Impact of DMEA size on scheme efficiency.

candidacy of a vehicle to the mule role depends on whether it belongs or not to a DMEA, the size of the latter influences the number of candidates, hence that of the designated DMs. The goal of this experiment is to determine the best values of this parameter to achieve a better efficiency of the scheme (assign DMs to the maximum of the AoI roads). Figure 4.12 shows the obtained results. If we consider the smaller DMEA size ($10 \times 10m^2$), which corresponds to the average size of the considered scenario intersections, we find that less than 50% of the AoI roads will obtain DMs which degrades the scheme efficiency. However, with the expansion of the DMEA size, the scheme efficiency improves. This demonstrates the strategy we used to define the area of DMEA (section 4.4.2). We find that OMS reaches its best performance (more than 85% of AoI roads will have their DMs) when the DMEA size is at least of $20 \times 20m^2$. Finally, the results show that sizes ranged from 20×20 to 50×50 allows a good efficiency of the scheme albeit, with the size of $40 \times 40m^2$, it achieves its optimum efficiency (more than 95% of AoI roads are assigned DMs).

In the sequel, we assess the scalability of OMS by applying it to wider AoIs of different sizes. The aim of these experiments is to evaluate the capability of the scheme to maintain its performance when considering larger AoIs. This allows assessing OMS behavior at large scale and within complex circumstances (more intersections, more road segments and longer road segments). We start by evaluating the ability of OMS to preserve its efficiency in terms of number of segments being assigned DMs by applying it to three AoIs respectively $1, 2km^2$, $2 km^2$, $3 km^2$ with different network densities: 5veh/km (sparse), 20veh/km (medium) and 40veh/km (dense). As it is observed from Figure 4.13, the increase of the AoI size does not cause degradation of OMS perfor-

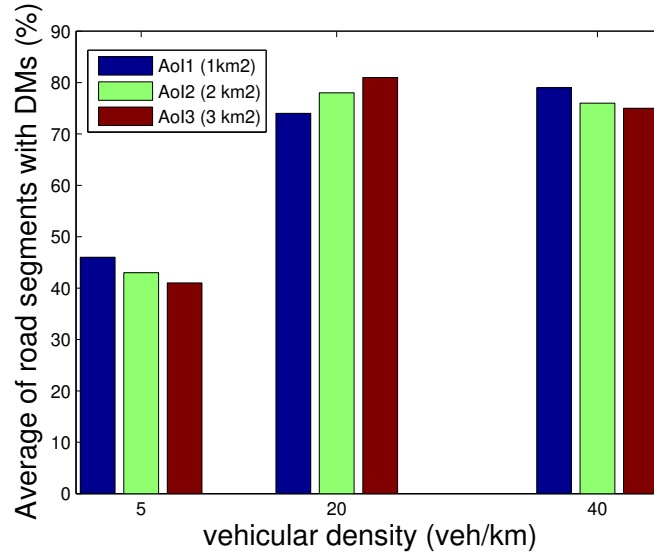


Figure 4.13: OMS efficiency in three AoIs with different sizes.

mance. On the contrary, we notice a slight improvement on the scheme efficiency in case of medium network density. However, the results exhibit a slight decrease of the number of segments covered with DMs in sparse and dense networks. Hence, we can deduce that the scheme preserves its efficiency when applied to large areas, because the total number of road segments having obtained DMs depends on the number of elected DMs within each DMEA, which in turn depends on the number of in-DMEA vehicles and their movements. Thus, OMS efficiency in terms of data mules coverage is independent of the AoI size.

Figure 4.14 shows the propagation delay of the request for DMs measured in three AoIs $1, 2 \text{ km}^2$, 2 km^2 , 3 km^2 . As we can observe, the query propagation delay increases with the increase of the AoI size. However, it remains acceptable in medium and dense networks (less than 30 sec). This demonstrates that the major factor that controls the query propagation delay is network connectivity, since OMS basically selects roads with high connectivity to disseminate the request for DMs to DMEAs. Indeed, short delays correspond to multi-hop transmission of the query which could only happen in good connectivity networks; whereas long delays can only be explained by weak connectivity which implies the carry of the query to the DMEA. Figure 4.15 that plots the network load corresponding to each scenario confirms that result. Indeed, looking at Figure 4.14 and Figure 4.15, we can observe that, regardless of the AoI size, long delays correspond to low network loads, referring to the carry of the query, and short ones correspond to high network loads; indicating that the query was multi-hopped to the DMEAs. From that, we find that OMS maintains its performance when applied to larger scales. This is due to the MST based query dissemination that selects always

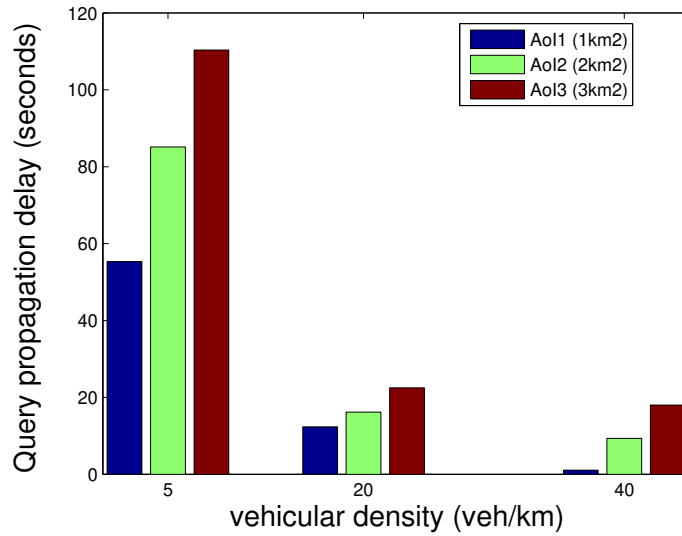


Figure 4.14: Impact of AoI size on query propagation delay.

roads with high connectivity to deliver the request for DMs to DMEAs.

4.6 Conclusion

The Data Mules approach has proved its efficiency in collecting data from static sensors. However, being applied to urban environment leads to high data delivery latency due to the characteristics of these environments. Moreover, the previous DM-based data collection works for urban environments assign long trajectories to the vehicles acting as data mules which amplifies the high latency drawback. To circumvent this issue, we proposed in this chapter to take benefit from the high and long presence of vehicles in urban areas to increase the distribution degree of the collection process by providing each road segment of an area of interest with a data mule to collect and deliver its data. We have proposed an optimized strategy to disseminate a request for Data Mules to roads intersections where Data Mules have to be opportunistically elected. The request dissemination strategy is based on minimum spanning tree (MST) where only a portion of road segments of the areas of interest are used to propagate the request which allow minimising the query propagation time with low overhead. The data mule election is opportunistically triggered at roads intersections where existing vehicles are designated according to their speed, heading, and position.

We showed, through NS3 simulations, that Data Mules are designated in an effective and efficient way. A vehicle density of 30 veh/km is sufficient to provide almost all the roads of the area with Data Mules within an opportune time. However, good results are also achieved with lower densities. Indeed, thanks to our strategy, the time has passed from the order of hours to the order of minutes. Furthermore, applying OMS

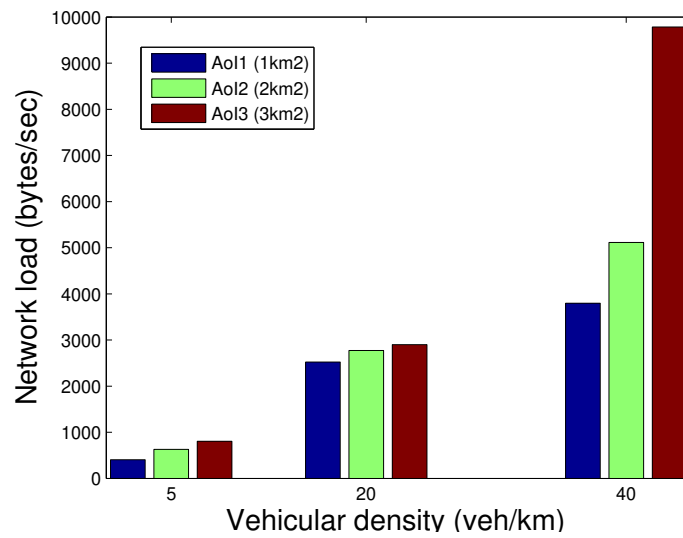


Figure 4.15: Impact of AoI size on network load.

to larger areas demonstrates its scalability. However, in sparse networks, such as in the nighttime scenarios, probabilities of obtaining opportunistic mules are reduced. To overcome this limitation, we propose to support OMS with dedicated vehicles that will act as Mobile Sinks in sparse areas and delegate opportunistic sinks in congested spaces as we will discuss in the next chapter.

