REPUBLIQUE DU CAMEROUN
Paix – Travail – Patrie
****
UNIVERSITE DE YAOUNDE I
FACULTE DES SCIENCES
DEPARTEMENT DE INFORMATIQUE
****

REPUBLIC OF CAMEROUN
Peace – Work – Fatherland
****
UNIVERSITY OF YAOUNDE I
FACULTY OF SCIENCE
DEPARTMENT OF COMPUTER SCIENCES
****

Autonomous Wireless Sensor Network Architecture for Vehicular traffic monitoring at an intersection

PHD THESIS operated within
University of Yaoundé 1, POST-GRADUATE SCHOOL OF SCIENCE, TECHNOLOGY AND GEOSCIENCE
In international joint supervision with
INSA Lyon, DOCTORAL SCHOOL InfoMaths

Par : **DOMGA KOMGUEM Rodrigue**

Sous la direction de
Nathalie MITTON
INRIA, Reviewer

Année Académique : 2019 - 2020

PHD THESIS

operated within

University of Yaoundé 1, POST-GRADUATE SCHOOL OF SCIENCE, TECHNOLOGY AND GEOSCIENCE

In international joint supervision with

INSA Lyon, DOCTORAL SCHOOL InfoMaths

Speciality : COMPUTER SCIENCE

Defended By :

DOMGA KOMGUEM Rodrigue

On :

July 06th, 2020

Autonomous Wireless Sensor Network Architecture for Vehicular traffic monitoring at an intersection

---

Composition of the Jury:

- **Marcel FOUDA NDJODO,** Professor, Université of Yaoundé 1, President
- **André-Luc BEYLOT,** Professor, ENSEEIHT Toulouse, Examinator
- **René NDOUNDAM,** Associate Professor, Université of Yaoundé 1, Examinator
- **Bernard TOURANCHEAU,** Professor, Université Grenoble Alpes, Reviewer
- **Nathalie MITTON,** Research Director, INRIA, Reviewer
- **Thomas DJOTIO,** Associate Professor, Université of Yaoundé 1, Reviewer
- **Fabrice VALOIS,** Professor, INSA Lyon, Co-Supervisor
- **Maurice TCHUENTE,** Professor, Université of Yaoundé I, Co-Supervisor
- **Razvan STANICA,** HDR, INSA Lyon, Guest

REPUBLIQUE DU CAMEROUN
Paix – Travail – Patrie
- . - . - . -
UNIVERSITÉ DE YAOUNDÉ I
Faculté des Sciences
Département d'Informatique
B.P. 812 Yaoundé

REPUBLIC OF CAMEROON
Peace – Work – Fatherland
- . - . - . -
UNIVERSITY OF YAOUNDÉ I
Faculté des Sciences
Department of Computer Science
P.O.Box 812 Yaoundé

ATTESTATION DE CORRECTION DE LA THÈSE DE DOCTORAT/PHD

Nous soussignés, Professeurs **FOUDA NDJODO Marcel, TCHUENTE Maurice, NDOUNDAM Réné**, membres du jury de la thèse de Doctorat/PhD présentée par Monsieur **DOMGA KOMGUEM Rodrigue**, matricule **04U282**, thèse intitullée :

« **Autonomous WSN architecture for vehicular traffic monitoring at an intersection** » et soutenue en vue de l'obtention du diplôme de **Doctorat/PhD en Informatique**, attestons que toutes les corrections demandées par le jury de soutenance en vue de l'amélioration de ce travail, ont été effectuées.

En fois de quoi, la présente attestation lui est délivrée pour servir et valoir ce que de droit.

Président	**Rapporteur**	**Examinateur**
Pr FOUDA NDJODO Marcel	**Pr. TCHUENTE Maurice**	**Pr. NDOUNDAM René**

The University of Yaounde I

Faculty of science

Division of Programming and follow-up of Academic Affaires

Université de Yaoundé I

Faculté des sciences

Division de la Programmation et du Suivi des Activités Académiques

List of permanent teaching staff | **Liste des enseignants permanents**

ACADEMIC YEAR 2019/2020

(By Department and by Grade)

UPDATING DATE: The 19th of Febuary, 2019

ADMINISTRATION

DEAN: TCHOUANKEU Jean Claude, Professor

VICE-DEAN DPSAA: DONGO Etienne, Professor

VICE-DEAN DSSE: AJEAGAH Gideon AGHAINDUM, Professor

VICE-DEAN DRC: ABOSSOLO Monique, Professor

Head of the Division of Academic Affairs, Scolarity and Research: MBAZE MEVA'A Luc, Professor

Head of the Administrative and Financial Division: NDOYE FOE Marie C. F., Associate professor

1- DEPARTMENT OF BIOCHEMISTRY (BC)(40)			
N^o	Names and Surnames	Grade	Observations
1	FEKAM BOYOM Fabrice	Professor	In service
2	MBACHAM FON Wilfried	Professor	In service
3	MOUNDIPA FEWOU Paul	Professor	Head of Department
4	NINTCHOM PENLAP V. épse BENG	Professor	In service
5	OBEN Julius ENYONG	Professor	In service
6	ACHU Merci BIH	Associate professor	In service
7	ATOGHO Barbara Mma	Associate professor	In service
8	BELINGA née NDOYE FOE M. C. F.	Associate professor	Head DAF/FS
9	BIGOGA DIAGA Jude	Associate professor	In service
10	BOUDJEKO Thaddée	Associate professor	In service
11	EFFA NNOMO Pierre	Associate professor	In service
12	FOKOU Elie	Associate professor	In service
13	KANSCI Germain	Associate professor	In service

14	NANA Louise épouse WAKAM	Associate professor	In service
15	NGONDI Judith Laure	Associate professor	In service
16	NGUEFACK Julienne	Associate professor	In service
17	NJAYOU Frédéric Nico	Associate professor	In service
18	AKINDEH MBUH NJI	Lecturer	In service
19	AZANTSA KINGUE GABIN BORIS	Lecturer	In service
20	BEBOY EDZENGUELE Sara Nathalie	Lecturer	In service
21	DAKOLE DABOY Charles	Lecturer	In service
22	DJOKAM TAMO Rosine	Lecturer	In service
23	DJUIDJE NGOUNOUE Marcelline	Lecturer	In service
24	DJUIKWO NKONGA Ruth Viviane	Lecturer	In service
25	DONGMO LEKAGNE Joseph Blaise	Lecturer	In service
26	EWANE Cécile Anne	Lecturer	In service
27	FONKOUA Martin	Lecturer	In service
28	BEBEE Fadimatou	Lecturer	In service
29	KOTUE KAPTUE Charles	Lecturer	In service
30	LUNGA Paul KEILAH	Lecturer	In service
31	MANANGA Marlyse Joséphine	Lecturer	In service
32	MBONG ANGIE M. Mary Anne	Lecturer	In service
33	MOFOR née TEUGWA Clotilde	Lecturer	IA4/MINESUP
34	PACHANGOU NSANGOU Sylvain	Lecturer	In service
35	Palmer MASUMBE NE-TONGO	Lecturer	In service
36	TCHANA KOUATCHOUA Angle	Lecturer	In service
37	MBOUCHE FANMOE Marceline Joëlle	Assistant lecturer	In service
2- DEPARTMENT OF BIOLOGY AND ANIMAL PHYSIOLOGIES (B.P.A.) (44)			
1	BILONG BILONG Charles-Félix	Professor	Head of Department
2	DIMO Théophile	Professor	In service
3	DJIETO LORDON Champlain	Professor	In service
4	ESSOMBA née NTSAMA MBALA	Professor	VDRC/FMSB/UYI
5	FOMENA Abraham	Professor	In service
6	KAMGANG René	Professor	C.E. MINRESI

7	KAMTCHOUING Pierre	Professor	In service
8	NJAMEN Dieudonné	Professor	In service
9	NJIOKOU Flobert	Professor	In service
10	NOLA Mose	Professor	In service
11	TAN Paul VERNYUY	Professor	In service
12	TCHUEM TCHUENTE Louis Albert	Professor	Coord. Progr. MINSANTE
13	AJEAGAH Gideon AGHAINDUM	Associate professor	VICE-DOYEN/DSSE
14	DZEUFIET DJOMENI Paul Désiré	Associate professor	In service
15	FOTO MENBOHAN Samuel	Associate professor	In service
20	JATSA BOUKENG Hermine épse MEGAPTCHE	Associate professor	In service
16	KEKEUNOU Sévilor	Associate professor	In service
17	MEGNEKOU Rosette	Associate professor	In service
18	MONY Ruth épse NTONE	Associate professor	In service
19	NGUEGUIM TSOFACK Florence	Associate professor	In service
21	TOMBI Jeannette	Associate professor	In service
22	ZEBAZE TOGOUET Serge Hubert	Associate professor	In service
23	ALENE Désirée Chantal	Lecturer	In service
24	ATSAMO Albert Donatien	Lecturer	In service
25	BELLET EDIMO Oscar Roger	Lecturer	In service
26	BILANDA Danielle Claude	Lecturer	In service
27	DJIOGUE Séfirin	Lecturer	In service
28	DONFACK Mireille	Lecturer	In service
29	GOUNOUE KAMKUMO Raceline	Lecturer	In service
30	KANDEDA KAVAYE Antoine	Lecturer	In service
31	LEKEUFACK FOLEFACK Guy B.	Lecturer	In service
32	MAHOB Raymond Joseph	Lecturer	In service
33	MBENOUN MASSE Paul Serge	Lecturer	In service
34	MOUNGANG Luciane-Marlyse	Lecturer	In service
35	MVEYO NDANKEU Yves Patrick	Lecturer	In service
36	NGOUATEU KENFACK Omer Bébé	Lecturer	In service
37	NGUEMBOK	Lecturer	In service
38	NJUA Clarisse Yafi	Lecturer	Head Div. UBA

39	NOAH EWOTI Olive Vivien	Lecturer	In service
40	TADU Zephyrin	Lecturer	In service
41	YEDE	Lecturer	In service
43	ETEME ENAMA Serge	Assistant lecturer	In service
44	KOGA MANG DOBARA	Assistant lecturer	In service
3- DEPARTMENT OF BIOLOGY AND VEGETAL PHYSIOLOGY (B.P.V.) (27)			
1	AMBANG Zachée	Professor	Head Division/UYII
2	BELL Joseph Martin	Professor	In service
3	MOSSEBO Dominique Claude	Professor	In service
4	YOUMBI Emmanuel	Professor	Head of Department
5	ZAPFACK Louis	Professor	In service
6	ANGONI Hyacinthe	Associate professor	In service
7	BIYE Elvire Hortense	Associate professor	In service
8	DJOCGOUE Pierre Franois	Associate professor	In service
9	KENGNE NOUMSI Ives Magloire	Associate professor	In service
10	MALA Armand William	Associate professor	In service
11	MBARGA BINDZI Marie Alain	Associate professor	CT/UDs
12	MBOLO Marie	Associate professor	In service
13	NDONGO BEKOLO	Associate professor	CE / MINRESI
14	NGONKEU MAGAPTCHE Eddy L.	Associate professor	In service
15	TSOATA Esae	Associate professor	In service
16	GOMANDJE Christelle	Lecturer	In service
17	MAFFO MAFFO Nicole Liliane	Lecturer	In service
18	MAHBOU SOMO TOUKAM. Gabriel	Lecturer	In service
19	NGALLE Hermine BILLE	Lecturer	In service
20	NGOUO Lucas Vincent	Lecturer	In service
22	NOUKEU KOUAKAM Armelle	Lecturer	In service
23	ONANA JEAN MICHEL	Lecturer	In service
24	NSOM ZAMO Annie Claude épse PIAL	Lecturer	Expert national/UNESCO
25	TONFACK Libert Brice	Lecturer	In service
26	DJEUANI Astride Carole	Assistant lecturer	In service
27	NNANGA MEBENGA Ruth Laure	Assistant lecturer	In service
4- DEPARTMENT OF INORGANIC CHEMISTRY (C.I.) (35)			
1	AGWARA ONDOH Mose	Professor	Vice Rector/Univ, Bamenda
2	ELIMBI Antoine	Professor	In service

3	Florence UFI CHINJE épouse MELO	Professor	Rector Univ.Ngaoundere
4	GHOGOMU Paul MINGO	Professor	Ministre Chargé de Miss.PR
5	NANSEU Njiki Charles Péguy	Professor	In service
6	NDIFON Peter TEKE	Professor	CT MINRESI/Head of Department
7	NDIKONTAR Maurice KOR	Professor	Vice-Dean Univ. Bamenda
8	NENWA Justin	Professor	In service
9	NGAMENI Emmanuel	Professor	DEAN/FS/UDs
10	BABALE née DJAM DOUDOU	Associate professor	Chargée Mission P.R.
11	DJOUFAC WOUMFO Emmanuel	Associate professor	In service
12	KAMGANG YOUBI Georges	Associate professor	In service
13	KEMMEGNE MBOUGUEM Jean C.	Associate professor	In service
14	KONG SAKEO	Associate professor	In service
16	NGOMO Horace MANGA	Associate professor	Vice Chancelor/UB
17	NJIOMOU C. épse DJANGANG	Associate professor	In service
18	NJOYA Dayirou	Associate professor	In service
19	YOUNANG Elie	Associate professor	In service
20	ACAYANKA Elie	Lecturer	In service
21	BELIBI BELIBI Placide Désiré	Lecturer	CS/ENS Bertoua
22	CHEUMANI YONA Arnaud M.	Lecturer	In service
23	EMADACK Alphonse	Lecturer	In service
24	KENNE DEDZO GUSTAVE	Lecturer	In service
24	KOUOTOU DAOUDA	Lecturer	In service
25	MAKON Thomas Beauregard	Lecturer	In service
26	MBEY Jean Aime	Lecturer	In service
27	NCHIMI NONO KATIA	Lecturer	In service
28	NDI NSAMI Julius	Lecturer	In service
29	NEBA nee NDOSIRI Bridget NDOYE	Lecturer	Inspecteur de Service MINFEM
30	NYAMEN Linda Dyorisse	Lecturer	In service
31	PABOUDAM GBAMBIE A.	Lecturer	In service
32	TCHAKOUTE KOUAMO Hervé	Lecturer	In service
5- DEPARTMENT OF ORGANIC CHEMISTRY (C.O.) (33)			
1	DONGO Etienne	Professor	Vice-Dean/PSAA

2	GHOGOMU TIH Robert Ralph	Professor	Dir. IBAF/UDS
3	NGOUELA Silvre Augustin	Professor	In service
4	NKENGFACK Augustin Ephrem	Professor	Head of Department
5	NYASSE Barthélemy	Professor	Director/UN
6	PEGNYEMB Dieudonné Emmanuel	Professor	Director/MINESUP
7	WANDJI Jean	Professor	In service
8	Alex de Théodore ATCHADE	Associate professor	DEPE/Rectorat/UYI
9	EYONG Kenneth OBEN	Associate professor	Head Service DPER
10	FOLEFOC Gabriel NGOSONG	Associate professor	In service
11	KEUMEDJIO Félix	Associate professor	In service
12	KEUMOGNE Marguerite	Associate professor	In service
13	KOUAM Jacques	Associate professor	In service
14	MBAZOA née DJAMA Céline	Associate professor	In service
15	MKOUNGA Pierre	Associate professor	In service
16	NGO MBING Joséphine	Associate professor	Sous/Direct. MINERESI
17	NOUNGOUE TCHAMO Diderot	Associate professor	In service
18	TABOPDA KUATE Turibio	Associate professor	In service
19	TCHOUANKEU Jean-Claude	Associate professor	Dean/FS/UYI
20	TIH née NGO BILONG E. Anastasie	Associate professor	In service
21	YANKEP Emmanuel	Associate professor	In service
22	AMBASSA Pantaléon	Lecturer	In service
23	FOTSO WABO Ghislain	Lecturer	In service
24	KAMTO Eutrophe Le Doux	Lecturer	In service
25	MVOT AKAK CARINE	Lecturer	In service
26	NGOMO Orléans	Lecturer	In service
27	NGONO BIKOBO Dominique Serge	Lecturer	In service
28	NOTE LOUGBOT Olivier Placide	Lecturer	Head Service/MINESUP
29	OUAHOUO WACHE Blandine M.	Lecturer	In service
30	TAGATSING FOTSING Maurice	Lecturer	In service
31	ZONDENDEGOUMBA Ernestine	Lecturer	In service
32	NGNINTEDO Dominique	Assistant lecturer	In service
6- DEPARTMENT OF COMPUTER SCIENCE (IN) (28)			
1	ATSA ETOUNDI Roger	Professor	Chef Div.MINESUP