Chapter 5

Congestion Aware Data Collection With Mobile Sinks in Smart City

5.1 Introduction

In the previous chapter, a vehicle-based data collection scheme for short delay smart city applications is presented. As we have seen, the scheme which we called OMS (for Opportunistic Mules for Short delay smart city applications), has reduced the data collection latency from 24 hours, which is the result obtained in [18], to few minutes. This significant improvement in the latency of the data collection process is due to the increase in the degree of distribution of the collection process, which is achieved by taking advantage of the opportunistic presence of urban vehicles to provide the city's roads, as much as possible, with data mules. Moreover, OMS is able to provide almost all of the area of interest with opportunistic mules in short time. However, in sparse networks, such as in night scenarios, obtaining opportunistic mules is likely reduced.

On the other hand, employing dedicated vehicles as Mobile Sinks for data collection from urban environments is also suggested [91]. Nevertheless, implying a large number of dedicated vehicles to obtain a reduced collection latency is costly. In addition, in traffic congestion situations, a mobile sink may take a long time to reach the targeted sensor, which will amplify the data drop-off problem (buffer overflow). Therefore, it would be interesting to support dedicated vehicle-based solutions with opportunistic vehicles in order to combine the advantages of both schemes, especially in terms of short data collection latency, high data collection rate and cost effectiveness. In this direction, we propose in this chapter, a congestion aware mobile sink-based scheme [25] that, in traffic jam situations, opportunistically delegates vehicles to act as slave sinks in congested roads and to forward the collected data to the mobile sink that will act as the Master of the area of interest. This increases the amount of mobile sinks operating in the same area, which therefore leads to fast data collection process.

The remainder of this chapter is organized as follows. Section 5.2 gives the motivation behind this proposal. Section 5.3 is devoted to define assumptions and the data collection architecture of the proposed scheme. Section 5.4 explains the design details of the proposed scheme and elaborates on how to designate slave sinks. Section 5.5 provides evaluation performance results. Finally, Section 5.6 concludes the chapter.

5.2 Motivation

In urban scenarios, sinks' mobility is constrained by traffic conditions. Moreover, in congested areas, the mobile sink (MS) may be stuck, resulting in delayed data collection. Meanwhile, roads could be populated by citizen vehicles that cross de facto the MS' targets. This motivates us to support the MS-based data collection approach by slave sinks (SSs) that could be designated by MS among its neighboring vehicles to collect data from congested roads and to forward it to the MS or to the closest Roadside unit (RSU) as depicted in Figure 5.1. Doing so, MS will be freed from collecting congested roads data. In addition, by running in less populated roads, it collects data and delegates new slaves further. Hence, increasing the number of slaves will speed up the collection process to meet the requirements of short delay applications. This new scheme provides the following advantages:

- It reduces data collection latency at low cost by taking benefit from congestion situations to create a collaboration between dedicated vehicles equipped with mobile sinks (the masters) and citizen cars (the slaves).
- It guarantees a uniform distribution of data collection by assigning masters to less congested roads and slaves to congested ones.

5.3 Assumptions

Lets consider a city with IoT nodes, roadside units (RSUs), communicating vehicles and at least one master mobile sink which could be a robot or a dedicated vehicle responsible for collecting data from a geographical area. The IoT nodes deployed across the city roads sense the urban area and periodically generate sensing data to be delivered to one of the RSUs which are positioned on most intersections. RSUs are responsible for forwarding collected data to the central controller via wireless/wired links. Vehicles are equipped with digital road-map, GPS device and multiple wireless interfaces to communicate with IoT sensors, other vehicles and with RSUs. Figure 5.1 illustrates the architecture of the proposed data collection scheme.

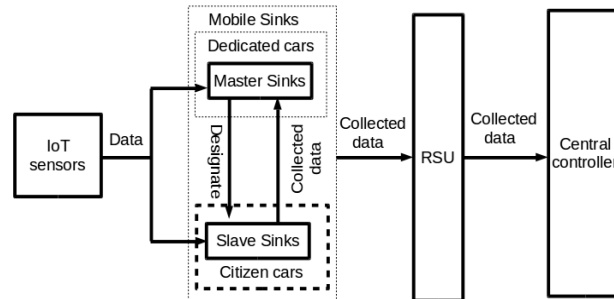


Figure 5.1: Data collection architecture.

5.4 The scheme design

As mentioned before, the goal of the new data collection scheme is to leverage congestion situations to support mobile sinks, that are mounted upon dedicated vehicles, with opportunistic collaborative sinks designated among citizen cars to collect data of congested roads. To this end, the scheme assigns a role to each vehicle contributing to the collection process as follows:

- The masters (*MasMS*, for Master Mobile Sink) are vehicles occupied with mobile sink initially dedicated for data collection and are responsible for designating opportunistic collaborative sinks (slave sinks). An MS is referred to as *MasMS* only after designating the first slave.
- Slaves (*SS*, for slave sinks) are vehicles of all types (personal, bus, ...) designated by *MasMS* as opportunistic collaborative sinks to collect data from dense roads.

Data collected by *SSs* is to be delivered to their masters or to RSUs. Therefore, the scheme operates in three phases: (1) Slave sinks designation, (2) Data collection, (3) Data delivery.

5.4.1 Slave sinks designation

The process of slave sinks designation only occurs in road intersections because they are considered as outgoing points for road segments. The process is triggered by a dedicated MS that aims to designate for each outgoing road, except of the less congested one, an *SS* according to its direction. Thus, the less congested road is to be taken in charge by MS, while the remaining outgoing roads will be ensured by *SSs*. To that end, MS asks neighboring vehicles for *SSs* by broadcasting the message *ASK_SS*. In order to limit the number of responses, thus avoiding conflicts, only vehicles belonging to the

Slaves Designation Area (SDA) are allowed to respond. The SDA includes the area of intersection and a part of each outgoing road segment of length l as it is depicted in Figure 5.2(a). Upon receiving ASK_SS message, vehicles belonging to SDA assess their susceptibility to be the slaves of the MS. A vehicle veh is prone to be a slave on an outgoing road R_i ($i = 1, n$, where n is the number of outgoing road segments) if it is moving on or towards R_i . Therefore, the direction and the location of veh with regards to each R_i are necessary to determine its qualification for the SS function on R_i . More precisely, based on simple geometrical considerations on the movement of veh with respect to the geographical coordinates of intersection segments, each veh can evaluate its ability to perform the SS function on R_i , by computing the following two metrics:

- $d_{loc}(i)$ is the distance between the location of veh and the beginning of the outgoing direction of R_i , as it is illustrated by Figure. 5.2(a). Note that if veh is already on R_i , $d_{loc}(i) = 0$; otherwise $d_{loc}(i) > 0$.
- $d_\alpha(i) \in [-\pi, +\pi]$ is the angular deviation with respect to R_i direction as it is depicted in Figure. 5.2(b). If veh is already in R_i , $d_\alpha(i) = 0$; otherwise $d_\alpha(i) > 0$;

These metrics are used as inputs to compute an index I_{veh} that measures the tendency of veh to head towards R_i . We propose the formula (5.1) to calculate the index I_{veh} .

$$I_{veh}(d_{loc}(i), d_\alpha(i)) = a * \frac{d_{loc}(i)}{D_{loc}(i)} + b * \frac{d_\alpha(i)}{D_\alpha(i)} \quad (5.1)$$

Where,

$D_{loc}(i)$ and $D_\alpha(i)$ are respectively the maximum tolerated errors in distance and deviation.

a and b are weighting coefficients of the location and direction metrics contributions, with $a + b = 1$. In fact, the lower the index value, the higher the tendency to move towards R_i . A detailed study on how weighting coefficients (a and b) are set is provided in OMS [26] previously presented in chapter 4.