2	FOUDA NDJODO Marcel Laurent	Professor	Chef Dpt ENS/Chef IGA.MINESUP
3	NDOUNDAM Réné	Associate professor	In service
4	AMINOU Halidou	Lecturer	In service
5	DJAM Xaviera YOUHEP KIMBI	Lecturer	In service
6	KOUOKAM KOUOKAM E. A.	Lecturer	In service
7	MELATAGIA YONTA Paulin	Lecturer	In service
8	MOTO MPONG Serge Alain	Lecturer	In service
9	TAPAMO Hyppolite	Lecturer	In service
10	ABESSOLO ALOO Gislain	Lecturer	In service
11	KAMGUEU Patrick Olivier	Lecturer	In service
12	MONTHE DJIADEU Valery M.	Lecturer	In service
13	OLLE OLLE Daniel Claude Delort	Lecturer	C/D Enset. Ebolowa
14	TINDO Gilbert	Lecturer	In service
15	TSOPZE Norbert	Lecturer	In service
16	WAKU KOUAMOU Jules	Lecturer	In service
17	BAYEM Jacques Narcisse	Assistant lecturer	In service
18	DOMGA KOMGUEM Rodrigue	Assistant lecturer	In service
19	EBELE Serge	Assistant lecturer	In service
20	HAMZA Adamou	Assistant lecturer	In service
21	JIOMEKONG AZANZI Fidel	Assistant lecturer	In service
22	KAMDEM KENGNE Christiane	Assistant lecturer	In service
23	MAKEMBE. S. Oswald	Assistant lecturer	In service
24	MEYEMDOU Nadge Sylvianne	Assistant lecturer	In service
25	NKONDOCK. MI. BAHANACK.N.	Assistant lecturer	In service
7- DEPARTMENT OF MATHEMATICS (MA) (31)			
1	BITJONG NDOMBOL	Professor	In service
2	DOSSA COSSY Marcel	Professor	In service
3	AYISSI Raoult Domingo	Associate professor	Head of Department
4	EMVUDU WONO Yves S.	Associate professor	CD Info/Chef division MINESUP
5	NKUIMI JUGNIA Célestin	Associate professor	In service
6	NOUNDJEU Pierre	Associate professor	In service
7	TCHAPNDA NJABO Sophonie B.	Associate professor	Director/AIMS Rwanda
8	AGHOUKENG JIOFACK Jean Gérard	Lecturer	Head Cellule MINPLAMAT

9	CHENDJOU Gilbert	Lecturer	In service
10	DJIADEU NGAHA Michel	Lecturer	In service
11	DOUANLA YONTA Herman	Lecturer	In service
12	FOMEKONG Christophe	Lecturer	In service
13	KIANPI Maurice	Lecturer	In service
14	KIKI Maxime Armand	Lecturer	In service
15	MBAKOP Guy Merlin	Lecturer	In service
16	MBANG Joseph	Lecturer	In service
17	MBEHOU Mohamed	Lecturer	In service
18	MBELE BIDIMA Martin Ledoux	Lecturer	In service
19	MENGUE MENGUE David Joe	Lecturer	In service
20	NGUEFACK Bernard	Lecturer	In service
21	NIMPA PEFOUNKEU Romain	Lecturer	In service
22	POLA DOUNDOU Emmanuel	Lecturer	In service
23	TAKAM SOH Patrice	Lecturer	In service
24	TCHANGANG Roger Duclos	Lecturer	In service
25	TCHOUNDJA Edgar Landry	Lecturer	In service
26	TETSADJIO TCHILEPECK M. E.	Lecturer	In service
27	TIAYA TSAGUE N. Anne-Marie	Lecturer	In service
28	MBIAKOP Hilaire George	Assistant lecturer	In service
8- DEPARTMENT OF MICROBIOLOGY (MB) (13)			
1	ESSIA NGANG Jean Justin	Professor	DRV/IMPM
2	ETOA Franois Xavier	Professor	Head of Department/FS/UYI; Recteur Université de Douala
3	BOYOMO ONANA	Associate professor	In service
4	NWAGA Dieudonné M.	Associate professor	In service
5	NYEGUE Maximilienne Ascension	Associate professor	In service
6	RIWOM Sara Honorine	Associate professor	In service
7	SADO KAMDEM Sylvain Leroy	Associate professor	In service
8	ASSAM ASSAM Jean Paul	Lecturer	In service
9	BODA Maurice	Lecturer	In service
10	BOUGNOM Blaise Pascal	Lecturer	In service
11	ESSONO OBOUGOU Germain G.	Lecturer	In service
12	NJIKI BIKO Jacky	Lecturer	In service
13	TCHIKOUA Roger	Lecturer	In service

9- DEPARTMENT OF PHYSICS (PH) (41)			
1	BEN- BOLIE Germain Hubert	Professor	In service
2	ESSIMBI ZOBO Bernard	Professor	In service
3	KOFANE Timoléon Crépin	Professor	In service
4	NDJAKA Jean Marie Bienvenu	Professor	Head of Department
5	NJANDJOCK NOUCK Philippe	Professor	Sous Directeur/ MINRESI
6	NJOMO Donatien	Professor	In service
7	PEMHA Elkana	Professor	In service
8	TABOD Charles TABOD	Professor	Dean Univ/Bda
9	TCHAWOUA Clément	Professor	In service
10	WOAFO Paul	Professor	In service
11	BIYA MOTTO Frédéric	Associate professor	DG/HYDRO Mekin
12	BODOBertrand	Associate professor	In service
13	DJUIDJE KENMOE épouse ALOYEM	Associate professor	In service
14	EKOBENA FOUDA Henri Paul	Associate professor	Head Division. UN
15	EYEBE FOUDA Jean sire	Associate professor	In service
16	FEWO Serge Ibrad	Associate professor	In service
17	HONA Jacques	Associate professor	In service
18	MBANE BIOUELE César	Associate professor	In service
19	NANA ENGO Serge Guy	Associate professor	Director/Students/Affairs. UB
20	NANA NBENDJO Blaise	Associate professor	In service
21	NOUAYOU Robert	Associate professor	In service
22	SAIDOU	Associate professor	Sous Director/Minresi
23	SIEWE SIEWE Martin	Associate professor	In service
24	SIMO Elie	Associate professor	In service
25	VONDOU Derbetini Appolinaire	Associate professor	In service
26	WAKATA née BEYA Annie	Associate professor	Sous Directeur/MINESUP
27	ZEKENG Serge Sylvain	Associate professor	In service
28	ABDOURAHIMI	Lecturer	In service
29	EDONGUE HERVAIS	Lecturer	In service
30	ENYEGUE A NYAM épse BELINGA	Lecturer	In service
31	FOUEDJIO David	Lecturer	Head Cell. MINADER
32	MBINACK Clément	Lecturer	In service
33	MBONO SAMBA Yves Christian U.	Lecturer	In service
34	MELII Joelle Larissa	Lecturer	In service
35	MVOGO ALAIN	Lecturer	In service
36	NDOP Joseph	Lecturer	In service
37	OBOUNOU Marcel	Lecturer	DA/Univ Inter Etat/Sangmalima

38	WOULACHE Rosalie Laure	Lecturer	In service
39	CHAMANI Roméo	Assistant lecturer	In service
10- DEPARTMENT OF EARTH SCIENCES (S.T.) (44)			
1	BITOM Dieudonné	Professor	Dean/FASA/UDs
2	FOUATEU Rose épse YONGUE	Professor	In service
3	KAMGANG Pierre	Professor	In service
4	MEDJO EKO Robert	Professor	CT/UYII
5	NDJIGUI Paul Désiré	Professor	Head of Department
6	NKOUMBOU Charles	Professor	In service
7	NZENTI Jean-Paul	Professor	In service
8	ABOSSOLO née ANGUE Monique	Associate professor	Vice-Doyen/DRC
9	GHOGOMU Richard TANWI	Associate professor	CD/UMa
10	MOUNDI Amidou	Associate professor	CT/MINIMDT
11	NDAM NGOUPAYOU Jules-Remy	Associate professor	In service
12	NGOS III Simon	Associate professor	DAAC/Uma
13	NJILAH Isaac KONFOR	Associate professor	In service
14	ONANA Vincent Laurent	Associate professor	In service
15	BISSO Dieudonné	Associate professor	Director/Projet Barrage Memveele
16	EKOMANE Emile	Associate professor	In service
17	GANNO Sylvestre	Associate professor	In service
18	NYECK Bruno	Associate professor	In service
19	TCHOUANKOUE Jean-Pierre	Associate professor	In service
20	TEMDJIM Robert	Associate professor	In service
21	YENE ATANGANA Joseph Q.	Associate professor	Head Div. /MINTP
22	ZOO ZAME Philémon	Associate professor	DG/ART
23	ANABA ONANA Achille Basile	Lecturer	In service
24	BEKOA Etienne	Lecturer	In service
25	ELISE SABABA	Lecturer	In service
26	ESSONO Jean	Lecturer	In service
27	EYONG JOHN TAKEM	Lecturer	In service
28	FUH Calistus Gentry	Lecturer	Sec. D'Etat/MINMIDT
29	LAMILEN BILLA Daniel	Lecturer	In service
30	MBESSE CECILE OLIVE	Lecturer	In service
31	MBIDA YEM	Lecturer	In service
32	METANG Victor	Lecturer	In service
33	MINYEM Dieudonné-Lucien	Lecturer	CD/Uma
34	MOUAFO Lucas	Lecturer	In service
35	NGO BELNOUN Rose Noël	Lecturer	In service

37	NGO BIDJECK Louise Marie	Lecturer	In service
38	NGUEUTCHOUA Gabriel	Lecturer	CEA/MINRESI
39	NOMO NEGUE Emmanuel	Lecturer	In service
36	NTSAMA ATANGANA Jacqueline	Chargé de Cours	In service
40	TCHAKOUNTE J. épse NOUMBEM	Lecturer	Head Cell./MINRESI
41	TCHAPTCHET TCHATO De P.	Lecturer	In service
42	TEHNA Nathanaël	Lecturer	In service
43	TEMGA Jean Pierre	Lecturer	In service

Numbered repartition of permanent teachers by Department

(The 19th of Febuary, 2019)

Department	**Number of teachers**				
	Pr	**AP**	**L**	**ASS. L**	**Total**
BCH	5 (1)	12 (6)	19 (11)	1 (1)	37 (19)
BPA	12 (1)	10 (5)	20 (07)	2 (0)	44 (13)
BPV	5 (0)	10(2)	9 (04)	2(02)	26 (9)
CI	9(1)	9(2)	14 (3)	0 (0)	32 (6)
CO	7 (0)	14 (4)	10 (4)	1 (0)	32(8)
IN	2 (0)	1 (0)	13 (0)	10 (3)	26 (3)
MAT	2 (0)	4 (1)	19 (1)	2 (0)	27 (2)
MIB	2 (0)	5 (2)	5 (1)	0 (0)	12 (3)
PHY	10 (0)	17(2)	11 (3)	1 (0)	39 (5)
ST	7 (1)	15 (1)	21 (5)	1 (0)	43(7)
Total	**61 (4)**	**97 (25)**	**141 (39)**)	**19(6)**	**318 (75)**

A total of: 318 (75) whose
- Professors 61 (4)
- Associate professors 97(25)
- Lecturers 141 (39)
- Assistant lecturers 18(5)
- ( ) = Number of women.

The dean of the Faculty of science

To DOMGA TCHIGO Gabrielle, DOMGA KOMGUEM Daniel and DOMGA GUEDIA Malia.
Receive it as a gift of your family.

Acknowledgements

Pursuing a PhD is undoubtedly an important evolutionary process for my career. It would not have been possible to complete without the guidance and support from many people and organizations.

I would like to express my sincere and immense gratitude to my supervisors Pr. Maurice TCHUENTE, Pr. Fabrice VALOIS and Dr. Razvan STANICA. I am deeply grateful for the time, the efforts, and the advises they gave me during these years. I would especially like to thank them for their rigor which will be very important for me in my career. Pr. TCHUENTE introduced me in the research since my master and have never stopped to support me. For this, receive my gratitude. I would like to thank Pr. Fabrice VALOIS for welcoming me in the CITI Laboratory, and providing support for many years, which resulted first in my master and then in this thesis. Thank to Dr. Razvan for his patience and his generosity. He always knew what to tell me during difficult times. I am grateful.

I gratefully acknowledge the funding received from the AUF ("Agence Universitaire de la Francophonie") under the program "Horizons Francophones Sciences Fondamentales" during the years 2012 to 2015. I also thank the University of Yaoundé I, INSA Lyon, INRIA, LIRIMA, IRD, UMMISCO for their support during these years.

I thank the University of Yaoundé I for the training and guidance I received. I especially thank the Department of Computer Science and its lectures who gave me all the knowledge necessary to engage in a PhD thesis. I would also like to thank all my colleagues from the University of Yaoundé I who helped and encouraged me throughout this thesis.

During these years, I made several research stays at INSA Lyon, within the CITI laboratory. I would like to thank the CITI laboratory, INSA Lyon, INRIA and particularly the Agora team for all the support provided to the realization of this thesis. I also thank the members of the Agora team and the administration of the CITI laboratory for the warm welcome they gave me during each of my stays with them.

I address my thanks to all the jury members for accepting and evaluating the works which have been done in this thesis.

Finally, I would like to thank my whole family for encouraging and supporting me patiently throughout this journey. Particularly, to Gerardine, Gabrielle, Daniel and Malia, thank you for always being present. You have given a meaning not only to this thesis, but also to all my projects.

Abstract

In many countries, because of the limited financial budget, the growth of road infrastructures is low compared to the growth of population and the number of vehicles in urban areas. Such a context does not make the task easy for authorities in charge of the management of transportation systems. The introduction of information and communication technologies (ICT) allows to better address these issues. Vehicular traffic management at intersections has an impact on the traffic jam observed in the whole city. In this thesis, our goal is to propose a low-cost, lightweight and autonomous Wireless Sensors Network (WSN) architecture for vehicular traffic monitoring, especially at an intersections. Vehicular traffic data collected can be used, for instance, for intelligent traffic lights management. In the WSN architecture proposed in the literature for vehicular traffic monitoring, underground sensors are used. In terms of network communication, these architectures are not realistic. Nowadays, surface-mounted sensors are proposed by manufacturers.

The first contribution of this thesis is an experimental characterization of wireless links in a WSN with sensors deployed at the ground level. We evaluate the impact of several parameters like the proximity of the ground surface, the communication frequency and the messages size on the link quality. Result show a poor link quality at ground level. Based on the conclusions of the experiments, the second contribution of this thesis is WARIM, a new WSN architecture for vehicular traffic monitoring at an intersection. In WARIM, the sensors deployed on a lane form a multi-hop WSN with a linear topology (LWSN). In this network, all the data are forwarded toward the sink. In a network with such properties, the computation and communication requirements are highest in the neighborhood of the sink. Thus, the third contribution of this thesis is a virtual nodes-based and energy efficient sensors deployment strategy for LWSN. Compared to a uniform deployment, this deployment improves the network lifetime by 40%. In our intersection monitoring application, it is important to correlate the messages generated by a sensor to its position with respect to the intersection. Therefore,the fourth contribution of this thesis is, a centroid-based algorithm for sensors ranking in a LWSN. We evaluate the performance of this algorithm considering a realistic channel model, a uniform deployment, as well as the virtual nodes based-deployment proposed in this thesis. Finally, putting all our contributions together, simulations show that WARIM can be used for reliable and real-time vehicular traffic monitoring at an intersection.

Keywords: Intelligent Transportation System, Vehicular Traffic Monitoring, Linear Wireless Sensor Network, Sensors Deployment, Sensors Ranking, Experiment

Résumé

Dans plusieurs pays à travers le monde, à cause des moyens financiers qui sont le plus souvent limités, la croissance des infrastructures de transport est généralement faible comparée non seulement à celle de la populations en zone urbaine, mais aussi à celle du parc automobile. Un tel contexte ne rend pas la tâche facile aux autorités en charge de la gestion des systèmes de transport. L'introduction des technologies de l'information et de la communication a permis à ces autorités de mieux adresser ce problème. Dans les centres urbains, la gestion du trafic véhiculaire aux intersections à un impact sur la congestion dans toute la ville. Dans cette thèse, notre objectif est de proposer une architecture de réseau de capteurs sans fil pour mesurer le trafic véhiculaire aux intersections. L'architecture proposée doit non seulement avoir un coût financier raisonnable, mais doit également être légère et autonome. Les architectures proposées dans la littérature utilisent des capteurs déployés sous la chaussée. D'un point de vu communication réseau, ces architectures sont irréalistes.

La première contribution de cette thèse est la caractérisation des liens radio dans un réseau avec des capteurs déployés en surface du sol. Les résultats montrent un lien radio de mauvaise qualité au sol. Partant des conclusions de ce premier travail, la seconde contribution de cette thèse est WARIM, une nouvelle architecture de réseau de capteurs sans fil pour le monitoring du trafic véhiculaire au niveau des intersections. Dans WARIM, les capteurs déployés sur une voie forment un réseaux de capteurs sans fil, multi saut et ayant une topologie physique linéaire. Dans ce réseau, toutes les données sont relayées en direction d'un mme point de collecte. Dans un tel réseau, les besoins en terme de traitement et de communication sont plus élevés dans le voisinage du point de collecte. Ainsi, la troisième contribution de cette thèse est, considérant un réseau de capteur linéaire, une stratégie de déploiement des capteurs basée sur le concept de noeuds virtuelles. Comparée a‘ un déploiement uniforme, la solution que nous proposons prolonge la durée de vie du réseau de l'ordre de 40%. Dans l'application de monitoring du trafic véhiculaire considérée dans cette thèse, il est important de corréler les messages générés par un capteur à sa position. La quatrième contribution de cette thèse est, considérant un réseau de capteur linéaire, un algorithme de classement des capteurs basé sur la centroide. Nous évaluons les performances de cet algorithme en considérant un modèle de communication réaliste, un déploiement uniforme, et aussi le déploiement proposé dans cette thèse. Finalement, en mettant ensemble toutes nos contributions, des simulations montrent que WARIM, l'architecture que nous avons proposée dans cette thèse, peut être utilisée pour une collecte fiable et en temps réel du trafic véhiculaire au niveau d'une intersection.

Mots clés: Systèmes de Transports Intelligents, Monitoring du trafic véhiculaire, Réseaux de capteurs sans fil linéaires, Déploiement des Capteurs, Classement des capteurs, Expérimentation

Contents

List of Figures

List of Tables

List of Algorithms

ACSLite	Adaptive Control Software Lite
ATLMS	Adaptive Traffic Light Management System
CDF	Cumulative Distribution Function
CO	Carbon Monoxyde
CSMA/CA	Carrier Sense Multiple Access with Collision Advoidance
FCD	Floating Car Data
GPS	Global Positioning System
ICT	Information and Communication Technologies
IEEE	Institute of Electrical and Electronics Engineers
ITS	Intelligent Transportation System
LoRa	Long Range
LPWAN	Low Power Wide Area Network
LQI	Link Quality Indicator
LWSN	Linear Wireless Sensor Network
MAC	Medium Access Control
MEMS	Micro-Electro-Mechanical System
MILP	Mixed Integer Linear Programming
MWW	Mann-Whitney-Wilcoxon
MSE	Mean Square Error
NRZ	Non-Return-to-Zero
OPAC	Optimized Policies for Adaptive Control
OS	Operating System
PDF	Probability Distribution Function
PDR	Packet Delivery Rate
PRR	Packet Reception Rate
QPSK	Quadrature Phase Shift Keying
RAM	Random Access Memory
RSSI	Receive Signal Strength Indicator
RF	Radio Frequence
SCATS	Sydney Coordinated Adaptive Traffic System
SCOOT	Split, Cycle and Offset Optimization Technique
UDG	Unit Disk Graph
UN	United Nation
WiFi	Wireless Fidelity
WSN	Wireless Sensor Network

Chapter 1

Introduction

1.1 Context

In the United Nations (UN) World Urbanization Prospect, the percentage of the population living in the urban area (urban population) was 36% in 1970 [1]. This percentage has increased to reach more than half of the World population in 2011 and will be more than 67% in 2050. The developing countries are the most concerned with the increase of urban population, which went from 0.68 billion (25% of the population) in 1970 to 2.67 billion (47% of the population) in 2011 and will be 5.12 billion (64% of the population) in 2050.

At the same time, the number of vehicles has increased exponentially but, in many countries, the capacities of roads and transportation systems have not increased in an equivalent way to efficiently cope with the number of vehicles traveling on them. Due to this, road jamming and traffic correlated pollution have increased with the associated adverse societal and financial effect on different markets worldwide [1, 2]. Another consequence is the increase of the travel time and fuel consumption for the same distances.

Since the early 1980s, with the evolution of the technology in various sectors, transport management authorities have introduced Information and Communication Technologies (ICT) in transport issues management: it is the beginning of the development of Intelligent Transport Systems (ITS) [3]. Traffic control is at the heart of ITS. Using the data collected concerning the vehicular traffic allows to provide various services that enable smoother, safer, and environmentally friendly transportation [5]. Examples of such services are: vehicles counting, monitoring of road infrastructures (tunnels, bridges, etc), messages diffusion (security alerts, prevention, etc), variable speed limits, route guidance, automatic tolls and adaptive traffic lights management.

Advances in Micro-Electro-Mechanical Systems (MEMS) technology, wireless communications, and digital electronics have enabled the development of low-cost, low-power, sensor nodes that are small in size and communicate over short distances. These components, deployed in large numbers in an area, communicating with each other to offer a given service, form a type of network called Wireless Sensor Networks (WSN) [6, 7]. The applications of WSN are numerous. Among the most significant are those related to the areas of health, home automation, smart cities (transport, pollution, electrical networks, etc.), industrial applications and environmental

monitoring. Thanks to their relatively low-cost, WSN with sensors able to detect a stop or a moving vehicle are widely used in ITS.

Energy and communication are the strongest constraints in WSN [6, 7]. Firstly, sensors are equipped with batteries of reduced capacity, which limits their functioning time. Thus, sensors energy must be efficiently used. Secondly, the wireless channel is negatively affected by the environment (obstacles, interferences, etc). Moreover, because of the limited energy, sensors usually communicate at low-power. Considering the environment constraints, the communication range of sensors is limited. In order to allow two distant sensors to communicate, a multi-hop mechanism is required. These constraints are taken into account by researchers when they develop communication protocols or others services for WSN.

1.2 The goal of the thesis

In cities, traffic jams are most of the time observed at intersections. Thus, intelligently managing traffic in these areas may reduce the negative impact of traffic jams. In this thesis, our goal is to propose a low-cost, lightweight and autonomous WSN architecture for vehicular traffic monitoring at intersections. We are interested in a WSN architecture which may be adopted by many municipalities worldwide, particularly those with limited financial budget. The proposed architecture must also be implemented by operators without a technical training in WSN. The node energy must be efficiently used in order to allow to the network to function as long as possible. Thus, an energy efficient and simple deployment strategy must be proposed. To reduce the maintenance cost, self-configuration and self-organization protocols must be developed. Finally, the limited energy of nodes and the communication constraints in WSN must be taken into account by solutions proposed in this thesis.

In the WSN architectures proposed in the literature for vehicular traffic monitoring, underground sensors are used [8, 9, 10, 11]. In terms of network communication, these architectures are not realistic. Nowadays, surface-mounted sensors are proposed by manufacturers [12].

1. Our first contribution is, considering surface-mounted sensors, an experimental characterization of wireless links in a WSN with physical nodes deployed at the ground level. **We evaluate the impact of several parameters, like the proximity of the ground surface, the communication frequency and the messages size.**

2. Based on the observations from the experiments, the second contribution of this thesis is a **new WSN architecture for vehicular traffic monitoring at an intersection.** In this architecture, several sensors are deployed on each lane. The distance between two nodes depends on the deployment environment and the radio module properties.

3. The sensors deployed on a lane form a category of WSN named Linear WSN (LWSN). This is a WSN in which nodes are deployed in order to form a linear physical topology. In this network, we have a multi-hop and convergecast communication pattern. In a network with such properties, the data traffic intensity is highest in the neighborhood of the sink.

Thus, our third contribution is a **virtual nodes-based and energy efficient nodes deployment strategy for LWSN.** This deployment significantly improves the network lifetime compared to a uniform deployment.

4. It is important to correlate the messages generated by a sensor to its position at the intersection. Indeed, if a sensor reports a problem (e.g. an unauthorized stop), it is interesting for the controller or the traffic management operator to know the location where the problem has been detected. Finally, the fourth contribution of this thesis is a **centroid-based algorithm for sensors ranking in a LWSN.** We evaluate the performance of this algorithm considering a realistic channel model (taking into account the link properties at ground level), a uniform deployment, as well as the virtual nodes based-deployment proposed in this thesis.

1.3 The organization of the thesis

This thesis is organized in 8 chapters. After the introduction in this chapter, Chapter 2 reviews the related work and the technical background concerning ITS and WSN. In this chapter, we present vehicle detector technologies and services offered by ITS. We also present existing WSN-based architectures for vehicular traffic monitoring at intersections. In Chapter 3, we characterize link properties with nodes deployed at ground level as function of many parameters. Based on the experimental observations, in Chapter 4, we propose WARIM, a new realistic WSN architecture for vehicular traffic monitoring at intersection. In this architecture, the number of nodes per lane and the distance between two consecutive nodes are defined taking into account not only the application and environment constraints, but also the radio properties. The network formed by nodes deployed on a lane form a LWSN with multi-hop and convergecast communication model. Thus, considering a LWSN, in Chapter 5, we propose a virtual nodes-based deployment strategy, which prolongs the network lifetime compared to a uniform deployment. Chapter 6 proposes an autonomous ranking algorithm for nodes in a LWSN. With this ranking algorithm, nodes autonomously find their ranks without the need for a human intervention. Chapter 7 validates WARIM. Considering a 24 hours dataset of the vehicular traffic of the city of Cologne in Germany, we propose an adaptive message generation strategy which reduces the number of messages generated in the network. Considering a realistic physical layer (built from empirical data), a simple opportunistic routing protocol, the deployment proposed in Chapter 5 and the ranks obtained by the ranking algorithm proposed in Chapter 6 used as geographical coordinates, we use a WSN simulator to evaluate the capability of WARIM to collect vehicular traffic information at an intersection. Finally, Chapter 8 concludes this thesis and presents several perspectives of our work.

Chapter 2

Intelligent Transportation Systems and Wireless Sensor Networks

The number of vehicles in the cities around the world increases in the same rhythm with the urban population. With such a high population density in urban areas, managing the transports or the mobility (of humans or cars) in urban areas becomes a very complex problem.

Using vehicular traffic information (e.g. the number of vehicle passing on a road segment), ITS offer many services like parking place management, access control to restricted area, route guidance, automatic tolls and intelligent traffic lights management. Regardless the considered service, vehicular traffic monitoring is the main component of ITS. This component is realized by vehicular detectors, which can be considered as the lowest level of an ITS architecture. Thus, this chapter reviews existing vehicle detection technologies proposed in the literature. Since our focus is on intersection management, the second goal of this chapter is to present some existing traffic lights management systems. We also investigate existing WSN-based architectures for vehicular traffic monitoring at an intersection.

The remainder of this chapter is organized as follow. In Section 2.1, we define ITS, their main components and the services offered. Concerning these services, Section 2.2 focues on some adaptive traffic lights management systems. Section 2.6 introduces WSN, a low-cost technology for vehicular traffic monitoring. In Section 2.7, we present available WSN-architectures for vehicular traffic monitoring at an intersection, Finally, Section 2.8 concludes this chapter.

2.1 Intelligent Transportation System

The two main concerns with the increasing number of vehicles on the roads are congestion and safety. In the USA, congestion accounts for 160 billion dollars in wasted time and fuel costs [1, 2], with similar figures in other developed countries. We did not find such data for Cameroon or other developing countries. But, given the traffic jams observed in the main cities of these countries, it is clear that this situation has a significant impact on their economy, environment and population. To address this problem, a possible solution is to increase highway and arterial capacity by increasing the number of lanes or by building new ones. But, due to limited budget, it is difficult to implement such a solution, particularly in developing countries. A second solution is to develop

alternatives that increase capacity by improving the efficiency of existing transportation systems. One option for this second solution may be to integrate ICT in the transport management issues. This can be done by investing in ITS while reasonably increasing the highways and arterials capacity, taking into account the available financial budget.

2.1.1 Definition

ITS designate the introduction of ICT in the transport sector [3, 4]. In this context, ITS aim at enhancing transportation efficiency and safety through the use of advanced information processing, communications, control, as well as new electronic technologies. The goals of ITS are to [2]:

- enhance public safety;
- reduce congestion;
- improve access to travel and transit information;
- generate cost saving to motor carriers, transit operators, toll authorities, and government agencies;
- reduce environmental impact.

ITS include automatic vehicles detection, communication, and traffic control technologies. These technologies are assisting states, cities, and towns worldwide to meet the increasing demand of transportation system. Vehicle detection and surveillance technologies are at the heart of ITS and can be considered as being their lowest level. These technologies have been improved to provide enhanced speed monitoring, traffic counting, presence detection, vehicles classification and weight-in-motion data. Some of them are presented in the following section.

2.1.2 Vehicle detection technologies

2.1.2.1 Presentation

Detection technologies can be classified as in-roadway detectors and over-roadway detectors [13].

An in-roadway detector is one that is either embedded in the pavement of the roadway, embedded in the subgrade of the roadway, or taped or otherwise attached to the surface of the roadway [13]. Examples of in-roadway detectors include:

- inductive loop detectors, which require sawcuts in the pavement;
- weigh-in-motion detectors, which are embedded in the pavement;
- magnetometers, which may be embedded or placed underneath a paved roadway or bridge structure;

- tape switches, microloops, pneumatic road tubes, and piezoelectric cables, which are mounted on the roadway surface.

The operation of most of these detectors is well understood as they generally represent applications of mature technologies to traffic surveillance.

An over-roadway sensor is one that is mounted above the surface of the roadway, either above the roadway itself, or alongside the roadway, offset from the nearest traffic lane by some distance [13]. Some over-roadway sensors include:

- video-image processors that utilize cameras mounted on poles adjacent to the roadway, on structures that span the roadway, or on traffic signal mast arms over the roadway;
- microwave radar sensors mounted adjacent to the roadway or over the lanes to be monitored;
- ultrasonic, passive infrared, and laser radar sensors normally mounted over the lanes to be monitored;
- passive acoustic sensors mounted adjacent to the roadway.

Most over-roadway sensors are compact and not roadway invasive, making installation and maintenance relatively easy. However, some installation and maintenance operations may still require the closing of the roadway to normal traffic to ensure the safety of the installer and motorist. All the sensors discussed operate under day and night conditions, although video-image processors may require street lighting at night [13, 14].

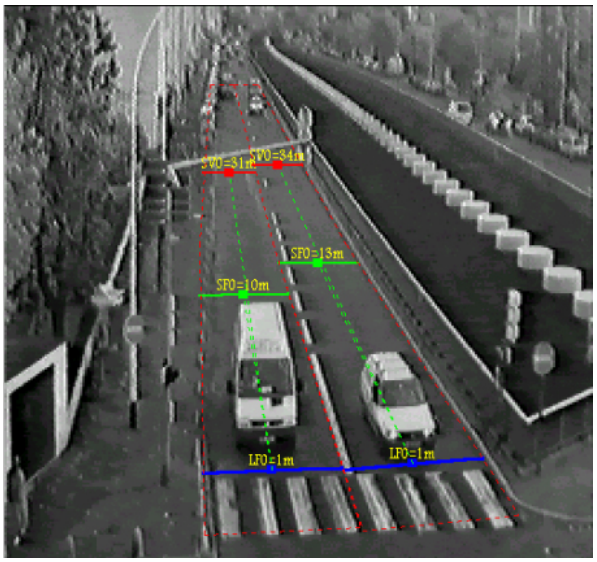
2.1.2.2 Some examples of vehicle detectors

Inductive loop, magnetometer and video-image processors are the most used vehicle detectors (See Figure 2.1). Inductive-loop detectors sense the presence of a conductive metal object (like a vehicle) by inducing electrical currents in the object. The induced current decreases the loop inductance, which is sensed by the inductive-loop electronics unit. The electronic unit interprets the decreased inductance as a vehicle detection. Inductive loops can be used to count the number of vehicles. They can also be used to determine a vehicle speed, length, weight, and the distance between two vehicles. This system can be used for very slow or stationary vehicles (like in a parking), as well as for vehicles traveling at high speed (like on a highway). With video-image processor, video-data recorded by a camera is transmitted to a specialized processor, which analyzes changes in image characteristics as a vehicle passes through motion detection and image processing algorithms. This technology requires an initial configuration of the processor, and among the parameters set, we can mention: the background image, lane separation lines, or the distance between different lanes. The main difficulty with this technology is the size of the collected data (video), that is very large.

Magnetic sensors are passive devices that indicate the presence of a metallic object (like a vehicle) by detecting the perturbation (known as a magnetic anomaly) in the Earth magnetic field created by the object. Two types of magnetic sensors are used for traffic flow parameter

measurement. The first type, two- and three-axis fluxgate magnetometers, detects changes in the vertical and horizontal components of the Earth magnetic field produced by a ferrous metal. The second type of magnetic field sensor is the magnetic detector, more properly referred to as an induction or search coil magnetometer. It detects the vehicle signature by measuring the distortion in the magnetic flux lines induced by the change in the Earth magnetic field produced by a moving ferrous metal. Magnetometers are capable of detecting large objects, such as cars, at tens of meters [15]. The closer the object is, the higher are the distortions. Thus, magnetometers are usually calibrated to detect only vehicles which pass over them. These sensors identify stopped or moving vehicles. As output data, the magnetometer produces the magnetic signature of each detected vehicle. This data is used for vehicle classification and re-identification [16].

(a) Inductive Loop



(b) Camera

(c) Magnetometer, under-ground

(d) Magnetometer, surface-ground

Figure 2.1: Some examples of vehicle detector technologies [13]

Table 2.1 compares the strengths and weaknesses of these three detection technologies, while Table 2.2 presents, for each detector, its output data. More details (technical, cost, etc) concerning the other detectors can be found in [13, 14, 17]. Table 2.2 shows that the video-image processor can be used to monitor many lanes, while the inductive loop and magnetometer are used only for a single lane. The analysis in [17] shows that, considering an intersection with two magnetometers or inductive loops per lane and a video-image processor per approach, the lowest cost is obtained with magnetometer sensors. It is important to note that this cost concerns not only the hardware, but also installation and maintenance.

Technologies	Strengths	Weaknesses
Inductive Loop	• Mature, well understood technology • Insensitive to inclement weather • Provides better accuracy for count data as compared with other commonly used techniques.	• Installation requires pavement cut. • Detection accuracy may decrease when design requires detection of a large variety of vehicle classes. • Installation and maintenance require lane closure.
Magnetometer	• Less susceptible than loops to stresses of traffic. • Insensitive to inclement weather. • Some models transmit data over wireless RF link.	• Installation requires pavement cut. • Models with small detection zones require multiple units for full lane detection.
Video Image Processor	• Monitors multiple lanes and multiple detection zones/lane. • Easy to add and modify detection zones. • Provides wide-area detection when information gathered at one camera location can be linked to another.	• Performance affected by inclement weather and day-tonight transition. • Requires 9- to 15-m camera mounting height for optimum presence detection and speed measurement. • Reliable nighttime signal actuation requires street lighting.