Since each R_i may have multiple candidates, every vehicle within SDA uses a back-off-based algorithm to advertise its involvement by sending SS_ADVERTISE message containing its identifier and R_i identifier. The back-off time $BCT_{veh}(R_i)$ is a random bounded value that depends on the vehicle's index I_{veh} . It is computed using the proposed formula (5.2).

$$BCT_{veh}(R_i) = Rand(T_s, T_{des}) + [I_{veh}(d_{loc}(i), d_\alpha(i))] * T_s \quad (5.2)$$

Where,

$Rand(A,B)$ is a function that returns a single uniformly distributed random number that is bounded between A and B . T_s is the message transmission time duration and

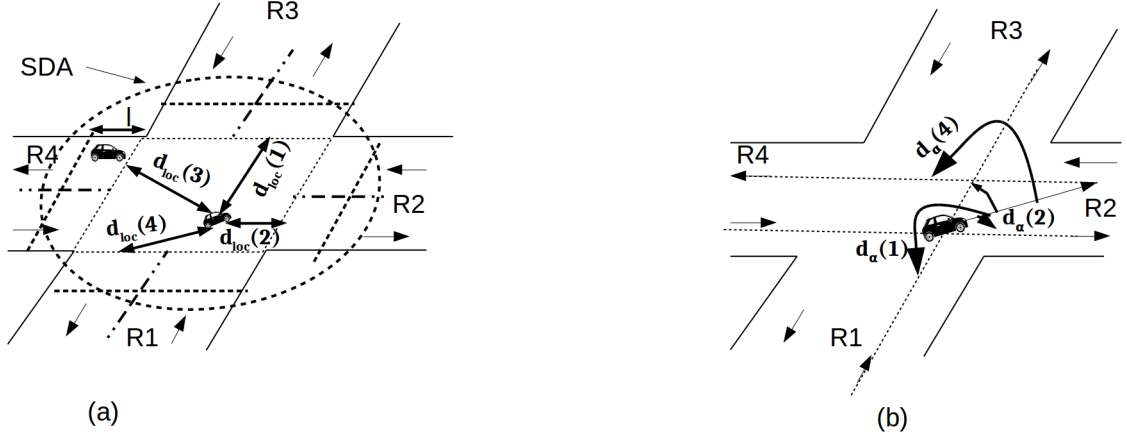


Figure 5.2: Slaves designation process.

T_{des} is the duration of slaves designation process.

After being designated as the slave on R_i , the vehicle veh collects the data of the encountered IoT sensors. Every SS can select other SSs by renewing the designation process when approaching a new intersection provided that the latter has not yet been managed. In order to inform the mobile sinks (the master and its slaves) about already managed intersections, each SS broadcasts a beacon message BEAC_INT after the reception of the first SS_ADVERTISE. The BEAC_INT message contains the intersection identifier and should be spread throughout the collection period within the Beaconsing Area (BA). We define the beaconsing area BA_{Ri} by the road portion of length LBA_{Ri} such as:

$$LBA_{Ri} = L_{Ri} - Rg \quad (5.3)$$

where L_{Ri} is the length of R_i and Rg is the communication range of a vehicle.

A vehicle that does not belong to the BA_{Ri} is not responsible of BEAC_INT re-broadcasting. Note that it is not necessary to spread BEAC_INT messages all over the region of interest since each SS is a private car that moves according to the driver's will. Therefore, SSs are only interested in intersections that belong to their paths. On the other hand, an $MasMS$ is not in need to know about already managed intersections because its movement strategy requires targeting less congested roads (in order to fasten the coverage of the region of interest with slaves) regardless of intersections state. By limiting the diffusion of BEAC_INT messages, we avoid overloading the network with additional messages.

5.4.2 Data collection

When in the communication range of an IoT node, mobile sinks (*MasMS* or *SSs*) download data from IoT sensors. Our scheme guarantees sufficient sensor-MS contact time in all cases. Indeed, the sensor data will be transmitted either to the *MasMS* which can control its speed to adjust the contact time, or to an *SS* which moves by default on congested roads at low speed allowing sufficient contact time.

5.4.3 Data Delivery

Data collected by *MasMS* and *SSs* is to be delivered to the nearest RSUs in order to transmit it to the central controller. *MasMS* is very likely to encounter more RSUs as it can sweep most of the collection area during its function of data gathering and slaves designation. For the sake of ease and reliability of data delivery, we classify *SSs* into two levels: the first level is composed of *SSs* which are designated by *MasMS* and the second level is composed of the *SSs* which are designated by the *SSs* of the first level. The first level of *SSs* are generally moving close to *MasMS* as they are directly designated by it. They can forward their data to *MasMS* using few hops. On the other hand, *SSs* designated by other *SSs* are more likely to move away from the *MasMS*. To facilitate the data delivery of this category of *SSs*, these latter forward data to RSUs using one of the delay optimal vehicular data forwarding algorithms such as [30] in order to achieve a reduced latency of the overall data collection process.

5.4.4 Flowcharts

Figure 5.3 and Figure 5.4 summarize the master and the slaves operation, respectively.

- **The master sink's actions:**

- (1): Dedicated MS asks for *SSs* for the first time.
- (2) to (5): Upon receiving the first slave involvement message, MS declares itself as the master (*MasMS*) of the area and broadcasts a beacon message to inform *SSs* of the treated intersection while moving towards the less congested road.
- (6): *MasMS* collects data of encountered IoT nodes.
- (7) to (8): If the new intersection is already managed, then *MasMS* moves towards the less congested road without asking for *SSs*.

Table 5.1: Simulation Settings

Parameter Name	Value
Size of simulation area	9.80 km ²
Traffic volume	5 – 50 veh/lane/km
Vehicle speed	18 – 50 km/h
Number of MS (dedicated vehicles)	1 veh
MAC protocol	IEEE802.11p (wave)
Propagation model	Nakagami_m
Bandwidth	6 Mbps
Simulation time	2000 Sec(33.33 min)

We first measure the ratio of roads having obtained sinks over time in the three traffic situations namely low, medium and congested. The objective of this experiment is to assess the evolution of the collection process through the evolution of slaves' designation process. Indeed, if the scheme is able to provide the largest number of roads with mobile sinks in a short time, then, it will be able to collect a large amount of data over the same period of time. Figure 5.5 depicts the ratio of roads to which mobile sinks (master or slaves) are assigned in three traffic volumes: 10 veh/lane/km for low traffic, 20 veh/lane/km for medium traffic and 40 veh/lane/km for congested traffic. The measurements were taken every 100 seconds during the simulation run time. We

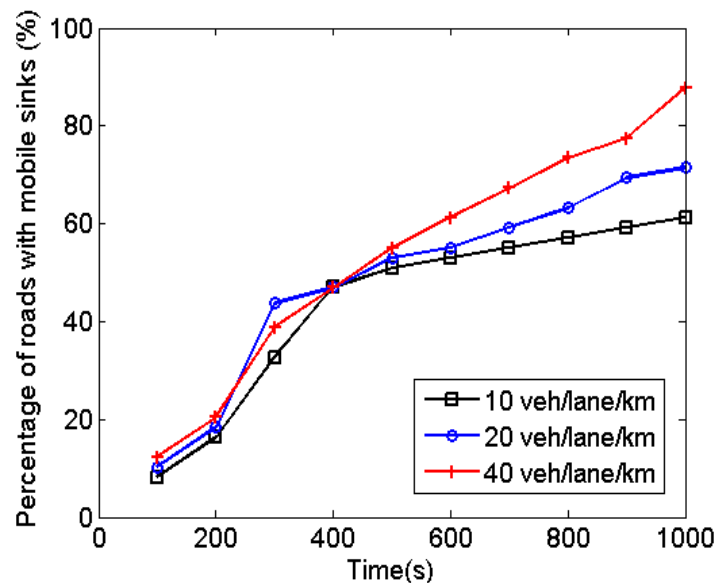


Figure 5.5: The evolution of slaves designation over time.

observe that the ratio of roads that have obtained sinks increases over time. In addition, this ratio exceeds 50% after 500 seconds (8.33 minutes) in all scenarios, indicating that in less than 9 minutes, our scheme is able to collect more than half of the data

of the region of interest regardless of traffic condition. We also note that the scheme performs better in the congestion scenario and can provide 88% of roads with sinks after 1000 seconds (16.66 minutes), what we consider as satisfactory given the length of the used map's roads (between 400m and 1500m). So, we can see that in the worst case (long roads with congestion), our scheme collected nearly 90% of the data after 16.66 minutes. This is due to the *MasMS'* movement that always takes the least crowded roads to quickly reach intersections and thus designating more slaves in a short time. Additionally, in crowded scenarios, the probability of obtaining slaves increases.