Table 2.1: Strengths and weaknesses of the three most used detector technologies.

	Inductive Loop	Magnetometer	Video Image Processor
Count	✓	✓	✓
Presence	✓	✓	✓
Speed	✓	✓	✓
Occupancy	✓	✓	✓
Classification	✓	✓	✓
Multiple Lanes Detection			✓

Table 2.2: Output data of the three most used detector technologies.

Detectors presented in this section are used to monitor the vehicular traffic. Using the data provided by detectors, ITS offer many services. Some of these services are presented in the next section.

2.1.3 Services offered by ITS

Electronic payment, emergency management, traffic safety, traffic law enforcement and smart parking are some examples of applications [5]. Electronic payment allows users to electronically pay for fuel, tolls or parking spots. Such services may allow to save users time, to secure payments and optimize the usage of the current infrastructure.

ITS are very useful in the case of accidents because, in such a situation, informations are required in real time. Indeed, in case of accidents, the main challenges are the intervention delay, the prevention of accident chains and especially the restoration of the vehicular traffic. A real-time reception of the information concerning the accident makes it possible to better organize and manage accident situations. Traffic safety applications deal with accidents prevention. In order to fulfill this purpose, they make detector devices work pro-actively. They warn drivers about potentially dangerous situations. These situations may be the presence of obstacles, animals, adverse road conditions (ice, water, oil) and vehicles either stopped or driving in the opposite direction (overtaking assistance, wrong-way driving warning). The collaboration among these devices enables to notify drivers of events beyond line-of-sight, thus increasing the available time of response. Some systems for traffic safety are presented in [18, 19].

Traffic law enforcement applications can be considered as a special case of traffic safety applications, since one of the final goals of traffic laws is to increase safety. Currently, traffic law violations are usually detected and put into effect when a police officer or a traffic enforcement vehicle is nearby. ITS though offer permanent monitoring of the locations where they are deployed, enabling to automate the process of reporting infractions. Some laws which can be supervised are speeding, illegal parking, traffic lights violation, unauthorized use of bus lanes or access to restricted areas. Applications such as [20, 21] are able to detect speed limit violations and illegal parking with high precision through the collaboration between adjacent nodes.

The lack of parking spaces in cities is a concern which leads to illegal parking, congestion due to low speed driving and long searching times suffered by drivers. In order to minimize inconvenience to drivers, numerous smart parking systems [22, 23] have been developed, which guide drivers to vacant parking spots and enable smart payment and reservation options. In these systems, a vehicle detector is installed at each parking spot. It thus tells if the parking spot is used (occupied by a vehicle) or free.

Traffic control involves applications directing vehicles within a road network. Their main objective is to smooth traffic on the roads. Detectors deployed along roads are used to measure the traffic flow, obtaining information such as vehicle density, speed, formation of platoons or distribution of vehicles according to different categories. Traffic guidance and intelligent traffic lights management are two examples of applications that may be offered by collected data.

In this thesis, our goal is to propose a simple and low-cost solution to monitor the vehicular traffic at intersections. Traffic information, collected in real time, may be used as input by the traffic lights management system to adapt the light plans to the vehicular traffic state on each road of the intersection. In the next section, we present some traffic lights management systems deployed worldwide.

2.2 Traffic lights management systems

The first ITS to be deployed worldwide (particularly in developed countries) were for traffic management purpose, and particularly for traffic lights management. Three approaches can be used to manage traffic lights in a city [24].

2.2.1 Static management

This is the oldest method in this area. There are no vehicles counting or detection involved. The traffic light plans are fixed. For different periods of the day, light plans are predefined and configured in the light controller. Therefore, it is possible to have a green light for lanes without vehicles, while there are other lanes with high congestion.

2.2.2 Semi-adaptive management

With this approach, vehicles passing on each entry of a given intersection at different times of the day are counted. Vehicles counting can be done manually or using a vehicle detector (See Section 2.1.2). Collected data are: lanes occupancy, vehicles queue length and the average waiting time. These data are then given as input to a software which produces the light plans for all intersections. These light plans are then implemented at each intersection. This can be done manually or automatically using a remote computer. This process can be repeated some years later (2 or 3 years) to take into account changes in road infrastructure and the volume of vehicular traffic, which continuously increases.

2.2.3 Adaptive management

This approach uses real time data provided by detectors to adapt the traffic light plans to the changes in vehicular traffic. Many Adaptive Traffic Light Management Systems (ATLMS) have been proposed and deployed in many cities of developed countries. The goal of any ATLMS is to reduce travel and waiting time, and also to reduce stops at intersections. Split, Cycle and Offset Optimization Technique (SCOOT) [25, 26], Sydney Coordinated Adaptive Traffic System (SCATS) [27, 28], Optimized Policies for Adaptive Control (OPAC) [29, 30], InSync [31], Adaptive Control Software Lite (ACSLite) [32] are some examples of ATLMS deployed worldwide.

Since our goal is to collect vehicular traffic data that can be used for real-time traffic lights management, we are interested by adaptive traffic lights management. In this section, our goal is not to give an exhaustive description of these ATLMS, but to focus on their hardware architecture. These very costly ATLMS are classified into two categories: centralized [25, 27, 28] and decentralized [29, 30, 31, 32] systems.

2.2.4 Centralized traffic lights management

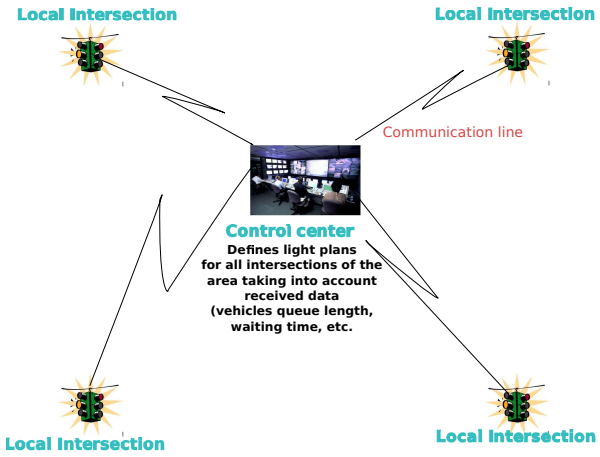


Figure 2.2: Centralized traffic lights management architecture.

2.2.4.1 SCOOT

In SCOOT, each intersection is equipped with a light controller and an inductive loop (one per lane) as vehicle detectors. In this system (See Figure 2.2), all the information collected at the intersection is transmitted through telephone lines to a computer located at the traffic lights management center. This computer processes received data and sends indications concerning the changes to be applied to the light plans to all the traffic light controllers. In SCOOT, the period at which the light plans are updated is calculated taking into account some parameters like the vehicles average waiting time, the vehicles queue length or vehicles stops on lanes.

SCOOT is one of the most used ATLMS in the world. It is currently deployed in 200 cities across 14 countries worldwide. The study presented in [33] shows that its cost per intersection is estimated at 50 000$. This study also shows that SCOOT can reduce the travel time by 8 %,

the average waiting time by 22% and the number of stops at intersections by 17%. But, because of its high cost, its deployment until now is limited only to some cities in developed countries.

2.2.4.2 SCATS

As SCOOT, SCATS is a centralized ATLMS. It divides the monitored area into regions. Each region is controlled by a *Regional Computer* and is composed of a set of intersections. At each intersection, there is a *Light Controller* and an inductive loop installed on each lane as vehicle detector. The light controller processes the data received from the inductive loop and sends to its *Regional Computer* only the information needed (vehicles queue length and waiting time) to calculate light plans. All the information concerning the vehicular traffic in the monitored area are centralized in a *Supervisor Computer*.

With SCATS, a region can control 250 intersections and SCATS can manage 64 regions. The study presented in [27, 33] shows that its cost per intersection is estimated to 60 000$. Compared to static traffic lights management, SCATS could reduce the driving time by more than 15.5%, the accidents by 20%, fuel consumption by 7%, and carbon monoxide (CO) emission by 18%. The study presented in [33] shows that SCATS reduces the travel time by 8%, the delay by 28% and the number of stops at intersections by 42%.

2.2.5 Decentralized traffic lights management

With the two ATLMS presented in the previous section, light plans are calculated on a central computer using data from detectors. The device deployed at the intersection is responsible only for the collection of data (like the vehicles queue length or the average vehicles waiting time) used by the central computer to calculate light plans. The evolution of the technology gives the possibility to manufacture small devices able to execute complex programs. Such devices, deployed at road intersections, have allowed the development of decentralized ATLMS. In these ATLMS, the light plans are calculated at the intersections level by the light controller. Special technologies are then used to establish communications between controllers deployed at neighboring intersections. Communications between intersections allow for example, to create green light tunnels, or to facilitate traffic flow between several intersections. OPAC [29, 30] , InSync [31], and ACSLite [32] are some examples of decentralized ATLMS. In this section, we present InSync, one of the most recent and most deployed decentralized ATLMS. The most important hardware components of InSync are:

Mainly installed in some cities in the USA, InSync is one of the latest traffic lights control systems. The most important hardware components of InSync are:

- **vehicles detector:** a vehicle detector which periodically evaluates, for each lane, the vehicles presence, the queue length and the waiting time. Three options are proposed for vehicles detection: video-image processor proposed as standard detector, an interface to use existing detectors (inductive loop, magnetometer, etc), and a third option which is the fusion of the standard detector (video-image processor) and the existing detectors if any;

- **processor:** the vehicle detectors transmit, through Ethernet, data concerning the queue length on each lane and the vehicle waiting time to the **Processor**;
- **light controller:** the light plans are calculated and transmitted by the **Processor** to the **Light Controller**. To do this, it takes into account the vehicular traffic at the local intersection and the informations received from neighboring intersections.

InSync takes advantage of progress in the field of communications to improve its performances. Indeed, at the intersection, the camera is connected to the processor via an Ethernet connection. And, adjacent intersections can be connected using optical fiber, cables or by using a wireless channel.

The financial cost of InSync is estimated to 30000$ [33], almost two times less than SCATS and SCOOT. Compared to these ATLMS, InSync gives better performance. Indeed, it can reduce the travel time by 25%, the delay by 60% and the number of stops at intersections (due to red lights) between the source and the final destination by 70%.

2.3 Thesis scope

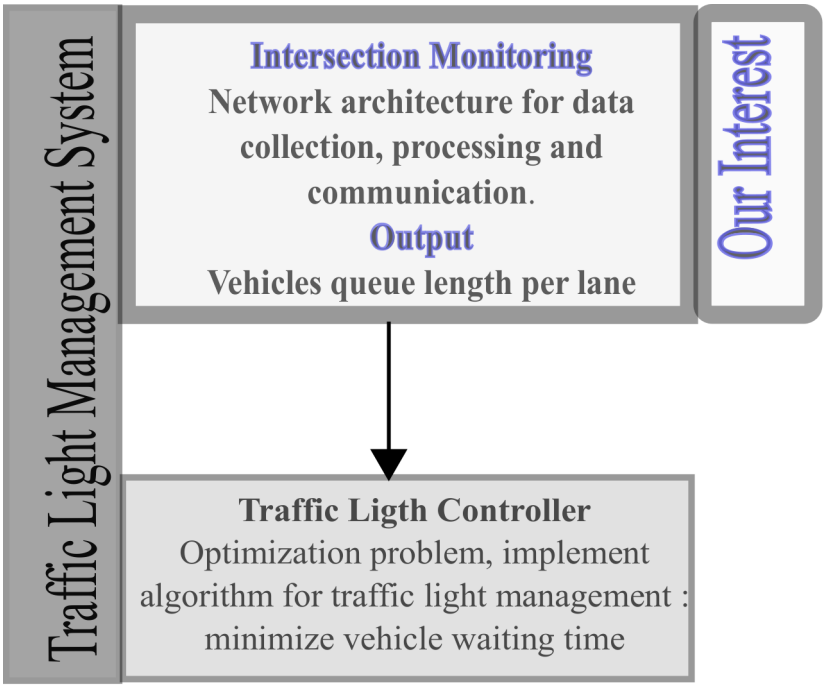


Figure 2.3: Traffic Light Management System Architecture

In the traffic lights management systems presented in the previous section, the controller or the central computer takes as input the number of vehicles on each lane of the intersection and their respective waiting time. It then produces as output the light plans for each direction. The collected data can also be used to offer other services, like vehicular traffic monitoring in the whole city or drivers guidance. Figure 2.3 presents a simplified architecture of the traffic lights

management system. It can be divided into two main components. The first component (***Intersection Monitoring***) is responsible of the monitoring of vehicular traffic at the intersection and provides data (allowing to determine vehicles queue length and waiting time) to the second component (***Traffic Light Controller***), which implements the traffic light management algorithm. Numerous algorithms for traffic light management have been proposed in the literature [8, 9, 10, 34]. In this thesis, we are only interested by the vehicular traffic monitoring at the intersection: i.e., how to measure vehicles traffic at intersection using a lightweight architecture, in order to provide relevant input to the traffic light controller. The goal of the traffic light controller being to reduce driving time and waiting time at each intersection by scheduling the different green lights, the quality of the input (measured as the closeness of the input to the ground truth) impacts the scheduling performance. Communication technologies and protocols play an important role because they define mechanisms that are used to report messages to the light controller. Important research questions are then: how to reliably and accurately monitor road traffic at an intersection? What autonomous and low financial cost solution can be used to reliably measure the vehicles queue length at an intersection? That is the problem addressed in this thesis. The solution must also be simple in terms of the installation duration or complexity.

2.4 Discussions

Even if, compared to SCOOT and SCATS, InSync presents the lowest financial cost, the cost of all the ATLMS presented in this section remains too high. The high cost of the ATLMS is due not only to the hardware purchase, but also to the system deployment and maintenance. Indeed, [27, 33] show that the time required to install one intersection is $24h$ for SCOOT, $13h$ for SCATS, and $18h$ for InSync. Moreover, all these systems require not only many hours of training to understand how the system function, but also a weekly maintenance.

In this thesis, we target vehicular traffic monitoring solutions which can be implemented in any cities worldwide, particularly in developing countries. Such a solution must present the following properties:

1. **low-cost:** the solution must be able to be adopted by most municipalities worldwide, even by those with limited budget;
2. **easy to implement:** the solution proposed must be implemented by an operator without a particular complex training;
3. **autonomous:** the solution must also be able to self-organize and self-configure, in order to reduce the maintenance efforts;
4. **long lifetime:** the new solution will function as long as possible.

Compared to the solutions proposed in this section, we present in the next section a more recent technology. This technology is widely integrated in ITS.

2.5 Vehicular traffic data collection technologies

With the evolution of the technology in various domains, many solutions have been proposed for vehicular traffic monitoring. These solutions include Floating Car Data (FCD) and WSN with magnetometer as vehicle detector.

2.5.1 Floating Car Data

FCD is a more recent technology for vehicular traffic monitoring [35, 36, 37]. The idea behind this technology is to use mobile phones or GPS modules that might be available in a moving car for localization. The principle of FCD is, by locating the vehicles via mobile phones or GPS over the entire road network, to collect real-time traffic data. This technology assumes that some vehicles are equipped with mobile phone or GPS which acts as a sensor for the road network. Data such as car location, speed and direction of travel are sent anonymously to a central processing center. When old mobile phones (without GPS) are used, the location of the vehicle is determined using triangulation or the hand-over data stored by the network operator. Based on these data, traffic congestion can be identified, travel times can be calculated, and traffic reports can be rapidly generated. In contrast to vehicle detector technologies presented in the previous sections, no additional hardware on the road network is necessary. But, this technology cannot be used to measure the vehicular traffic at lane level, since the location precision it offers is very low.

2.5.2 Other technologies

By combining one of the vehicle detector solution presented in Section 2.1.2.2 to a micro-controller and a wireless communication module, it is possible to deploy a solution to collect data on vehicular traffic at a particular road segment. IEEE 802.11 [38], IEEE 802.15.4 [39], LoRa [40] and SIGFOX [41] are some examples of wireless communication technologies operating in the Industry Sciences Medicine (ISM) band that may be used. A permanent energy source might not be available in the deployment environment to power vehicle detection modules. Thus, given the huge energy requirement of IEEE 802.11 [38] module, this technology is not appropriated to the use case considered in this thesis. LoRa [40], SIGFOX [41] and IEEE 802.15.4 [39] are low power technologies. The long range of LoRa [40] and SIGFOX [41] made them privileged communication technologies for a wide application range. But the limited number of messages that can be transmitted with this technology and its high transmission latency are not factors encouraging its deployment, particularly in the scenario considered in this thesis. Finally, IEEE 802.15.4 [39] is nowadays a mature, low power and low cost communication technology. The deployment of a large number of IEEE 802.15.4 [39] communicating modules in an area forms a WSN.

2.6 Wireless Sensor Networks

2.6.1 Definition

Advances in wireless communications and electronics have enabled the development of low-cost, low-power, multi-functional sensor nodes that are small in size and communicate over short distances. These tiny sensor nodes, which consist of sensing, data processing, and communicating components, leverage the idea of WSN [6, 7]. Adding a micro-controller and a wireless transceiver to a magnetic detector (Section 2.1.2.2) allows to build a conventional wireless sensor network. The architecture of a single sensor node is shown in Figure 2.4.

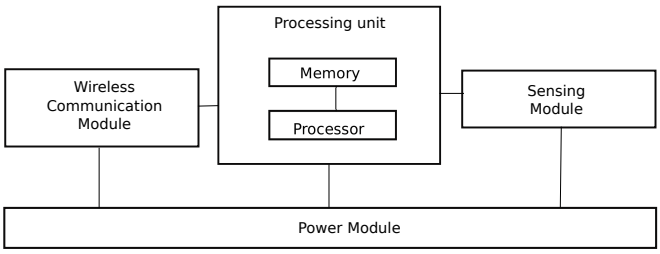


Figure 2.4: Sensor architecture

2.6.2 Applications

Sensor networks may consist of many different types of detectors such as seismic, magnetic (like the one described in Section 2.1.2.2), thermal, visual, infrared, acoustic and radar [6, 7]. These sensors are able to monitor a wide variety of ambient conditions that include the following: temperature, humidity, vehicular movement, lighting condition, pressure, noise levels, the presence or absence of certain kinds of objects or mechanical stress levels on attached objects.

WSN have been adopted in a large number of diverse application domains. It is envisioned that, in the future, everyday objects will be embedded with sensors to make them smart. Smart objects can explore their environment, communicate with other smart objects, and interact with humans. In general, WSN applications can be of two types: monitoring and tracking. The leading application domains of WSNs include military and crime prevention, environment, health, industry and agriculture, and urbanization and infrastructure.

Concerning ITS, WSN can be used as technology to deliver many services [5, 42, 43, 44, 45]. These services include vehicular traffic monitoring, parking spots management, access control to restricted area, automatic tolls management, traffic lights management, and so on. Since, in this thesis, we are interested by the problem of vehicular traffic monitoring at an intersection, Section 2.7 presents existing WSN architectures which address this particular problem.

2.6.3 WSN features

Intended large scale deployments will be made possible by the small price of WSN devices. Wireless sensors are embedded systems with limited resources: a low-power (up to 10 mW), a

low-range (less than 100 m for $IEEE802.15.4$ [**?**] radio widely used in WSN), the low-bandwidth communication ($250kbit/s$ for $IEEE802.15.4$ standard [**?**]), a small memory (few MB), and finally, a small battery (few thousand mAh). Since radio communication consumes most of the power [46], the communication protocols must incorporate energy-efficient communication techniques. Obviously, the characteristic of WSN radio chips are far inferior compared to a concurrent WiFi technology. The original idea to deploy nodes over a large area, combined with a small radio range, lead to multi-hop functioning of WSN.

WSN represent a collection of sensor nodes, with no fixed infrastructure support. Self-organization is thus a real challenge for WSN. It allows to a set of individual sensors to independently create a fully autonomous network, without needing an human intervention.

Similarly to the plug-and-play concept from personal computers, WSN should offer a deploy and-forget experience to the final users. Self-organization over the long operational periods should be taken into account for the link breakage, appearance of new nodes, and dying out of nodes due to the battery exhaustion or malfunctioning. Therefore, WSN should be self-healing.

Low-cost, ease of deployment and nodes self-organization capabilities are some properties that have widely motivated the integration of WSN in the domain of ITS. Indeed, the installation time of a sensor is about 10 minutes, and WSN deployed to monitor a typical intersection can operate for about 10 years [47, 12]. Concerning the vehicular traffic monitoring at an intersection, existing WSN-based architectures are presented in the next section.

2.7 WSN architectures for vehicular traffic monitoring at an intersection

We classify these solutions into two categories: detector-based architectures and link-quality-based architectures. In detector-based architectures, dedicated types of sensors, e.g. magnetic sensors (Section 2.1.2.2), are used to detect vehicles at a given location of the intersection [8, 10, 11]. In the second approach [48, 49], only radio units are deployed, and traffic flows at the intersection are estimated based on the properties of the radio link between these devices.

2.7.1 Detector-based solutions

Most of the proposed solutions for intersection monitoring are based on vehicle detection. Detection technologies can be classified as in-roadway sensors and over-roadway sensors [13]. The main and most visible difference between in-roadway and over-roadway sensors is that the latter can be deployed without disturbing traffic for a long period. However, regarding the sensor financial cost, the most deployed over-roadway sensors, video cameras, are much more expensive compared to magnetic sensors, which are the most popular in-roadway solution [13]. Apart from the cost, the specific placement required alongside the roadway, usually on lampposts, is another strong constraint of over-roadway detectors. Our analysis focuses on solutions using in-roadway detection technologies, and particularly magnetic sensors. Detector-based architectures can be further classified into two categories : infrastructure-free and infrastructure-based.

2.7.1.1 Infrastructure-free architectures

In most of the architectures proposed in the literature [8, 10, 11], magnetic sensors are the main detection technology used for traffic light control applications. With this solution, nodes equipped with 2 or 3-axis magnetometers are deployed on the road to detect the presence or passing of vehicles. Based on the $IEEE802.15.4$ standard, which theoretically provides a communication range in the order of 70-100m, this kind of architecture is adopted by most of the researchers (see Fig. 2.5). In this architecture, two sensors are used per lane : Arrival Detector (AD) sensor, which is used to detect the arrival of a vehicle at the intersection, and Departure Detector (DD) sensor, which is used to detect vehicles exiting the intersection. When AD sensor detects a vehicle, it sends a message to DD sensor. Then DD sensor uses arrival and departure information to compute the number of vehicles per lane. If Nv_i denotes the number of vehicles on the lane at time slot i, A_i and D_i, the number of vehicles detected during time slot i by AD and DD sensors respectively, Nv_i is expressed by Equation2.1.

$$Nv_i = Nv_{i-1} + A_i - D_i \tag{2.1}$$

This architecture has several drawbacks which directly impact the performance of the traffic light controller. The first drawback is the low precision on the number of vehicle reports per lane. Indeed, once a vehicle has been detected by the AD sensor, nothing prevents it to change lanes, especially given the large distance (up to 100 m) between the two sensors. A second drawback with this architecture is represented by the *unauthorized* stops at the intersection. Indeed, if after passing above the AD sensor, a vehicle stops for technical or human reasons, it is not possible with this architecture to automatically and quickly detect or localize this situation. The last but most important drawback with this kind of architecture is to guarantee a reliable wireless communication using the IEEE 802.15.4 standard between two sensors placed over or under the roadway and at a distance of 100 m. We will investigate link properties with sensor nodes deployed at ground level in Chapter 3.

2.7.1.2 Infrastructure-based architectures

To deal with these communication issues, the use of an infrastructure deployed along the road was proposed [50]. In this example, in order to allow line-of-sight conditions favorable to radio communication, at least one wireless node is not located at the road level but higher above the ground to collect data: reliability is improved and a longer radio range is achieved. Therefore, in such solutions, repeaters are installed along the road, at a given height, to relay messages coming from the monitoring sensors. AD and DD sensors, located at ground level, report messages to the repeaters, which relay received packets to the controller. This is currently the approach adopted by most industrial solutions [47, 12, 51]. Even if the radio communication issue is well-addressed by this kind of architecture, all the other problems mentioned for infrastructure-free architectures remain. Note that such an architecture also leads to an increased financial cost, as additional infrastructure has to be purchased, deployed and maintained, which can significantly increase the cost of the system. Moreover, these repeaters are sources of obstruction of the trotory.

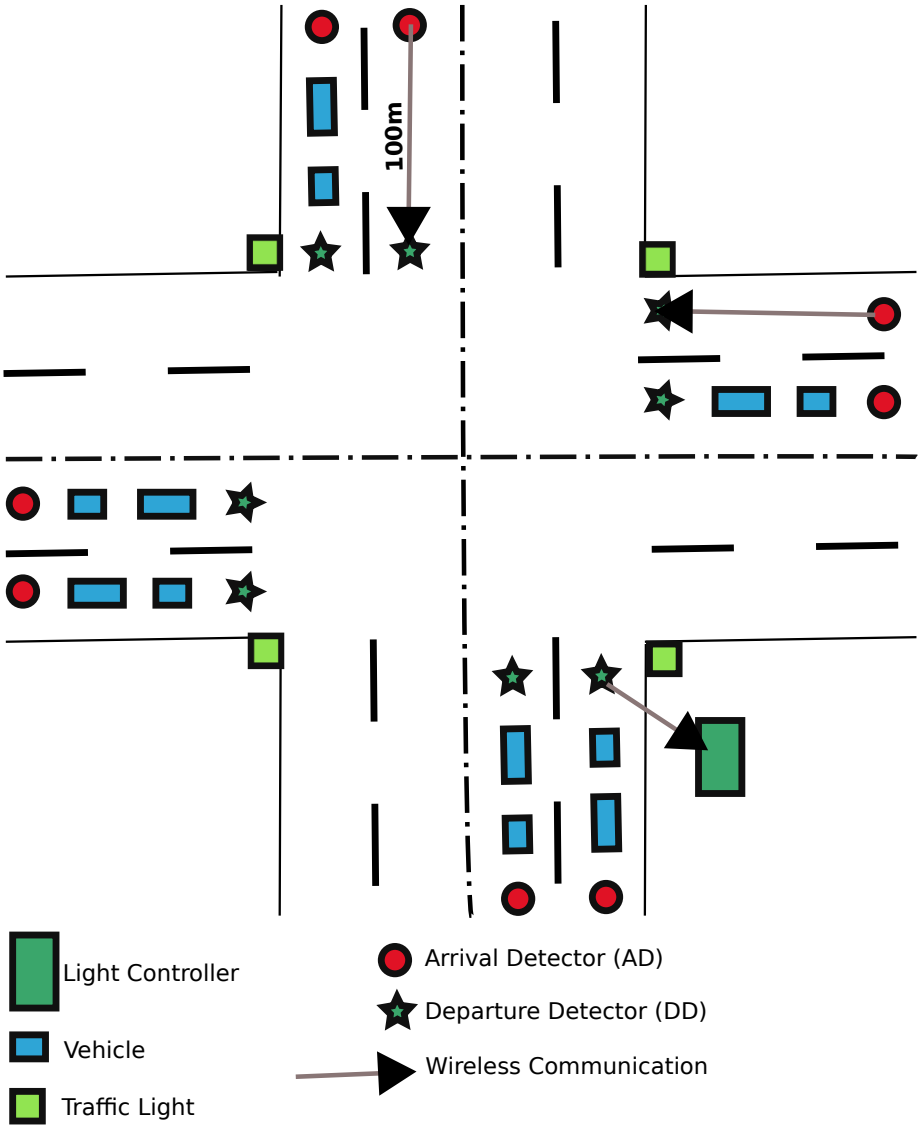


Figure 2.5: Infrastructure-free WSN architecture for traffic light control.

2.7.2 Link-Quality based solution

In the previously discussed approaches, vehicles are detected using either in-roadway or over-roadway sensors, and repeaters along the roadside can be used to improve the reliability of the data collection to the controller. In [48, 49], authors proposed a WSN architecture for traffic estimation based on the monitoring of the quality of radio links, where no dedicated monitoring sensors are used (see Fig. 2.6). The main idea of this solution is that vehicle flow influences the radio link quality between transmitter (TX) and its associated receiver (RX). Both are installed on the two sides of the road at a height of $0.5m$. Received Signal Strength Indicator (RSSI) measured between TX and RX node is used to classify traffic condition as free flow or congested based on a threshold.

Unfortunately, this architecture allows only to give a general view of the traffic conditions on each road of the intersection, without any detail of the number of vehicles per lane. This implies that the system cannot be used in a situation where a high accuracy on the number of vehicles is required or when intersection monitoring is integrated in a global traffic management

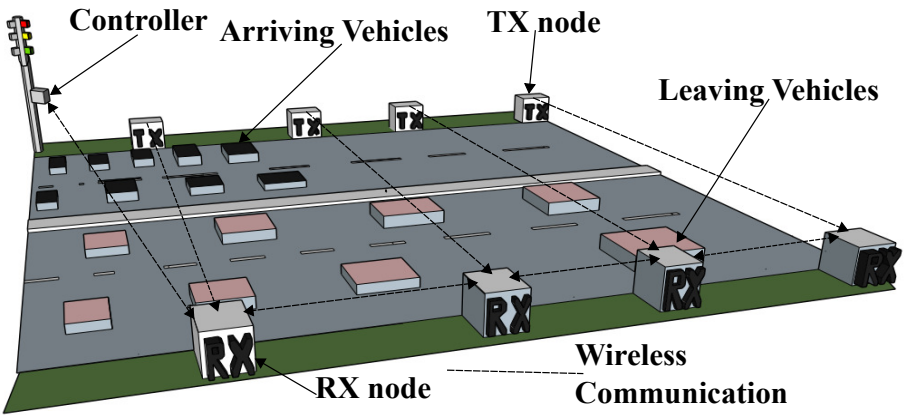


Figure 2.6: Link-Quality based architecture for traffic light control

at a city scale. Moreover, both arriving and leaving vehicles impact the radio link quality; this leads to overestimate the traffic condition because both arriving and leaving vehicles are counted together, while only arriving vehicles should be taken into account in traffic light management. If this solution does not disrupt traffic flow during deployment, the installation of transmitter and receiver nodes increases significantly not only the financial cost, but also the obstruction of the trotory.

2.8 Conclusion

The goal of this thesis is to propose a simple solution for vehicular traffic monitoring at an intersection. The principal target of the collected data is the adaptive traffic lights management system at a local intersection. These data may also be used at a higher level to offer other services.

Thus, we started this chapter by presenting ITS. Firstly, this presentation concerns the vehicles detection technologies and services offered by ITS. Secondly, we presented some traffic light management systems. Indeed, traffic light management systems are the principal target of the vehicular traffic data collected at an intersection. The high cost of these solutions limits their deployment only to some cities in developed countries. Thus, we also presented WSN, a low-cost technology widely used in the literature and by industrials for ITS. We also presented existing WSN-based architectures for vehicular traffic monitoring at the intersections. In these solutions, nodes are deployed on road surface or under the road surface. In this thesis, we are driven by three main requirements:

1. **the realism:** all our contributions must take into account the deployment environment constraints. In other words, our solution must be able to be implemented in the real world;

2. **the cost:** 2.compared to existing adaptive traffic light management systems, the cost of the proposed solution must be low in order to be implemented by many municipalities. This cost include the hardware, the deployment and the maintenance cost;

3. **the easy implementation and maintenance:** the solution must be simple, i.e., easy to implement and maintain. Any operator, without a specific training (contrary to the traffic light management systems presented in this chapter), must be able to implement the proposed solution, and within a reasonable time. The maintenance of the solution must also be simplified. Thus, the network will autonomously organize in order to solve any link failure. The limited energy of nodes must also be efficiently used in order to allow the solution to function as long as possible.

Related work presented in this chapter shows that a sensor can be installed in 10 minutes (i.e., an intersection of 30 sensors can be installed in less than 6 hours, compared to 24 hours installation required for old intersection control systems), and a typical WSN deployed for vehicular traffic monitoring at intersections can operate during 10 years. For these reasons, WSN, with magnetometer as vehicle detector, without additional infrastructures, can satisfy the previous requirements.

Most infrastructure-free WSN-based solutions presented in this chapter are evaluated through simulations, but may give poor performances in a real deployment, due to communication constraints. Our goal in this thesis is to propose, considering the three previous requirements (the realism, the cost, and the easy implementation and maintenance), an infrastructure-free and autonomous WSN-based solution for vehicular traffic monitoring at an intersection.

Concerning the realism requirement, the next chapter presents the first contribution of this thesis: through extensive experiments, we try to understand and to characterize the constraints of wireless links deployed at ground level. The goal being to integrate these constraints when designing a WSN architecture for vehicular traffic monitoring, or when designing communication protocols or mechanisms in a WSN with nodes deployed in such conditions.

Chapter 3

Ground Level Deployment of Wireless Sensor Networks: Experiments, Evaluation and Engineering Insight

In this thesis, our goal is to propose an infrastructure-free, realistic (in terms of network communication), autonomous, low-cost and lightweight WSN architecture for vehicular traffic monitoring at a road intersection. We define an infrastructure-free architecture as an architecture which uses only sensors deployed on the road, i.e. an architecture without pre-existing communication infrastructures deployed along the road. We focus on architectures that can be adopted by many municipalities worldwide, particularly in those with limited financial budget. For an architecture to satisfy all these properties, many challenges must be addressed. Energy-efficient and simple deployment (easy to implement) solutions must be proposed, as well as self-configuration/self-organization and reliable data collection mechanisms. We define self-configuration/self-organization as the capability of a sensor to autonomously discover its neighbors, to configure some parameters like the output power or the communication frequency or to communicate with other sensors in order to find its relative position related to the light controller. In the previous chapter, we reviewed existing WSN architectures for vehicular traffic monitoring at intersections. In the infrastructure-free architectures proposed in the literature, two sensors supporting the IEEE 802.15.4 standard are deployed per lane (under the road surface). These solutions consider, as defined in the IEEE 802.15.4 standard, that the separation distance between the two sensors deployed per lane is $70 - 100\ m$. Our hypothesis at the beginning of this chapter is that it is more difficult to establish reliable communication links between sensors deployed at the road surface, separated with such distances, and using the IEEE 802.15.4 communication standard.