Figure 5.6 shows the percentage of roads that have obtained mobile sinks according to the diameter of the region of interest, starting from 0.5 km and increasing this area by 0.5 km in both axes (x, y) until scanning the entire area. The purpose of these measurements is to investigate the nature of the distribution of mobile sinks in order to conclude on the distribution of IoT nodes reached by our collection scheme. It should be noted that the ratio of roads having obtained mobile sinks increases almost uniformly up to a diameter of 2.5 km in the three traffic situations. Beyond this diameter, the growth rate drops but not in congested scenario which seems to be favorable since it reaches 88% after 3.5 km. We can therefore deduce that the mobile sinks are almost uniformly distributed over the region of interest, which means that our method manages to target almost all IoT nodes of the region of interest. This is due to the fact that, whatever the traffic conditions, each road is likely to obtain a mobile sink.

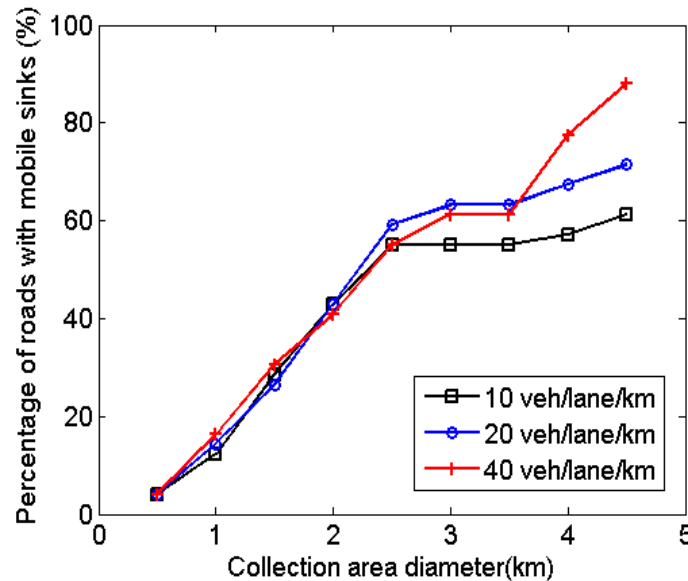


Figure 5.6: The evolution of slaves designation in space.

We then assess the impact of traffic volume on performance in terms of percentage

of accessibility to IoT nodes compared to the relevance-based scheme (RL-based) [91]. We chose RL-based because it is the most recent work close to our proposal. Indeed, both works rely on dedicated vehicles to collect sensors data in urban areas with the difference that RL-based targets sensors with higher sensing relevance to reduce the collection latency, whereas, our scheme leverages the opportunistic presence of vehicles to create a distributive data collection process. Both schemes were performed using a single dedicated vehicle. The objective of this comparison is to prove the advantage of exploiting non dedicated vehicles. The results of the simulation are shown in Figure 5.7. We find that, our scheme exceeds RL-based in all traffic volumes. For collection duration of 33.33 minutes (2000 seconds), our scheme reaches more than 70% of IoT nodes against less than 51% with RL-based. This is due to the fact that with RL-based, the shortest paths to the relevant sensors often include the same road segments, which decreases chances of MS to explore the entire region of interest. This ratio increases in our scheme with the increase of traffic volume to reach 96% in congestion situations (50 *veh/lane/km*). In fact, with the increase in traffic volume, the probability of designating *SSs* for almost all the roads increases, which increases the ratio of accessed IoT nodes. This confirms the relevance of the proposed scheme for congestion situations.

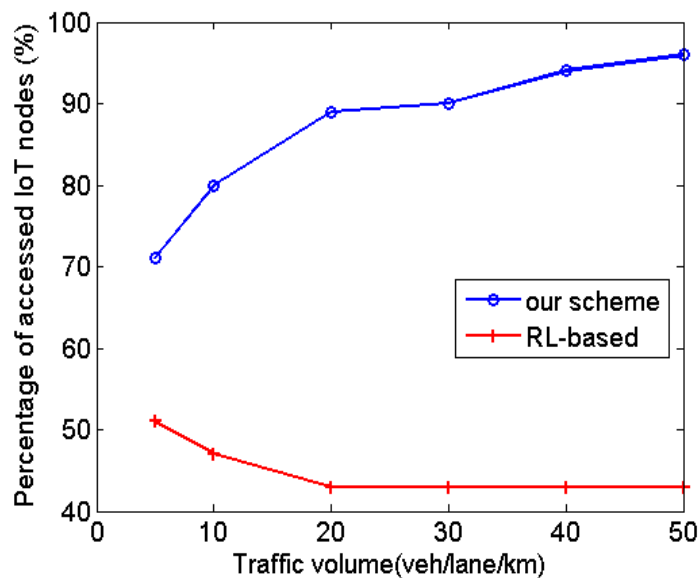


Figure 5.7: Percentage of accessibility to IoT nodes with different traffic volumes.

Figure 5.8 shows the data collection latency which corresponds to the time required to target all IoT nodes in the region of interest. As we considered collection duration of 33.33 minutes, the collection latency will represent the time required to reach the last IoT node among reachable nodes in the collection duration. We observe that our scheme allows lower collection latency compared to RL-based that consumes all the collection

duration to reach less than 51% of IoT nodes as illustrated in Figure 5.7. We note that with our scheme, data collection latency is reduced from 32.8 minutes in a very low traffic scenario ($5\text{veh}/\text{lane}/\text{km}$) to 29.9 minutes which corresponds to the delay of collecting the data of 94%-96% of IoT nodes in congested scenario ($40-50\text{veh}/\text{lane}/\text{km}$) as illustrated in Figure. 5.7. This is due to the distribution of the collection process over several collaborative sinks, unlike RL-based that relies on shortest paths to access the sensors.

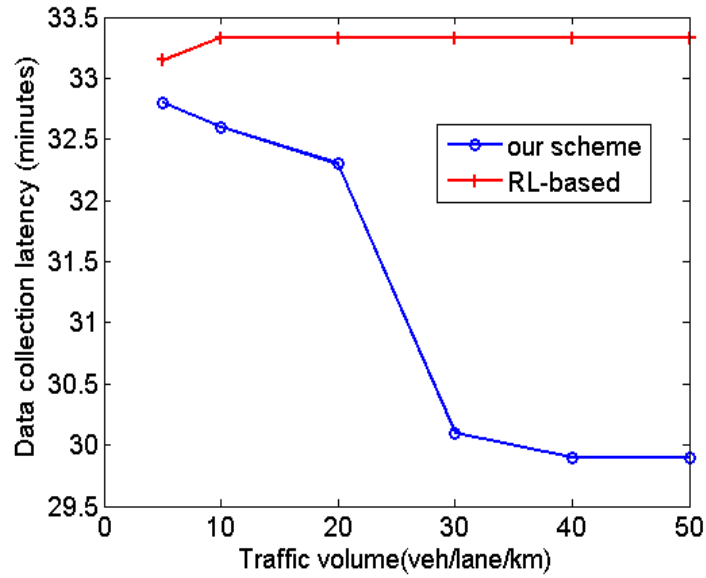


Figure 5.8: Data collection latency with different traffic volumes.

5.6 Conclusion

Using vehicles as mobile sinks to collect data from roadside IoT sensors is both cost effective and energy efficient. However, the high density of urban environments leads to high data delivery delays particularly in congestion situations. To overcome this issue, the work presented in this chapter takes benefit from the high vehicular density in congestion situations to create a collaborative data collection process aiming at reducing the collection latency. To this end, the concept of master and slave was introduced. This concept consists in supporting MS (master) with slaves which are designated among neighboring citizen cars to collect data of congested roads. Simulations using NS3 demonstrate that the scheme significantly reduces the data collection latency while increasing the reach-ability ratio of IoT sensors. We estimate that this contribution is relevant given that it allows a limitation of data collection time and takes benefit from congestion situations.

Chapter 6

Conclusion

Smart city projects involve acquiring new ways of managing cities, which generally rely on technological solutions that collect and process large amounts of city-data. As a result, WSNs and IoT objects have been widely deployed on city streets to gather information about the operation of metropolitan infrastructure. The efficiency of any smart city service depends on the quality and quantity of the data gathered. Accordingly, data collection issue is of a considerable interest for the design of smart city projects. Nevertheless, collecting urban sensors' data raises many challenges: First, network fragmentation due to the complexity of urban environment. Second, high communication cost due to the need for a permanent infrastructure to cover all the nodes. Third, the energy consumed for the data collection process is highly dependent on the quality of the required communication. Therefore, designing data collection mechanisms that consider the aforementioned challenges is of a great importance.

The objective of this work has been to study the data collection issue in vehicular sensor networks with the aim of designing efficient data collection strategies that consider urban environment characteristics. In particular, we focused on collecting the data of roadside sensors/IoT objects using the vehicle-based data collection approach. In this context, the research carried out in this thesis aims to provide smart cities with cost-effective and energy-efficient urban data collection strategies, while meeting the requirements of a wide range of application scenarios. The contributions of this thesis are summarised in Section 6.1 while perspective directions are presented in Section 6.2.

6.1 Summary of contributions

We first reviewed the related works. The objective of this investigation was to define a set of criteria for evaluating vehicle-based urban data collection works to avoid the drawbacks of existing solutions when designing new collection strategies. The considered criteria include: the type of the path of the mobile element, the length of the mobile element's path, the number of hops to the mobile element, the data collection latency, the type of the data collection process, the type of vehicles used as mobile ele-

ments, exclusive vehicles requirement, etc. We then provided in Chapter ??, to the best of our knowledge, the first taxonomy for roadside sensor data collection using vehicular networks.

In reviewing related work, we noted the effectiveness of the vehicle network-based data collection approach in terms of reduced energy consumption and deployment cost. However, we found that most proposals result in a significant delay in data delivery due to the characteristics of the urban environment. Furthermore, in analyzing the reviewed proposals, we noticed that the capabilities of the vehicular network were often underutilized and that vehicles acting as mobile elements were assigned long paths, which led to high latency. The variety of smart city scenarios drove us to address the high latency drawback of this approach in order to include a wide range of smart city applications, in particular short delay applications. To that end, we suggested, in this thesis, to take advantage of the richness of urban environment in vehicles to involve as many vehicles as possible in the collection process, as well as exploiting the mobility and connectivity capabilities of vehicular networks to obtain a better distribution of the data collection process thus, reducing its latency.

As a first contribution in this dissertation, we have proposed a vehicle-based data collection scheme that leverages the opportunistic presence of urban transport vehicles (citizen cars, busses, and taxicabs) to create a reduced-latency distributed data collection process for urban WSNs. In this way, this scheme contributes to cost-effective and energy-efficient data collection for a wide range of smart city scenarios including those that require rapid responses. The scheme design focuses on the requirements of a specific class of smart city applications, namely short delay applications, by simultaneously collecting IoT object data from different road segments of an area of interest to speed up the collection task. To that end, the scheme relies on an optimized query dissemination strategy that provides, in a short time, each road of the area of interest with a data mule designated from existing vehicles.

The query dissemination strategy is based on minimum spanning tree (MST) where only a portion of road segments of the areas of interest are used to propagate the request for data mules. This allows disseminating the query in an opportune time with low communication cost. The data mule election is opportunistically triggered at specific zones with low communication cost using only one messages in each election zone. The scheme manages to provide almost all of the roads of the area of interest with data mules in an opportune time, in medium and high vehicle density, while it achieves acceptable results in low density.

Although this scheme has effectively reduced the data collection latency from hours to minutes, it has the shortcoming that data mules are designated without considering the contribution factors of the election parameters, which can lead to inappropriate designations. Therefore, it was necessary to revise the election procedure by taking

into account the contribution factor of each election parameter. The priority order established on election parameters in the improved version of the scheme, called OMS, allows electing the best data mules in almost all scenarios compared to the primary version of the scheme. Furthermore, applying OMS to larger areas demonstrates its scalability. However, in sparse networks, such as in the nighttime scenarios, probabilities of obtaining opportunistic mules are reduced. In this regard, we have proposed a new scheme that support OMS with dedicated vehicles that will act as Mobile Sinks in sparse areas and delegate opportunistic sinks to collect data in congested spaces.

The new scheme leverages the high vehicular density in congestion situations to create a collaborative data collection process aiming at reducing the collection latency. To this end, we have introduced the concept of master and slave which consists in supporting the MS (master) with slaves which are designated among neighboring citizen cars to collect data of congested roads. The scheme can significantly reduce the data collection latency while increasing the reach-ability ratio of IoT sensors. In addition it provides the following advantages:

- Reduction of the collection cost by reducing the number of dedicated vehicles.
- Reduction of collection latency thanks to the concept of opportunistic designation of slave sinks for congested roads which makes it possible on the one hand to increase the distribution of the collection process and on the other hand to devote dedicated vehicles to less congested roads.

We estimate that this contribution is relevant given that it allows a limitation of data collection time and takes benefit from congestion situations.

The schemes proposed in this thesis are abstracted from any specific smart city configuration. In addition each scheme can complement the other to provide a reduced-latency data collection strategy suitable to all traffic situations and network connectivity levels (spars, medium, and high).

As a result of our work, we found that data collection in smart cities is a complex problem. A black-box scheme with a multipurpose collection strategy that covers the requirements of most smart city scenarios in terms of response time, energy consumption, and implementation cost is not feasible. To the best of our knowledge, our work in the context of this thesis is the first contribution that deals with the high collection latency of the vehicle-based approach to design cost effective and energy efficient data collection solutions for short delay smart city applications. Our work is not intended to include a complete data collection solution capable of collecting the overall IoT objects data in any smart city scenario. Nevertheless, smart city administrators can combine the schemes proposed in this thesis to adapt them to the particular cases of their cities.

6.2 Perspectives

Following the same interest in vehicle-based approach, our future work will include:

- Designing a distributed propagation strategy that makes use of local traffic information to deliver the request for data mules to election zones.
- Designing a dedicated routing strategy for data delivery phase that makes use of the previous OMS phases (query propagation, data harvesting) to estimate the current state of traffic in order to deduce the optimal routes for forwarding collected data to the BS.
- Decreasing delays in order to improve the reached performances by exploring multiple cooperative data mules in the same segment.
- Investigating on slaves designation near IoT nodes locations by leveraging inter-vehicular communications in order to provide a road segment with multiple slaves, thereby taking advantage of the congestion situation to reduce the collection latency and therefore, improve the application response time. This will meet the requirements of other smart city scenarios, especially those that require rapid responses.
- Designing adaptable and optimal delay data forwarding algorithms for data delivery from slave sinks to Master sinks or to RSUs.
- Adapting the proposed schemes to the vehicular fog computing paradigm by deploying fog stations in the proximity of vehicles to reduce the delivery delay of data carried by data mules and slave sinks.
- Adapting the system to disaster situations using artificial intelligence techniques to provide vehicles with appropriate algorithms that allow them to learn about the environment around them. This allows vehicles to be equipped with anomaly detection capabilities and thus to initiate collaborative data collection in a reactive manner.

Bibliography

- [1] Network simulator ns3. <<https://www.nsnam.org/ns-3-26/download>>.
- [2] Openstreetmap. <https://www.openstreetmap.org/>.
- [3] Simulation of urban mobility. <<http://www.sumo.dlr.de/daily/userdoc/Downloads.html>>.
- [4] F. Adelantado, X. Vilajosana, P. Tuset-Peiro, B. Martinez, J. Melia-Segui, and T. Watteyne. Understanding the limits of lorawan. *IEEE Communications Magazine*, 55(9):34–40, September 2017.
- [5] M. Ahmad, Q. Chen, Z. Khan, M. Ahmad, and F. Khurshid. Infrastructure-based vehicular congestion detection scheme for v2i. *International Journal of Communication Systems, Wiley*, 32(3):1–14, December 2018.
- [6] K. Akkaya and M. Younis. Energy-aware routing to a mobile gateway in wireless sensor networks. *In Proceedings of the 47 th IEEE Global Telecommunications Conference Workshops (GlobeCom), Dallas, TX, USA*, page 16–21, November–December 2004.
- [7] M. Al-Hader, A. Rodzi, A. R. Sharif, and N. Ahmad. Smart city components architecture. *In Proc. of International Conference on Computational Intelligence, Modelling and Simulations, Brno, Czech Republic*, pages 93–97, September 2009.
- [8] F. Al-Turjman and J. P. Lemayian. Intelligence, security, and vehicular sensor networks in internet of things (iot)-enabled smart-cities: An overview. *Computers and Electrical Engineering*, 87:1–13, October 2020.
- [9] H. Ammari and S. Das. Data dissemination to mobile sinks in wireless sensor networks: an information theoretic approach. *In Proc. of the 2 nd IEEE International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, pages 8 pp.–314, November 2005.
- [10] G. Anastasi, E. Borgia, M. Conti, and E. Gregori. A hybrid adaptive protocol for reliable data delivery in wsns with multiple mobile sinks. *The Computer Journal*, 54(2):213–229, February 2011.