The performance of any WSN architecture, communication protocols, or network mechanisms depends mainly (but not only) on the link quality. Thus, in this chapter, we evaluate and analyze the quality of links with sensors deployed at ground level. Our goal is to highlight the impact of the ground proximity on the links, as well as the impact of network parameters like the communication frequency, the transmission power, the packet size and the deployment

environment. We also want to investigate if an infrastructure-free, lightweight and low-cost WSN architecture for vehicular traffic monitoring at an intersection is achievable.

The remainder of this chapter is organized as follows. In Section 3.1, the problem statement and motivations of this chapter are presented. In Section 3.2, we present relevant literature results concerning the analysis of radio link properties in WSN. Theoretical radio channel models and their application in the case of a ground level deployment are discussed in Section 3.3. In Section 3.4, we present our experiment methodology and setup, and we also describe the platform used. Section 3.5 contains the results of our experimental evaluation and a thorough analysis of the impact of different parameters on the link quality. In Section 3.6, measured data are statistically compared to theoretical ones. In Section 3.7, we summarize our recommendations for WSN deployment and protocols design. Finally, Section 3.8 concludes this chapter.

3.1 Problem statement

The main particularity of WSNs is their limited resources: these networks are constrained in energy, storage, processing and communication capability. Energy is the strongest constraint because sensors are usually equipped with batteries of limited capacity [6, 7]. In addition to the energy constraint, because of the environment constraints (obstacles, variable meteorological conditions, etc), the instability of the radio link is another factor which significantly influences the performance of these networks [72, 73]. Topology construction and maintenance, localization, data aggregation and routing are some key operations in WSN which need to exploit reliable wireless links. Link quality has an important impact on these upper-layer protocols. Then, understanding the link properties in order to integrate them in the design of communication protocols is a real challenge in WSN.

Like the application considered in this thesis, several other applications require sensors to be deployed at the ground level. Bridge and tunnel monitoring [57], or dam monitoring [58] are some examples among others. By using off-the-shelf WSN platforms, our objective is to reduce the financial cost required to deploy a WSN architecture for vehicular traffic monitoring at an intersection. As we will show in the following, a sensor deployment at ground level has a significant negative impact on the properties of the wireless links that form the WSN. Our main goal in this chapter is to assess whether off-the-shelf WSN platforms, not designed specifically for this use case, can be used in a ground level deployment. To reach this goal, we answer many questions: i) what are the key differences between a road surface radio link and a radio link with sensors deployed at a given height? ii) What are the radio link properties between two sensors deployed at the ground level? iii) What is the impact (in terms of packet reception ratio (PRR) or received signal strength indicator (RSSI)) of parameters such as transmission power, packet size or communication channel frequency? iv) What is the impact of the topography of the deployment environment on the link properties? A very important parameter that can have an impact on the link properties is the antenna position or orientation. Unfortunately, we did not studied this aspect in this thesis.

Another important parameter that could impact the radio link properties is the antenna position or orientation.

However, to our knowledge, none of the previous experimental studies considers the case where sensors are deployed at the ground level. Our main motivation in this chapter is to highlight the significant differences that exist between a wireless link at the ground level and one situated at height. To understand the properties of radio links in a WSN with sensors deployed in such conditions, we have conducted an extensive experimental campaign in which we evaluate the impact of different parameters. Considering the PRR and the RSSI as metrics, our observations and recommendations are presented in this chapter.

3.2 Literature review

Numerous studies on the evaluation of radio links in WSN are proposed in the literature. Some authors propose theoretical models [69, 70, 71] of the wireless channel, while others adopt experimental approaches [77].

In [71], the authors propose a coexistence model of two standards (IEEE 802.15.4 and IEEE 802.11b/g), which exposes the interactions between these technologies. They analyze the vulnerability of low power IEEE 802.15.4 technology face to technologies like 802.11b/g, which use a higher transmission power and work in the same frequency band. Their results show that, when IEEE 802.11b/g interference occurs, the throughput of IEEE 802.15.4 decreases drastically. Our experimental results presented later in this chapter also highlight that the link quality in IEEE 802.15.4 standard is influenced by interference from other technologies.

Experimental studies have demonstrated that there is a large transitional region (a geographical area) in the communication area of a wireless node, characterized by a significant level of unreliability and asymmetry. Therefore, in [69, 70], the authors derive the analytical expression of the packet reception rate distribution as a function of distance, as well as the location and extent of the transitional region. Their results show that the environment (path-loss exponent, noise floor) and the hardware (hardware variance, output power) have an impact on the size of the transitional region. In this thesis, we use a single (and homogeneous) hardware platform. Nevertheless, our results show that, from one deployment area to another, the link quality can drastically change, due to the local environment and the sensor location.

Because of the complexity of wireless sensor networks, the designers of communication protocols generally use simplifying assumptions when building their protocols. Thus, in most medium access control (MAC) or routing protocols, a common approach is to assume that the area covered by a signal emitted by a sensor S_i is a circle of radius R, where R is the maximum communication range of sensor S_i. This implies that a message transmitted by S_i is received by all sensors located at a distance less than or equal to R from S_i. This implicitly assumes an isotropic communication area and symmetrical link between a couple of sensors. However, it is established that such assumptions are far from the reality [78]. Our experimental results show that asymmetric links are even more common when sensors are deployed at the ground level.

In a comprehensive experimental study [72], the authors measure the quality of the radio link in two indoor and one outdoor environments. They test the validity of several assumptions usually made in conceptual models used to theoretically evaluate communication protocols. These include the *stability* assumption (i.e. the link quality in terms of packet reception ratio changes slowly compared to data rate), the *channel* assumption (i.e. the link quality is the same on all frequency channels), the *spatial* assumption (i.e. losses on different links are independent) and the *acknowledgment* assumption (i.e. acknowledgment and packet delivery ratios are the same). Their results show that none of these assumptions holds.

Focusing on outdoor deployments [73], the authors evaluate the link quality in three different environments: two tunnels (one with vehicular traffic and another without vehicular traffic) and a vineyard. Their study shows that the link quality in the tunnels is significantly different from other classical WSNs. Indeed, the authors observed that, in tunnels, links are generally stable and long-range compared to the vineyard deployment, the tunnel behaving like a waveguide. While the experiments described in [72, 73] do not consider the scenario of ground level deployment, an important conclusion, that we observe in our experiments as well, is the high dependence of the link quality on the deployment environment.

Moreover, not only the spatial properties of links have been observed, but also seasonal properties [74, 75, 76]. In [75, 76], the authors evaluate the impact of the weather on the link quality in a WSN. Their results show a daily and seasonal fluctuation of link quality. In [74], the authors measured over a period of six months the effects of weather on IEEE 802.15.4 links. By measuring the correlation between the variations in the PRR and the RSSI with four selected meteorological factors (temperature, absolute humidity, precipitation and sunlight), the authors show that the PRR and the RSSI correlate the most with the temperature (better link quality when the temperature is lower). In [75], the authors design a low-cost experimental infrastructure to vary the on-board temperature of sensors and their results also show a correlation between link quality and temperature. In [74], the authors propose a new platform and use it to estimate the parameters of an over-water radio link, through two experiments: one conducted in the morning and another in the afternoon. From the morning to the afternoon experiment, the temperature and the relative humidity have reduced by about 2°C and 25% respectively. Their results shown an increasing of the radio range from 40 m in the morning to 70 m in the afternoon. Results presented in [74, 75, 76] show a poor link quality in presence of high temperature. In our case, since temperature is higher at ground level, this brings a negative impact on link quality compared to links placed higher above ground.

Experimenting at large scale requires a lot of hardware and is a fastidious and time-consuming task. Therefore, several testbed platforms have been deployed all around the world to allow faster experimentations, with various sizes, hardware, topologies, and degrees of flexibility [79, 80, 81, 82]. These testbeds are usually deployed in indoor environments, but allow to conduct complex and easily reproducible experiments. In these testbeds, a user usually selects and configures the set and type of sensors wanted for the experiment. For some testbeds, it might be possible to select sensors in order to guarantee the main radio link property considered in this thesis:

radio link at ground level. However, in our work, we preferred conducting experiments in a real environment, which allowed us to evaluate, for example, the impact of the topography.

To summarize, previous results show that link properties in WSN usually depend on hardware properties, radio communication parameters, meteorological conditions and deployment environment. This should be taken into account when designing communication protocols or mechanisms for WSN. In most experimental setups, the deployment choices usually depend on the desired objectives and the platform used: sensors are usually attached on some support, the sensors may be configured with an internal or external antenna. In our study, we consider an off-the-shelf WSN platform [53], with sensors configured with internal antennas and deployed outdoor on the ground surface. Our goal is to characterize the wireless links in such settings, taking into account radio parameters such as the packet size or the communication channel frequency, as well as environment properties like its topography. Another parameter which may has an impact on the radio link properties, but not investigated in this thesis is the transmission power [83].

3.3 Communication at ground level: Radio channel modeling

As discussed above, a number of applications require sensors to be deployed at ground level. However, the existing WSN platforms were never evaluated in such conditions.

Different metrics can be used to measure the quality of a communication link. Taking a networking point of view, we adopt the PRR as the main metric in our study, as this allows a straightforward derivation of application layer metrics; however we also provide extensive RSSI results in the following, as a complementary metric. As well established metrics, the PRR and RSSI can also be computed theoretically, using standard communication models. Therefore, before detailing our experimental results, we use a similar reasoning to [69, 70] and we derive analytically the PRR and RSSI, to allow a fair comparison between experimental and theoretical results in Section 3.6.

3.3.1 Received power

When an electromagnetic signal propagates, its strength decreases not only because of the distance between the transmitter and the receiver, but also because of other physical phenomena, such as diffraction, reflection or scattering [84]. The signal attenuation, or path-loss, is related to the environment characteristics. Its value depends on the distance between the transmitter and the receiver. It also depends on the position, the shape and the dielectric properties of objects on the signal path. One of the most common radio propagation models, widely used in WSN, is the log-normal path-loss model [70, 84] which is given by Eq. 3.1:

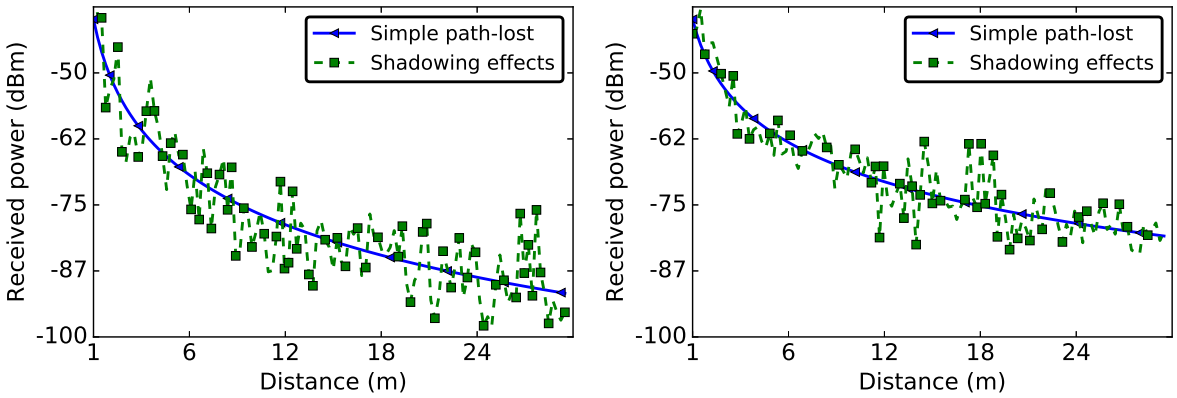
$$PL(d) = PL(d_0) + 10\psi log_{10}\left(\frac{d}{d_0}\right) + X_\phi \tag{3.1}$$

In Eq. 3.1, $PL(d)$ represents the path-lost at a distance d between the transmitter and the receiver. On the right side of Eq. 3.1, the first two terms ($PL(d_0) + 10\alpha log_{10}\left(\frac{d}{d_0}\right)$) represent the

signal path loss produced by the distance d between the transmitter and the receiver (ideal path-loss), while X_ϕ is a Gaussian random variable with mean 0 and variance ϕ (standard deviation due to multipath effect), which represents the path loss due to shadowing effects. The parameters ψ (also known as the path loss exponent) and ϕ depend on the propagation environment and they are generally obtained from field tests in different scenarios. However, from our knowledge, suitable values have never been determined for these parameters in a ground level deployment scenario. Instead, for propagation that approximately follows a free-space or two-ray model, ψ can be set to 2 or 4, respectively [84]. Furthermore, in [84, 85], the authors propose an approach to calculate the value of ψ in a given environment, given measured data. This approach is used to estimate the value of ψ and ϕ in Section 3.6.

The parameter $PL(d_0)$ is the path loss at a reference distance (distance at which the communication area can be assumed to be a perfect disc) d_0; its value can be obtained empirically or analytically using the free-space propagation model at distance d_0. Using an analytical approach,$PL(d_0) = 20log_{10}(\frac{4\pi d_0}{\delta})$, where δ is the wave-length of the transmitted signal. Finally, when a signal is transmitted with a power P_t, the received power$P_r(d)$ at a distance d from the transmitter is given by Eq. 3.2:

$$P_r(d) = P_t - PL(d) \tag{3.2}$$



(a) Ground deployment: $\psi = 3.56$, $\phi = 6.06$ (b) Height deployment: $\psi = 2.75$, $\phi = 3.71$

Figure 3.1: Theoretical received signal strength over a WSN link at ground level (3.1a) and on at a height of 57 cm (3.1b): $P_t = 0\ dBm, d_0 = 1\ m$

Figure 3.1 shows theoretical values of the received power Pr (see Eq. 3.2) as a function of distance, using a transmission power of 0 dBm and a reference distance d_0 of 1 m. $PL(d_0)$ is evaluated using the free space propagation model [84] and is equal to 40.33 dB. Results are shown considering sensors deployed at ground level and then at a given height. The values of ψ and ϕ (3.56 and 6.06 respectively at ground, and 2.75 and 3.71 at height) are computed from our empirical data, as discussed in Section 3.6. To demonstrate the impact of the multi-path fading effect, this figure also depicts the received power in an ideal scenario, with no multi-path fading (i.e. $X_\phi = 0$ in Eq. 3.1). In this simple ideal channel, the path-loss is a logarithmic function of

the distance between the transmitter and the receiver. This theoretical result shows a poor link quality at ground compared to a height deployment.

3.3.2 Packet reception rate

The probability of successfully receiving a packet over a wireless link depends on the packet length and the probability of successfully receiving each bit of the packet. The number of bits transmitted at the physical layer depends on the encoding scheme. The probability of successfully receiving a bit, or its complementary, the bit error rate (BER), depends on the used modulation and on the signal-to-noise ratio (SNR) at the receiver side.

For a received power $P_r(d)$ at distance d, the SNR γ_d is given by Eq. 3.3:

$$\gamma_d = \frac{P_r(d)}{P_n} \tag{3.3}$$

where P_n is the noise floor (the total sum of noise in area and frequency range). In interference-free environments, P_n is given only by the thermal noise and it is constant. Nevertheless, in most environments, P_n varies with time, either because of interference or because of changes in temperature. Similar to [132], we consider this realistic hypothesis (variable P_n) and model P_n as a Gaussian random process of mean 0 and standard deviation X_ϕ.

For a quadrature phase-shift keying (QPSK) modulation, such as the one used by platforms used in our experiments, described in Section 3.4, the BER at distance d, P_e^d is given by Eq. 3.4 [70]:

$$P_e^d = \frac{1}{2} erfc\left(\sqrt{\gamma_d \frac{B_N}{R}}\right) \tag{3.4}$$

where R is the bit data rate, B_N is the bandwidth of the additive white Gaussian noise, and $erfc(x)$ is the complementary error function[1].

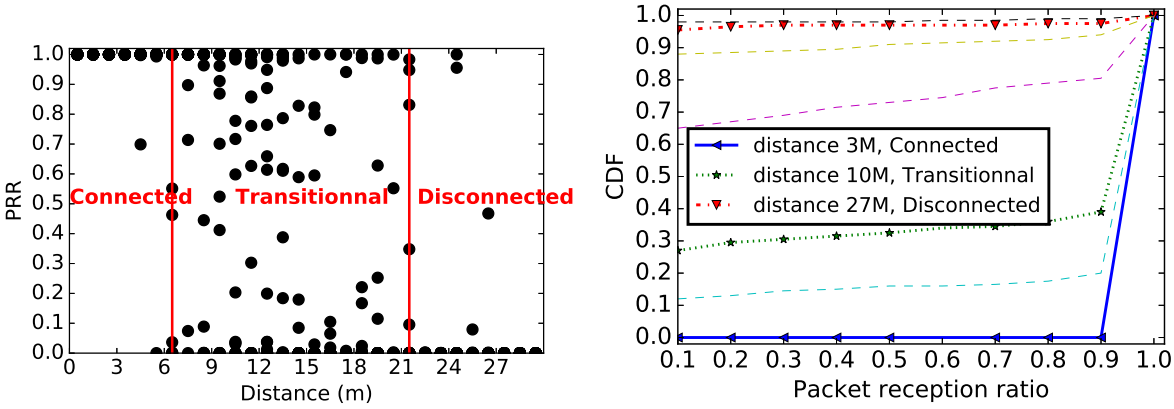
Then, the probability p_d of successfully receiving a packet of size f bits at distance d is given in Eq. 3.5:

$$p_d \simeq (1 - P_e^d)^f \tag{3.5}$$

For several encoding schemes and modulations, the BER and the PRR are given in [70].

Figure 3.2a shows the theoretical PRR values for links of different distances. For the log-normal shadowing model, we used the same parameters values used for Figure 3.1. We consider a QPSK modulation, NRZ encoding and a packet sized of 100 Bytes. For each distance, we generate many PRR values. As in [69, 70], we distinguish three different areas in this figure: *a connected area* in which links are almost perfect ($PRR \approx 1$), *a transitional area* in which the links are unstable and *a disconnected area* in which we have very poor or no links. Figure 3.2b shows the cumulative distribution function (CDF) of link PRR for several distances. It highlights

[1]The complementary error function of x is defined as $ercf(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} \, \mathrm{d}t$



(a) PRR values vs distances

(b) CDF of the PRR

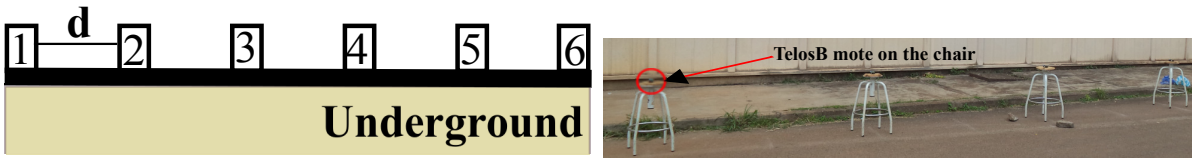
Figure 3.2: Theoretical packet reception ratio over a WSN link: non-coherent $QPSK$ modulation, NRZ encoding, packet size 100 Bytes, $\psi = 3.56$, $\phi = 6.06$, $P_t = 0dBm$

the distances of 3 m, 10 m and 27 m, located in connected, transitional and not connected areas respectively. Results presented in Figure 3.2b show that most of the links in the network are expected to be either of poor quality (PRR less than 10%) or of good quality (PRR superior to 90%), with very few links in the transitional category.

3.4 Experiment overview

In order to understand whether off-the-shelf hardware [53] behaves as predicted by theoretical models, we conduct an extensive evaluation campaign. This section discusses the deployment scenarios, the experimentation material and methodology, as well as the selected evaluation metrics.

3.4.1 Sensors characteristics and deployment



(a) Ground level deployment

(b) Deployment on a support

Figure 3.3: Sensors deployment.

For all the experiments discussed in this chapter, we used the TelosB platform [53]. This platform is equipped with a 8 MHz microprocessor, a 10 KB RAM, a 1 MB external flash memory and a set of sensors (temperature, humidity and light). In all our experiments, the collected data are stored in the external memory.

For communication, TelosB uses the CC2420 radio module. This module is IEEE 802.15.4 [39] compatible and it is widely used for industry and research in the field of low power wireless sensor networks. The CC2420 radio shares with other technologies, e.g. IEEE 802.11 (WiFi),

Bluetooth, the unprotected ISM frequency band (2400-2483.5 MHz). In other words, if a WSN is deployed in an area covered by equipments using these different technologies, the communication of sensors in the network will interfere with these other devices. Following the IEEE 802.15.4 protocol [39], the spectrum is divided into 16 channels of 2 MHz bandwidth and 5 MHz spacing between consecutive channels. The CC2420 module also offers 32 different output power levels, ranging from -25 dBm to 0 dBm. The transmission power and the communication channel can be programmed by changing the value of particular CC2420 registers.

In our experiments, a set of sensors is linearly deployed with a fixed distance d between two consecutive sensors. In each experiment, the same communication protocol, avoiding in-network interferences, is executed by all sensors. To assess the impact of ground level deployment on the quality of wireless links, we also study a scenario where sensors follow a similar linear deployment, but placed on a support, tens of centimeters above ground. We do not control the human movement in the deployment area, nor the external radio interference. Figure 3.3 shows the sensors deployment in the two context: ground level and on a support.

3.4.2 Communication protocol

In all our experiments, each sensor executes the same protocol, which we implemented in Contiki OS 2.7 [55]. For a deployment using N sensors, all sensors are initialized with identifiers ranging from 1 to N. In order to avoid internal interferences, sensors communicate using a round-robin algorithm. When a sensor gets its turn to send messages, it broadcasts M messages at a frequency of one message per 500 ms. This upper layer synchronization allows us to disable the carrier sense mechanism and practically use a transparent MAC protocol.

The transmitted messages contain the identifier of the sender, a sequence number, the transmission power used by the message sender and a dummy application payload. Upon message reception, a sensor stores in its external flash memory all the information found in the message, as well as information on the RSSI and the link quality indicator (LQI) provided by the receiver radio module. This communication protocol is described by Algorithms 3.1, 3.2 and 3.3. Algorithm 3.1 is executed by each sensor at initialization while Algorithms 3.2 and 3.3 are callbacks called for messages transmission and reception respectively.

Algorithm 3.1 Sensor initialization //Called at sensor initialization

Input: N: number of sensors, F: communication frequency, P_t: transmission power, φ: transmission period, i: sensor rank

1: ConfigureCommunicationFrequency(F)
2: ConfigureTransmissionPower(P_t)
3: TxStartDate $\leftarrow$ GetTxStartDate(i, N, φ)
4: SetTimer(TxStartDate, TxCallBack)

Algorithm 3.2 TxCallBack //Called when a sensor want to send messages

Input: $PayLoadSizeList$: the list of different payload sizes, P_t: transmission power

1: $sequenceNumber \leftarrow 0$
2: **for** $plsize$ in $PayLoadSizeList$ **do**
3: SendBroadCastMessage($plsize$, $sequenceNumber$, P_t)
4: SetTimer($Now + \varphi$, TxCallBack)
5: $sequenceNumber \leftarrow sequenceNumber + 1$
6: **end for**

Algorithm 3.3 RxCallBack //Called each time a message is received

Input: $RcvMessage$: the new message received

1: RSSI $\leftarrow$ GetRSSI($RcvMessage$)
2: LQI $\leftarrow$ GetLQI($RcvMessage$)
3: SenderID $\leftarrow$ GetSender($RcvMessage$)
4: sequenceNumber $\leftarrow$ GetSequenceNumber($RcvMessage$)
5: PayLoadSize $\leftarrow$ GetPayLoadSize($RcvMessage$)
6: $P_t \leftarrow$ GetTransmissionPower($RcvMessage$)
7: SaveDataToExternalMemory($RSSI, LQI, SenderID, sequenceNumber, PayLoadSize, P_t$)

3.4.3 Evaluation criteria

As explained, for each received message, the CC2420 module allows logging the RSSI and LQI values. The RSSI is an estimation of the signal power at the receiver. In the TelosB data-sheet [53], the received signal power in dBm, Pr is defined by:

$$P_r = RSSI\_VAL - RSSI\_OFFSET \tag{3.6}$$

where $RSSI\_OFFSET$, empirically found in the data-sheet [53], is equal to -45 dBm.

In addition to the RSSI, the LQI is also a widely used metric to characterize the quality of the radio link in low power wireless sensor networks. However, in this study, we prefer to consider the PRR instead of LQI as a link quality metric, as the PRR is more meaningful at the network layer, and it is calculated over a longer time interval (i.e. several message receptions). Practically, we use the sequence number of the received messages to calculate the PRR value over batches of 20 messages.

Figure 3.4 shows how the link is defined as a function of the measured PRR, following guidelines similar to those in [72]: if no packet is received between two sensors, there is No Link between this pair of sensors. A link on which less than 10% of packets are received is considered as Poor. If 10 90% of packets are received, the link is considered as Intermediate. A link on which more than 90% of packets are received is considered as Good, and when 100% of packets are received we consider the link as Perfect.

Another important property is the link symmetry: between two sensors A and B, we have the oriented link A $\rightarrow$ B that materializes the messages transmitted by A, and B $\rightarrow$ A for messages where B is the sender. Depending on the quality of links A $\rightarrow$ B and B $\rightarrow$ A, we have a symmetrical (the two links exist) or asymmetrical (only one of the two links exists, either A

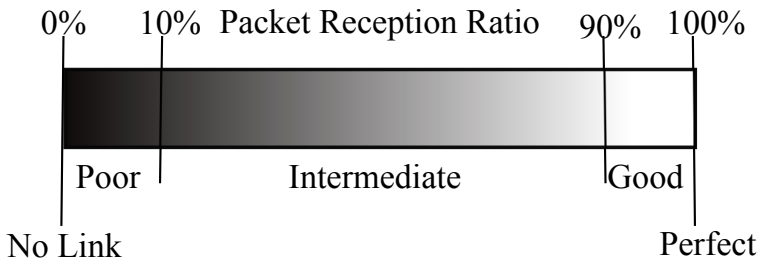


Figure 3.4: Link Definition

$\rightarrow$ B or B $\rightarrow$ A) link between sensors A and B. In order to measure the asymmetry of a link, we use both the PRR and RSSI, by designing two metrics, asm_{AB}^{PRR} and asm_{AB}^{RSSI} , as follows:

$$asm_{AB}^{PRR} = (|PRR_{AB} - PRR_{BA}|) \tag{3.7}$$

where PRR_{AB} is the PRR measured from A to B, and

$$asm_{AB}^{RSSI} = \left(|\overline{RSSI}_{AB} - \overline{RSSI}_{BA}|\right) \tag{3.8}$$

where $\overline{RSSI}_{AB}$ is the average RSSI measured for messages transmitted from A to B for each PRR measured by B. Please note that the latter metric is only meaningful if at least one message is received in both directions. Otherwise, the link is considered unidirectional. Finally, the percentage of unidirectional links, U_l, is also an indicator we use to measure the symmetry of the links in the network.

3.4.4 Experiment methodology

Our methodology, for each experiment, can be summarized by the following steps:

1. a set of N TelosB sensors is used, with all the sensors configured to use one of the 16 communication channels defined by the IEEE 802.15.4 standard;
2. the sensors are linearly deployed, at ground level or on a support at a height of 57 cm above ground, with a fixed distance d between two consecutive sensors;
3. in order to evaluate the impact of the message size, each sensor sends M messages for each application payload from 2 bytes to 100 bytes;
4. all messages are sent in broadcast mode, at a frequency of 2 messages per second;;
5. all sensors used the communication protocol described in Section 3.4.2;
6. when a sensor receives a message, it stores in its external flash memory the message sequence number and the RSSI. The message sequence number is used to calculate the PRR;
7. the experiment ends when every sensor in the deployment has sent all its messages.

In the following sections, we will give, for each experiment we discuss, the number of sensors N, the number of messages M and the distance d between two consecutive sensors.

3.5 Experimental results

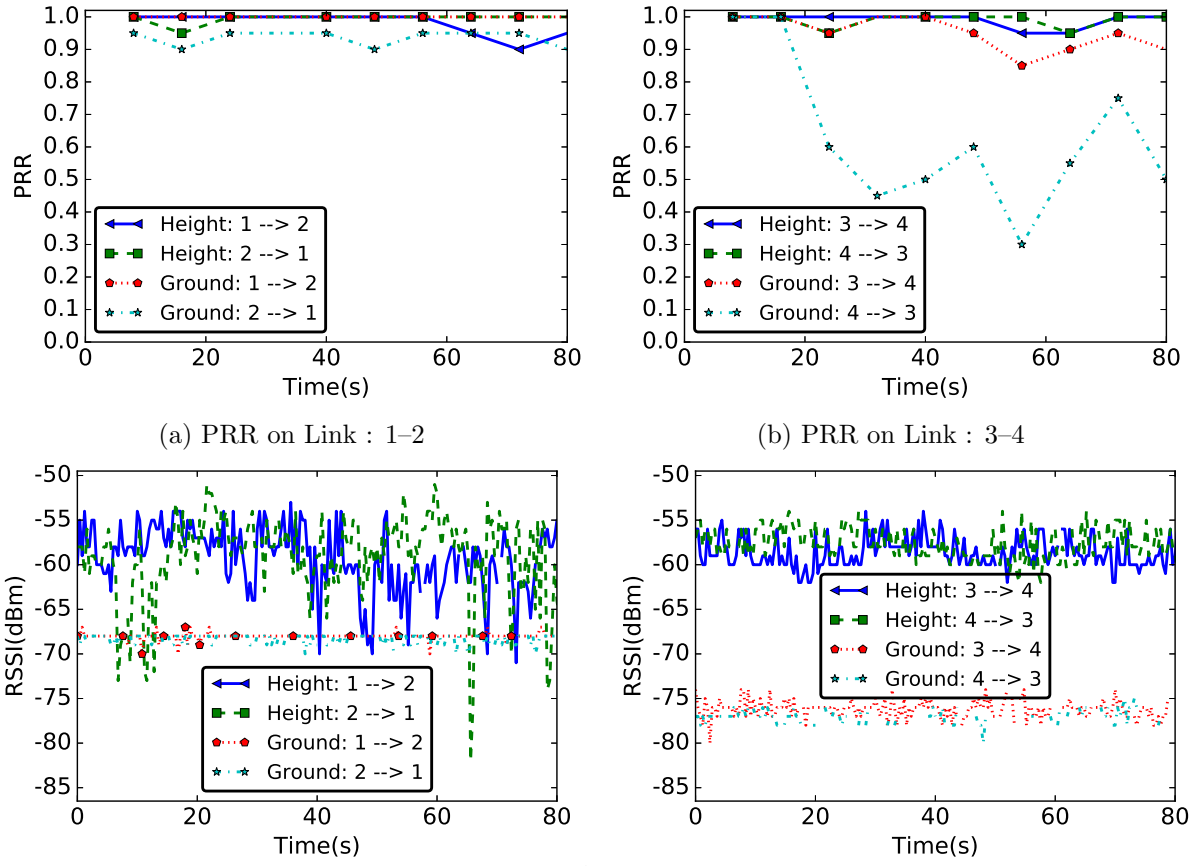
Having discussed the experiment set-up, our goal in this section is to characterize the radio links in a WSN with sensors deployed on the ground surface, taking into account some parameters like the deployment environment, the communication channel frequency and the packet size. We start by comparing the wireless link properties at ground level and in an above-ground deployment, followed by an analysis of the different parameters with an impact on link quality, and we finish by integrating these experimental results with theoretical values obtained using the model described in Section 3.3.

3.5.1 Comparative study of ground-level and above-ground deployment

As noticed in Section 3.2, to our knowledge, none of the experimental results presented in the literature considers the case where sensors are deployed at ground level. For this evaluation, we conducted two sets of experiments. In the first set of experiments, 6 sensors are linearly deployed on the ground with a fixed distance of 3 m between them. In the second set of experiments, sensors are linearly deployed but, in this case, each sensor is deployed on a support of height 57 cm. In both experiments, channel 26 (which does not interfere with communications from other technology like IEEE 802.11) of the IEEE 802.15.4 standard is used by sensors for communication. However, note that Section 3.5.3 evaluates in detail the impact of the communication channel in a WSN with sensors deployed on the ground surface. Sensors are configured to use 0 dBm as transmission power and an application payload of 82 bytes is added as user data for each packet transmitted in the network. In Section 3.5.2, we further evaluate the impact of the packet size on the link quality. Figure 3.5 presents the PRR and RSSI values on forward and backward links of pairs of sensors **(1,2)** and **(3,4)** as a function of time, when sensors are deployed at ground level and at height. From sensor **1** point of view, related to sensor **2**, the forward link is the link $1 \rightarrow 2$ and the backward link is the link $1 \leftarrow 2$. We consider sensors located at different positions in the network, but with the same separation distance of 3 m. This allows to analyze and compare not only the temporal properties of the links in the two deployments, but also some spatial properties.

3.5.1.1 PRR and RSSI values

As mentioned in Section 3.4, we consider, in this work, the RSSI and PRR values to characterize the radio communication links. These metrics can be influenced not only by hardware properties, but also by deployment area characteristics. The values of PRR and RSSI presented in Figure 3.5 show that, in our experiments, the links are most affected when sensors are deployed at ground level. Indeed, the values of measured RSSI or PRR at ground level are always smaller compared to the values measured for a deployment at height. This is particularly true in the case of RSSI, where the difference can reach 10-20 dBm. Since the same hardware is used for both deployments, this significant difference in the value of RSSI and PRR measured in the



(a) PRR on Link : 1–2

(b) PRR on Link : 3–4

(c) RSSI on Link : 1–2

(d) RSSI on Link : 3–4

Figure 3.5: Variation of PRR and RSSI values over time on wireless links **1–2** and **3–4**, with the same distance of 3 m, in two different deployments: one with sensors deployed at ground level (labeled "Ground") and another with sensors deployed on a support of height 57 cm (labeled "Height")

two experiments is mainly due to the environment properties. These properties are time and linkposition dependent. In the following, we compare, for the two deployments, the temporal and spatial properties.

An important observation can be made in Figure 3.5: the two metrics, RSSI and PRR, have different behaviors. While the difference in terms of RSSI between ground level and above-ground scenarios is always significant, this is not the case for PRR, meaning that a similar PRR can be obtained at highly different RSSI levels. Moreover, the opposite is true as well and, at similar RSSI levels, the PRR can be quite different, as shown for the two directions of link **3–4**. The results presented in this section show that the PRR is the best metric to use at network layer to select the link which guarantees the relevant metric transmission reliability.