-
- [11] G. Anastasi, M. Conti, and M. Di Francesco. Data collection in sensor networks with data mules: an integrated simulation analysis. *In Proc. of the 13 th IEEE Symposium on Computers and Communications (ISCC), Marrakech, Morocco*, page 1096–1102, July 2008.
- [12] G. Anastasi, M. Conti, M. Di Francesco, and A. Passarella. Energy conservation in wireless sensor networks: A survey. *Ad Hoc Networks*, 7(3):537–568, may 2009.
- [13] G. Anastasi, M. Conti, E. Gregori, C. Spagoni, and G Valente. Motes sensor networks in dynamic scenarios: an experimental study for pervasive applications in urban environments. *International Journal of Ubiquitous Computing and Intelligence*, 1(1):9–16, April 2007.
- [14] J. Ansari, D. Pankin, and P. Mahonen. Radio-triggered wake-ups with addressing capabilities for extremely low power sensor network applications. *In Proc. of the 19 th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Cannes, France*, page 1–5, September 2008.
- [15] M. Aşçı, C. U. İleri, and O. Dağdeviren. An energy-efficient capacitated minimum spanning tree algorithm for topology control in wireless sensor networks. *In proc. of IEEE 25th Signal Processing and Communications Applications Conference (SIU), Antalya, Turkey*, pages 1–4, May 2017.
- [16] P. Baruah, R. Urgaonkar, and B. Krishnamachari. Learning-enforced time domain routing to mobile sinks in wireless sensor fields. *In Proc. of the 29 th IEEE International Conference on Local Computer Networks (LCN), Tampa, FL, USA*, page 525–532, November 2004.
- [17] A. Bazzi, B. M. Masini, and G. Zanella, A. and Pasolini. Ieee 802.11 p for cellular offloading in vehicular sensor networks. *Computer Communications*, 60:97–108, February 2015.
- [18] M. Bonola, L. Bracciale, P. Loreti, R. Amici, A. Rabuffi, and G. ppeBianchi. Opportunistic communication in smart city: Experimental insight with small-scale taxi fleets as data carriers. *Ad Hoc Networks, Elsevier*, 43(43):43–55, February 2016.
- [19] D. Borsetti, C. Casetti, C. F. Chiasserini, M. Fiore, and J. M. Barceló-Ordinas. Virtual data mules for data collection in road-side sensor networks. *In Proc. of the Second International Workshop on Mobile Opportunistic Networking, ACM, (MobiOpp '10)*, pages 32–40, February 2010.
- [20] J. O. Briony. Researching information systems and computing. *Sage*, 2005.

- [21] J. L. Burbank, J. Andrusenko, J. S. Everett, and W. T. M. Kasch. Wireless metropolitan area networks. *Wireless Networking: Understanding Internetworking Challenges, First Edition, Chapter 5, John Wiley Sons, Inc*, pages 195–249, 2013.
- [22] N. Cenerario, T. Delot, and S. Ilarri. A content-based dissemination protocol for vanets: Exploiting the encounter probability. *IEEE Transactions on Intelligent Transportation Systems*, 12(3):771–782, September 2011.
- [23] M. Chahal and S. Harit. A stable and reliable data dissemination scheme based on intelligent forwarding in vanets. *International Journal of Communication Systems, Wiley*, 32(3):851–864, December 2018.
- [24] A. Chakrabarti, A. Sabharwal, and B. Aazhang. Using predictable observer mobility for power efficient design of sensor networks. In: *Zhao F., Guibas L. (eds) Information Processing in Sensor Networks. IPSN 2003. Lecture Notes in Computer Science*, 2634:129–145, April 2003.
- [25] F. Chehbour, Z Doukha, S. Moussaoui, and M. Guerroumi. Congestion aware data collection with mobile sinks in smart city. *2020 International Symposium on Networks, Computers and Communications (ISNCC)*, pages 1–7, October 2020.
- [26] F. Chehbour, Z Doukha, S. Moussaoui, and M. Guerroumi. Oms: Opportunistic mules for short latency data collection in smart cities. *International Journal of Communication Systems, Wiley*, 33(10):e4207, July 2020.
- [27] Z. Moussaoui S. Chehbour, F. Douka and M. Guerroumi. Opportunistic data mules for short delay smart city applications. In *proc. of IEEE 2018 International Conference on Smart Communications in Network Technologies (SaCoNeT), eloued, Algeria*, pages 126–131, October 2018.
- [28] C. Cheng, C. Tse, and F. C. M. Lau. A delay-aware data collection network structure for wireless sensor networks. *IEEE Sensors Journal*, 3(11):699 – 710, June 2011.
- [29] M. Chochul and P. Ševčík. A survey of low power wide area network technologies. In *Proc. of the 18th IEEE International Conference on Emerging eLearning Technologies and Applications (ICETA), Košice, Slovenia*, pages 69–73, November 2020.
- [30] S. Jeong J. lee H. W. Choi, O. Kim and S. Chong. Delay-optimal forwarding in vehicular sensor networks. *IEEE transactions on vehicular technology*, 65(8):6389–6402, August 2016.

-
- [31] M. P. Deisenroth, A. A. Faisal, and C. S. Ong. Mathematics for machine learning. *Cambridge University Press, book chapter*, 2020.
- [32] M. Di Francesco, K. D. Sajal, and G. Anastasi. Data collection in wireless sensor networks with mobile elements: A survey. *ACM Transactions on wireless Sensor Networks*, 8(1):1–31, August 2011.
- [33] M. Di Francesco, K. Shah, M. Kumar, and G. Anastasi. An adaptive strategy for energy-efficient data collection in sparse wireless sensor networks. *In: Silva J.S., Krishnamachari B., Boavida F. (eds) Wireless Sensor Networks. EWSN. Lecture Notes in Computer Science*, 5970:322–337, April 2010.
- [34] E. W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik, Elsevier*, 1(1):269–271, December 1959.
- [35] S. Djahel, R. Doolan, G. Muntean, and J. Murphy. A communications-oriented perspective on traffic management systems for smart cities: Challenges and innovative approaches. *IEEE Communications Surveys Tutorials*, 17(1):125–151, January 2015.
- [36] R. Du, C. Chen, B. Yang, N. Lu, X. Guan, and X. Shen. effective urban traffic monitoring by vehicular sensor networks. *IEEE Transactions on Vehicular Technology*, 64(1):273–286, January 2015.
- [37] E. Ekici, Y. Gu, and D. Bozdog. Mobility-based communication in wireless sensor networks. *IEEE Communications Magazine*, 44(7):56–62, September 2006.
- [38] J. E. Fadul, K. M. Hopkinson, T. R. Andel, and C. A. Sheffield. A trust-management toolkit for smart-grid protection systems. *IEEE Transactions on Power Delivery*, 29(4):1768–1779, August 2014.
- [39] M. C. Ferreira, T. G. Dias, and J. F. Cunha. Is bluetooth low energy feasible for mobile ticketing in urban passenger transport? *Transportation Research Interdisciplinary Perspectives*, 5(5):1–10, May 2020.
- [40] V. G. Font. Anomaly detection in smart city wireless sensor networks. *PH.D. thesis in the doctoral program in network and information technologies, Open University of Catalonia*, February 2017.
- [41] V. G. Font, C. Garrigues, and H. R. Pous. Attack classification schema for smart city wsns. *Sensors*, 4(17):1–24, April 2017.
- [42] S. Gao, H. Zhang, and S. K. Das. Efficient data collection in wireless sensor networks with path-constrained mobile sinks. *IEEE Transactions on Mobile Computing*, 10(4):592–608, April 2010.

- [43] R. A. Gorcitz, Y. Jarma, P. Spathis, M. Dias De Amorim, R. Wakikawa, J. Whitbek, V. Conan, and S Fdida. Vehicular carriers for big data transfers. *In Proc. of the IEEE Vehicular Networking Conference (VNC), Seoul, Korea (South)*, pages 1–8, November 2012.
- [44] R. L. Graham and P. Hell. On the history of the minimum spanning tree problem. *IEEE Annals of the history of computing*, 7(1):43–57, January 1985.
- [45] L. Gu and J. Stankovic. Radio-triggered wake-up for wireless sensor networks. *Real-Time Systems Journal*, 29:157–182, March 2005.
- [46] Z. J. Haas and T. Small. A new networking model for biological applications of ad hoc sensor networks. *IEEE/ACM Transactions on Networking (TON)*, 14(1):27–40, February 2006.
- [47] L. He, Z. Chen, and J. D. Xu. Optimizing data collection path in sensor networks with mobile elements. *International Journal of Automation and Computing, Springer Link*, 8(1):69–77, February 2011.
- [48] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva. Directed diffusion for wireless sensor networking. *IEEE/ACM Transactions on Networking*, 11(1):2–16, February 2003.
- [49] S. Jain, R. Shah, W. Brunette, and S. Borriello, G.and Roy. Exploiting mobility for energy efficient data collection in wireless sensor networks. *IACM/Springer Mobile Networks and Applications*, 11(3):327–339, April 2006.
- [50] J. Ji, S.and He and Z. Cai. Data gathering in wireless sensor networks. *Amari H. (eds) The Art of Wireless Sensor Networks. Signals and Communication Technology. Springer, Berlin, Heidelberg.*, 1:535–565, 2014.
- [51] N. Jindaprakai and S. Nuchitprasitchai. Intelligent parking system using multiple sensor detection. *In Proc. of the IEEE 2019 Research, Invention, and Innovation Congress (RI2C), Bangkok, Thailand*, pages 1–4, December 2019.
- [52] P. Juang, H. Oki, M. Wang, Y.and Martonosi, L. Peh, and D. Rubenstein. Energy-efficient computing for wildlife tracking: Design tradeoffs and early experiences with zebranet. *ACM SIGPLAN Notices*, 37(10):96–107, October 2002.
- [53] H. Jun and E. Ammar, M.and Zegura. Power management in delay tolerant networks: A framework and knowledge-based mechanisms. *In Proc. of the 2nd IEEE Conference on Sensor and Ad Hoc Communications and Networks (SECON), Santa Clara, CA, USA*, page 418–429, September 2005.

- [54] A. Kansal, A. Somasundara, D. Jea, M. Srivastava, and D. Estrin. Intelligent fluid infrastructure for embedded networks. *In Proc. of the 2nd ACM International Conference on Mobile Systems, Applications, and Services (MobiSys)*, page 111–124, June 2004.
- [55] M. Khabazian, S. Aissa, and M. Mehmet-Ali. Performance modeling of message dissemination in vehicular ad hoc networks with priority. *IEEE Journal on Selected Areas in Communications*, 29(1):61–71, January 2011.
- [56] H. S. Kim, T. F. Abdelzaher, and W. H. Kwon. Minimum-energy asynchronous dissemination to mobile sinks in wireless sensor networks. *In Proc. of the 1st ACM Conference on Embedded Networked Sensor Systems (SenSys), Los Angeles, California, USA*, page 193–204, November 2003.
- [57] H. S. Kim, T. F. Abdelzaher, and W. H. Kwon. Dynamic delay-constrained minimum energy dissemination in wireless sensor networks. *ACM Transactions on Embedded Computing Systems*, 4(3):679–706, August 2005.
- [58] J. Kim, X. Lin, S. B. Shroff, and P. Sinha. Minimizing delay and maximizing lifetime for wireless sensor networks with anycast. *IEEE/ACM Transactions on Networking*, 18(2):515–528, April 2010.
- [59] S. Kim, S. Pakzad, D. Culler, J. Demmel, G. Fenves, S. Glaser, and M. Turon. Health monitoring of civil infrastructures using wireless sensor networks. *In Proc. of the 6th International Symposium on Information Processing in Sensor Networks, Cambridge, MA, USA*, page 254–263, April 2007.
- [60] C. Konstantopoulos, G. Pantziou, D. Gavalas, and B. Mamalis. A rendezvous-based approach enabling energy-efficient sensory data collection with mobile sinks. *IEEE Transactions on Parallel and Distributed Systems, Transportation Research Board of the National Academies*, 23(5):809–817, May 2012.
- [61] C. Konstantopoulos, G. Pantziou, D. Gavalas, A. Mpitziopoulos, and M. Basilis. A rendezvous-based approach enabling energy-efficient sensory data collection with mobile sinks. *IEEE Transactions on Parallel and Distributed Systems, Springer*, 23(5):809–817, May 2012.
- [62] P. Kortoçi, L. Zheng, C. Joe-Wong, M. Di Francesco, and M. Chiang. Fog-based data offloading in urban iot scenarios. *In proc. of IEEE INFOCOM-IEEE Conference on Computer Communications, Paris, France*, pages 784–792, April 2019.

- [63] J. B. Kruskal. On the shortest spanning sub tree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society*, 7(7):48–50, November 1956.
- [64] T.M.Vinod. Kumar and B. Dahiya. Smart economy in smart cities. *In book: Smart Economy in Smart Cities. Edition: 1st. Chapter: 1. Publisher: Springer*, pages 3–76, 2017.
- [65] Broadband Communications Research Laboratory. Montreal smart-city pilot system (mcsps). *Technical report, Department of Electrical Computer Engineering, McGill University*, July 2017.
- [66] Y. Lalle, L. C. Fourati, M. Fourati, and J. P. Barraca. A comparative study of lorawan, sigfox, and nb-iot for smart water grid. *In Proc. of the 2019 IEEE Global Information Infrastructure and Networking Symposium (GIIS), Paris, France*, pages 1–6, December 2019.
- [67] K. C. Lan and Z. M. Wu. On the feasibility of using public transport as data mules for traffic monitoring. *IEEE Intelligent Vehicles Symposium*, 2634:979–984, April 2008.
- [68] F. Leccese, M. Cagnetti, and D. Trinca. A smart city application: A fully controlled street lighting isle based on raspberry-pi card, a zigbee sensor network and wimax. *Sensors*, 12(14):24408–24424, December 2014.
- [69] A. Lee, H. Keshavarzian and H. Aghajian. Near-lifetime-optimal data collection in wireless sensor networks via spatio-temporal load balancing. *ACM Transactions on Sensor Networks*, 6(3):1–32, June 2010.
- [70] H. Li. An energy efficeint routing algorithm for heterogenuous wireless sensor networks. *In proc. of IEEE International Conference on Computer Application and System Modeling (ICCASM 2010), Taiyuan, China*, pages 612–616, October 2010.
- [71] K. W. Lim and Y. B. Ko. Multi-hop data harvesting in vehicular sensor networks. *IET Communications*, 4(7):768–775, May 2010.
- [72] Inc. LinkLabs. A comprehensive look at low power, wide area networks for internet of things engineers and decision makers. *Tech. rep. LinkLabs, Inc.*, 2016.
- [73] J. Liu, Z. Yang, and I. Stojmenovic. Receiver consensus: on-time warning delivery for vehicular ad-hoc networks. *EEE transactions on emerging topics in computing*, 1(1):57–686, July 2013.

- [74] X. Liu and P. Zhang. Data drainage: A novel load balancing strategy for wireless sensor networks. *IEEE Commun. Lett.*, 22(1):125–128, January 2018.
- [75] H. Luo, F. Ye, J. Cheng, S. Lu, and L. Zhang. Ttdd: two-tier data dissemination in large-scale wireless sensor networks. *Wireless Networks*, 11:161–175, January 2005.
- [76] J. Luo, J. Panchard, M. Piorkowski, M. Grossglauser, and J.P. Hubaux. Mobicroute: Routing towards a mobile sink for improving lifetime in sensor networks. *In Proc. of the 2nd IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS)*, page 480–497, May 2006.
- [77] P. Luoto, M. Bennis, P. Pirinen, S. Samarakoon, K. Horneman, and M. Latvaaho. System level performance evaluation of lte-v2x network. *In Proc. of the 22th European wireless conference. VDE*, page 1–5, May 2016.
- [78] L. Marsal. Measuring the standardized definition of “smart city”: A proposal on global metrics to set the terms of reference for urban “smartness”. *In Proc. of International Conference on Computational Science and Its Applications (ICCSA)*, pages 593–611, June 2015.
- [79] J. Marín, J. Rocher, L. Parra, S. Sendra, J. Lloret, and P. V. Mauri. Autonomous wsn for lawns monitoring in smart cities. *In Proc. of the 5th International workshop on Big Data and Social Networking Management and Security, Hammamet, Tunisia*, pages 501–508, Oct 30 - Nov 3 2017.
- [80] C. Meering and P. Balella. Smart cities and the internet of things. municipal transformation with the hpe universal iot platform. *Tech. rep. Hewlett Packard Enterprise Development LP*, 2016.
- [81] A. Mehto, S. Tapaswi, and K. K. Pattanaik. A review on rendezvous based data acquisition methods in wireless sensor networks with mobile sink. *Wireless Networks, Springer*, 26(4):2639–2663, May 2020.
- [82] B. Miles, E. B. Bourennane, S. Boucherkha, and S. Chikhi. A study of lorawan protocol performance for iot applications in smart agriculture. *Computer Communications*, 164:148–157, December 2020.
- [83] L.H.M.K. Mitton, N.and Costa and B. et al Krishnamachari. Green data collection and processing in smart cities. *Annals of Telecommunications*, 75(75):269–270, July 2020.
- [84] T. Nam and T. A. Pardo. Conceptualizing smart city with dimensions of technology, people, and institutions. *In Proc. of the 12th Annual International Con-*

- ference on Digital Government Research, College Park, Maryland, USA*, pages 282–291, June 2011.
- [85] M. Naphade, C. Banavar, G. Harrison, J. Paraszczak, and R. Morris. Smarter cities and their innovation challenges. *Computer*, 44(6):32–39, June 2011.
- [86] P. Neirotti, A. De Marco, A. C. Cagliano, G. Mangano, and F. Scorrano. Current trends in smart city initiatives: Some stylized facts. *Cities, Elsevier*, 38(38):25–36, June 2014.
- [87] The Royal Academy of Engineering. Smart infrastructure: the future. *Tech. rep. The Royal Academy of Engineering, available online at www.raeng.org.uk/smartinfrastructure*, page 16–17, January 2012.
- [88] K. Il. Park. Fundamentals of probability and stochastic processes with applications to communications. *Springer, book chapters*, pages 1–275, 2018.
- [89] L. Parra, S. Sendra, J. Lloret, and I. Bosch. Development of a conductivity sensor for monitoring groundwater resources to optimize water management in smart city environments. *Sensors*, 15(9):20990–21015, August 2015.
- [90] H. Pavel and W. Daniel. On the communication complexity of secure function evaluation with long output. *In Proc. of the 2015 conference on innovations in theoretical computer science. ACM*, page 163–172, January 2015.
- [91] J.P.J. Peixoto and D.G. Costa. Wireless visual sensor networks for smart city applications: A relevance-based approach for multiple sinks mobility. *Future Generation Computer Systems, Elsevier*, 76(76):51–62, November 2017.
- [92] R. C. Prim. Shortest connection networks and some generalizations. *Bell system technical journal, Nokia Bell Labs*, 36(6):1389–1401, November 1957.
- [93] J. Rao, T. Wu, and S. Biswas. Network-assisted sink navigation protocols for data harvesting in sensor networks. *In Proc. of the 2008 IEEE Conference on Wireless Communications and Networking WCNC, Atlanta, GA, USA*, page 2887–2892, May 2008.
- [94] A. Antonio. Reshme. Delay and energy efficient data collection in wireless sensor networks. *In Proc. of the 2019 International Conference on Recent Advances in Energy-efficient Computing and Communication (ICRAECC), Nagercoil, India*, pages 1–6, March 2019.
- [95] K. Rostamzadeh and S. Gopalakrishnan. Analysis of message dissemination in vehicular networks. *IEEE Transactions on Vehicular Technology*, 62(8):3974–3982, May 2013.

- [96] K. Rostamzadeh and S. Gopalakrishnan. Analysis of message delivery delay in vehicular networks. *IEEE Transactions on Vehicular Technology*, 64(10):4770–4779, October 2015.
- [97] P. M. Santos, J. G. P. Rodrigues, S. B. Cruz, T. Lourenço, P. M. d’Orey, Y. Luis, C. Rocha, S. Sousa, S. Crisóstomo, C. Queirós, S. Sargento, A. Aguiar, and J. Barros. Portolivinglab: An iot-based sensing platform for smart cities. *IEEE Internet of Things Journal*, 5(2):523–532, January 2018.
- [98] T. Guimarães D. Condeixa T. Sargento S. Aguiar A. Barros J. Santos, P.M. calada. Demo: Platform for collecting data from urban sensors using vehicular networking. *In Proc. of the 21st Annual International Conference on Mobile Computing and Networking (MobiCom ’15)*, page 167–169, September 2015.
- [99] L. Sarakis, T. Orphanoudakis, H. C. Leligou, S. Voliotis, and A. Voulkidis. Providing entertainment applications in vanet environments. *IEEE Wireless Communications*, 23(1):30–37, July 2016.
- [100] C. Schurgers, V. Tsiatsis, S. Ganeriwal, and M. B. Srivastava. Optimizing sensor networks in the energy-latency-density design space. *IEEE Transactions on Mobile Computing*, 1(1):70–80, January-March 2002.
- [101] R. C. Shah, S. Roy, S. Jain, , and W. Brunette. Data mules: Modeling a three-tier architecture for sparse sensor networks. *In Proc. of the 2nd ACM International Workshop on Wireless Sensor Networks and Applications (SNPA), Anchorage, AK, USA*, page 30–41, May 2003.
- [102] S. A. A. Shah, E. Ahmed, M. Imran, and S. Zeadally. 5g for vehicular communications. *IEEE Communication Magazine*, 56(1):111–117, January 2018.
- [103] P. K. Singh. Influences of tworayground and nakagami propagation model for the performance of adhoc routing protocol in vanets. *International Journal of Computer Applications*, 45(22):1–6, August 2012.
- [104] A. Somasundara, A. Kansal, D. Jea, D. Estrin, and M. Srivastava. Controllably mobile infrastructure for low energy embedded networks. *IEEE Transactions on Mobile Computing*, 5(8):1536–1233, August 2006.
- [105] R. Sugihara and R.K. Gupta. Path planning of data mules in sensor networks. *ACM Transactions on Sensor Networks (TOSN), Volume: 8 Issue: 1, pp. 1–27*, 8(1):1–27, June 2011.
- [106] C. Wang, H. Ma, Y. He, and S. Xiong. Approximate data collection for wireless sensor networks. *In Proc. of the 2010 IEEE 16th International Conference on*

- Parallel and Distributed Systems, Shanghai, China*, pages 164–171, December 2010.
- [107] F. Wang, D. Wang, and J. Liu. Traffic-aware relay node deployment: maximizing lifetime for data collection wireless sensor networks. *IEEE Trans. Parallel Distrib. Syst.*, 8(22):1415–1423, August 2011.
- [108] F.Y. Wang. Scanning the issue and beyond: Transportation and mobility transformation for smart cities. *IEEE Transactions on Intelligent Transportation Systems*, 16(2):525–533, June 2015.
- [109] J. Wang, C. Jiang, K. Zhang, T. Q. S. Quek, Y. Ren, and L. Hanzo. Vehicular sensing networks in a smart city: Principles, technologies and applications. *IEEE Wireless Communications Magazine*, 25(1):122–132, February 2018.
- [110] Z. M. Wang, S. Basagni, E. Melachrinoudis, and C. Petrioli. Exploiting sink mobility for maximizing sensor networks lifetime. *In Proc. of the 38 th Hawaii International Conference on System Sciences*, pages 287a–287a, January 2005.
- [111] A. Woo, T. Tong, and D. Culler. Taming the underlying challenges of reliable multi-hop routing in sensor networks. *In Proc. of the 1 st ACM Conference on Embedded Networked Sensor Systems (SenSys), Los Angeles, California, USA*, page 14–27, November 2003.
- [112] X. Yang and N. Vaidya. A wakeup scheme for sensor networks: Achieving balance between energy saving and end-to-end delay. *In Proc. of the 10 th IEEE Real-Time and Embedded Technology and Applications Symposium, Toronto, ON, Canada*, page 19–26, May 2004.
- [113] D. Ye and M. Zhang. A self-adaptive sleep/wake-up scheduling approach for wireless sensor networks. *IEEE Transactions on Cybernetics*, 48(3):979–992, March 2018.