3.5.1.2 Temporal properties

As observed, the RSSI or the PRR values are better at height compared to a ground level deployment. Figure 3.5 also shows that the temporal properties of links are different in the two deployments, with an opposite behavior when considering PRR and RSSI values. At ground

level, while we have a stable radio link, but poor quality, in terms of RSSI values, we observe a relatively high fluctuation of the measured PRR values. A different behavior can be noticed for the deployment at height, where we have stable PRR values and a high fluctuation of RSSI values.

We also note that the relatively high RSSI values obtained at height result in better PRR results than the stable RSSI values measured at ground level. As hinted by Figure 3.5b and further discussed in Section 3.5.1.4, the ground level deployment also results in a significant number of unidirectional links. Indeed, despite very similar RSSI values, link **3–4** shows highly different PRR values on the two directions, sensor **4** receiving many more messages than sensor **3**.

An explanation for the low RSSI values in the ground level deployment might come from the increased number of signal reflections, hence result in multi path [84]. As both the transmitter and the receiver are close to the ground, this phenomenon is to be expected. On the other hand, in the case of a deployment at height, the propagation environment is more open space (compared to a deployment at ground level), but also more sensitive to different sources of reflexions (e.g. a person moving around the deployment area). We guess that the fluctuations of the RSSI values in the case of a deployment at height is due to this property of the environment.

3.5.1.3 Spatial properties

In practice, the reliability of many hardware-based localization algorithms [91, 92] depends on the correlation of the RSSI with the distance. However, this correlation is questioned by our experimental results, as well as by other previous field tests [78]. For example, links **(1-2)** and **(3-4)** presented in Figure 3.5 cover the same distance (3 m), but they are located at different positions. However, as Figure 3.5 highlights, the properties of these two links are different, especially in the case of the ground level deployment. Indeed, the shape of the road around each sensor may be different from one sensor to another, which leads to different link properties at different points in the network. We argue that such spatial considerations, as well as temporal variations, must be taken into account by WSN designers.

3.5.1.4 Link asymmetry

Link asymmetry is another property that must be taken into account when designing WSN protocols. Looking at the RSSI or PRR measured on forward and backward links of pairs of sensors **(1,2)** and **(3,4)**, we notice a significant difference: for a ground surface deployment, Figure 3.5b shows a much higher PRR on the forward link $3 \leftarrow 4$ than on the backward link $4 \leftarrow 3$.

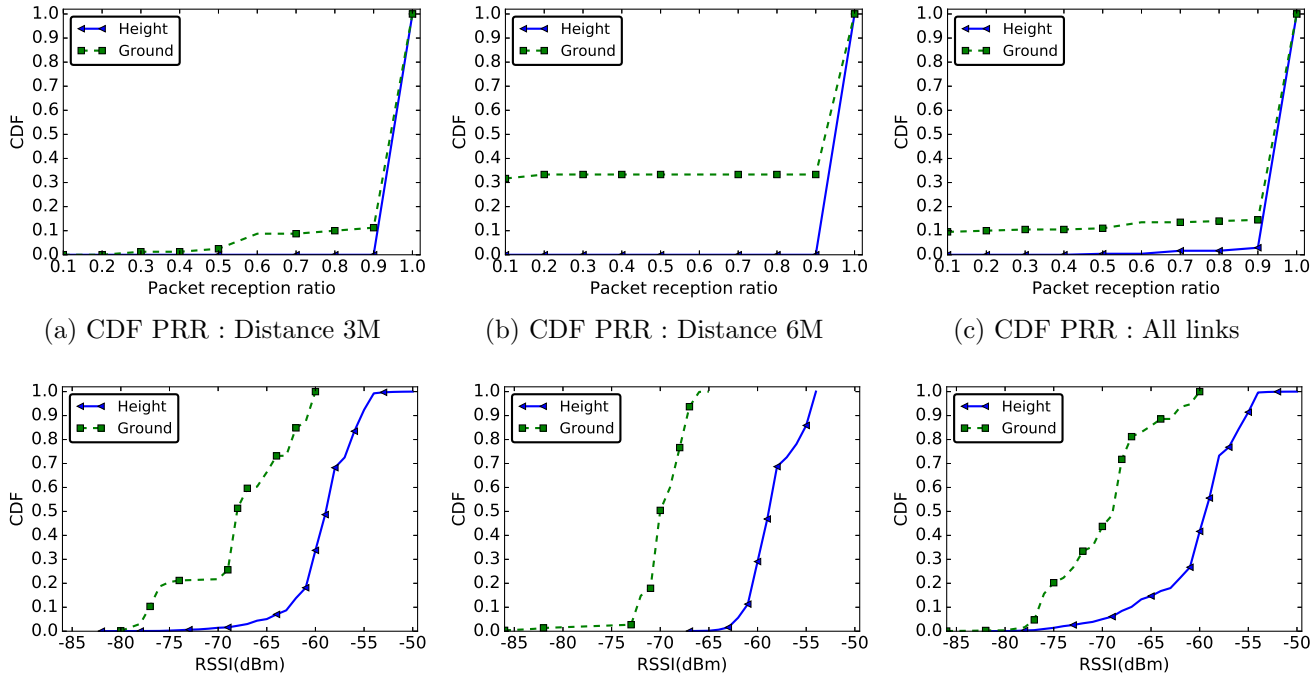
Beyond this illustrative example, Table 3.1 presents statistics concerning the difference of PRR and RSSI values measured on forward and backward links on all sensors, as well as the number of unidirectional links. These results show that the asymmetry problem is general in WSN, and not restricted to ground level deployments. Indeed, Table 3.1 shows that, while the average difference of measured PRR on forward and backward links is higher at ground level, the

maximum difference of the measured PRR and the maximum or the average difference of measured RSSI on forward and backward links are even higher for a deployment at height. However, we also note that, while there are no unidirectional links when the sensors are deployed at height, we have **8%** of unidirectional links in the case of ground surface deployment. This means that, for a deployment at height, even if there might temporarily exist a significant difference in values measured between two sensors, the forward and backward links are most of the time available in the two communication directions.

	Ground	**Height**
$AVG_{\lvert PRR_{AB}-PRR_{BA}\rvert}$	**0.0505**	0.04375
$MAX_{\lvert PRR_{AB}-PRR_{BA}\rvert}$	0.55	**0.6**
$AVG_{\lvert \overline{RSSI}_{AB}-\overline{RSSI}_{BA}\rvert}$	1.67	**2.67**
$MAX_{\lvert \overline{RSSI}_{AB}-\overline{RSSI}_{BA}\rvert}$	4.7	**14.001**
$\%Unidirectional$	**8**	0

Table 3.1: Link asymmetry for deployments at height and at ground level.

3.5.1.5 PRR and RSSI distribution



(a) CDF PRR : Distance 3M (b) CDF PRR : Distance 6M (c) CDF PRR : All links

(d) CDF RSSI : Distance 3M (e) CDF RSSI : Distance 6M (f) CDF RSSI : All links

Figure 3.6: CDF of PRR and RSSI for sensors for distance 3 m, 6 m and for all links in the network in two different deployments: sensors deployed at ground level and another with sensors deployed on a support of height 57 cm.

Figure 3.6 shows the CDF of the PRR (Figure 3.6a-3.6c) and RSSI (Figure 3.6d-3.6f) values measured for the two deployments. These results are shown for sensors from 3 m and 6 m apart, and for all existing links between sensors in the network. As in the case of Figure 3.5, these

results also confirm the degraded quality of a link in a ground surface deployment compared to a height deployment.

For example, in Figure 3.6f, nearly 80% of the messages are received with a RSSI less than -65 dBm at ground level, while in a deployment at height, more than 80% of the messages are received with RSSI higher than this value. The differences are even more visible when one considers just links of distance 6 m, where the ranges of the measured RSSI values in the two deployments are actually disjoint.

Figure 3.7 gives the distribution of the PRR of radio links in the network according to their quality (see Figure 3.4) and deployment, taking into account all the network links. Regardless the deployment conditions, these results show that most of the radio links in the network are good with very few intermediate links (see Figure 3.4). These results are consistent with the theoretical ones presented in [70] and in Section 3.3, and with other experimental results [72]. We note, however, that almost 10% links at ground level are poor, while all links in the case of a deployment at height are either intermediate or good (with almost 95% of good links).

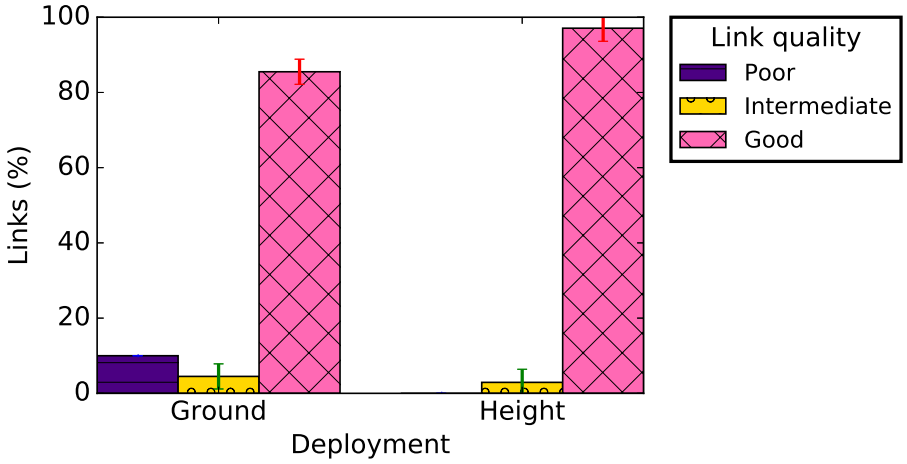


Figure 3.7: Link quality classification for a deployment at ground level and at height.

3.5.1.6 Discussion

Considering the temporal and spatial properties of radio links, as well as the link asymmetry, our results conclude to a deteriorated radio link quality when sensors are deployed at ground level. The relatively poor quality of wireless links in such deployment is due to the proximity of the ground, which creates multiple signal reflections and can even act as an obstacle to radio wave propagation. Since reflected and original signals are summed at the receiver, this can produce distortion in the resulting signal and then negatively impact link quality [84].

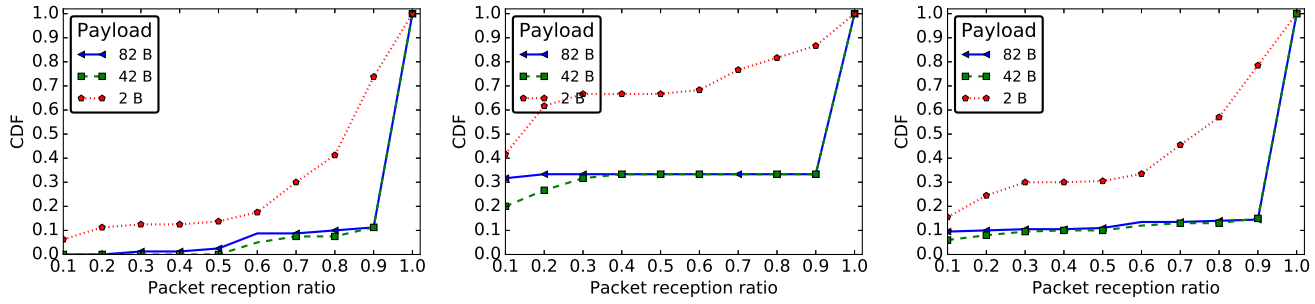
The temperature at ground level can also be an explanation for this result. Indeed, several studies [74, 75, 76] show that the higher is the temperature, the lower the quality of the link, because of the thermal noise. In our case, with sensors deployed on a road, often in full sunlight, the temperature at ground level is significantly higher than the temperature for the deployment at height.

It is well known [72, 73, 74, 75, 76, 78] that the radio link in low-power WSN is very unstable because of many factors like the radio properties or the propagation environment. The results presented in this section show that, at the ground level, the radio links quality (measured by the PRR and RSSI values) are more severely affected by the environment. This must be taken into account, not only during the deployment of such networks, but also for designing communication protocols.

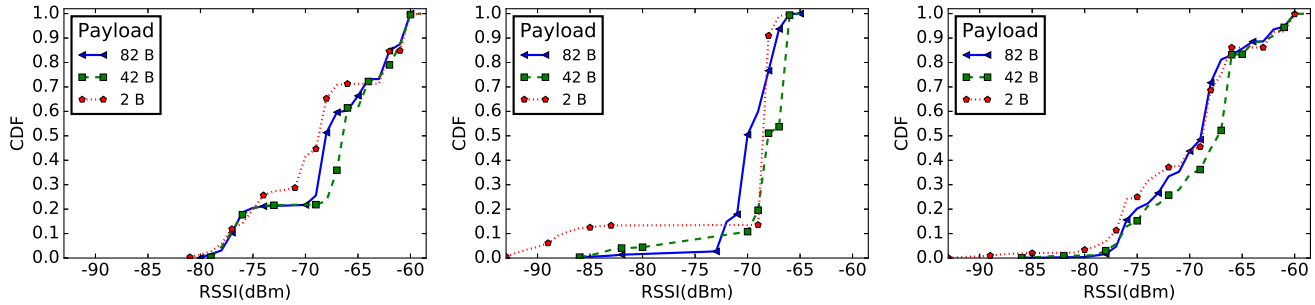
On the good side, our results demonstrate that communication between off-the-shelf sensor is possible at ground level. Thus, by using relaying capability of sensors, the door is opened to the numerous applicative use-cases discussed in Section 3.1. However, results presented in this section only consider one transmission power (0 dBm), one application payload size (82 bytes), one communication channel (number 26), and one deployment area (a relatively flat area). All these parameters may have an impact on the link quality. Therefore, in the following, except the transmission power which is kept to 0 dBm, we analyze the impact of these parameters on the link properties.

3.5.2 The impact of packet size

The maximum size of an IEEE 802.15.4 frame is 127 bytes. In this section, many experiments are conducted with different packet sizes. Our goal is to evaluate the impact of this parameter on the radio link when sensors are deployed at ground level. We consider five sensors linearly placed on the ground with a distance of 3 m between two consecutive sensors. The sensors are configured to use channel 26 for communications and 0 dBm as transmission power. We vary the application payload from 2 bytes to 102 bytes. Figure 3.8 shows the PRR and RSSI CDF for radio links between 2 sensors separated by a distance of length 3 m (Figure 3.8a and 3.8d), 6 m (Figure 3.8b and 3.8e), and for all links in the network (Figure 3.8c and 3.8f). For presentation purpose, only three application payloads are shown in this figure. The results do not show a significant difference concerning the RSSI distribution, regardless the application payload. Nevertheless, Figures 3.8a - 3.8c present a clear difference between the PRR distribution for a very small application payload (2 bytes; i.e. only the sequence number as payload in the packet) and larger ones (42 and 82 bytes). Figure 3.9 shows that, with an application payload of 2 bytes, there are more intermediate radio links in the network, while for other message sizes more good radio links are present in the network. Similar remarks can be made concerning the radio link asymmetry. Indeed, data presented in Table 3.2 shows that, when the application payload is 2 bytes, the average difference of the measured RSSI or PRR on forward and backward radio links is the highest. The percentage of unidirectional wireless links is also larger for such a small message size. We could not explain the poor quality of links when the packet payload is only 2 bytes, which might be even due to some operating system bug. However, we can confirm that this phenomenon manifested in all the different settings we tested, and it disappears for applicative payloads larger than 10 bytes.



(a) CDF PRR : Distance 3M (b) CDF PRR : Distance 6M (c) CDF PRR : All links



(d) CDF RSSI : Distance 3M (e) CDF RSSI : Distance 6M (f) CDF RSSI : All links

Figure 3.8: CDF of PRR and RSSI for distances 3 m, 6 m and for all links in the network for different packet sizes.

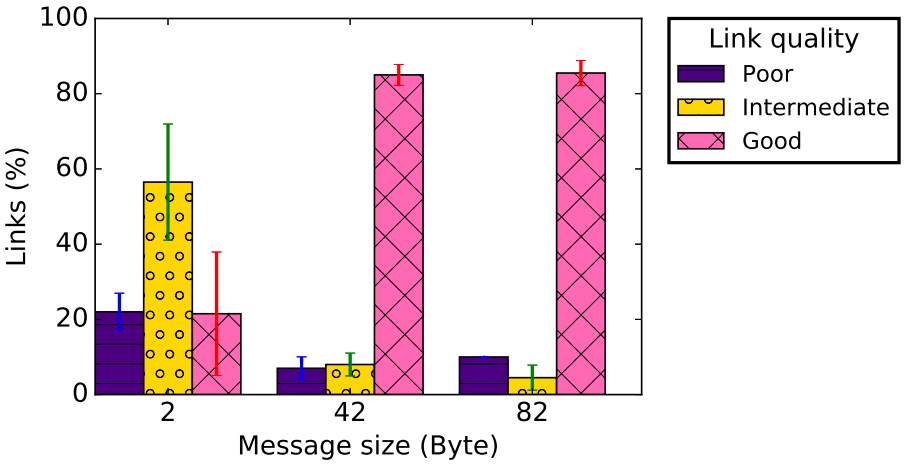


Figure 3.9: Link quality classification for different packet sizes.

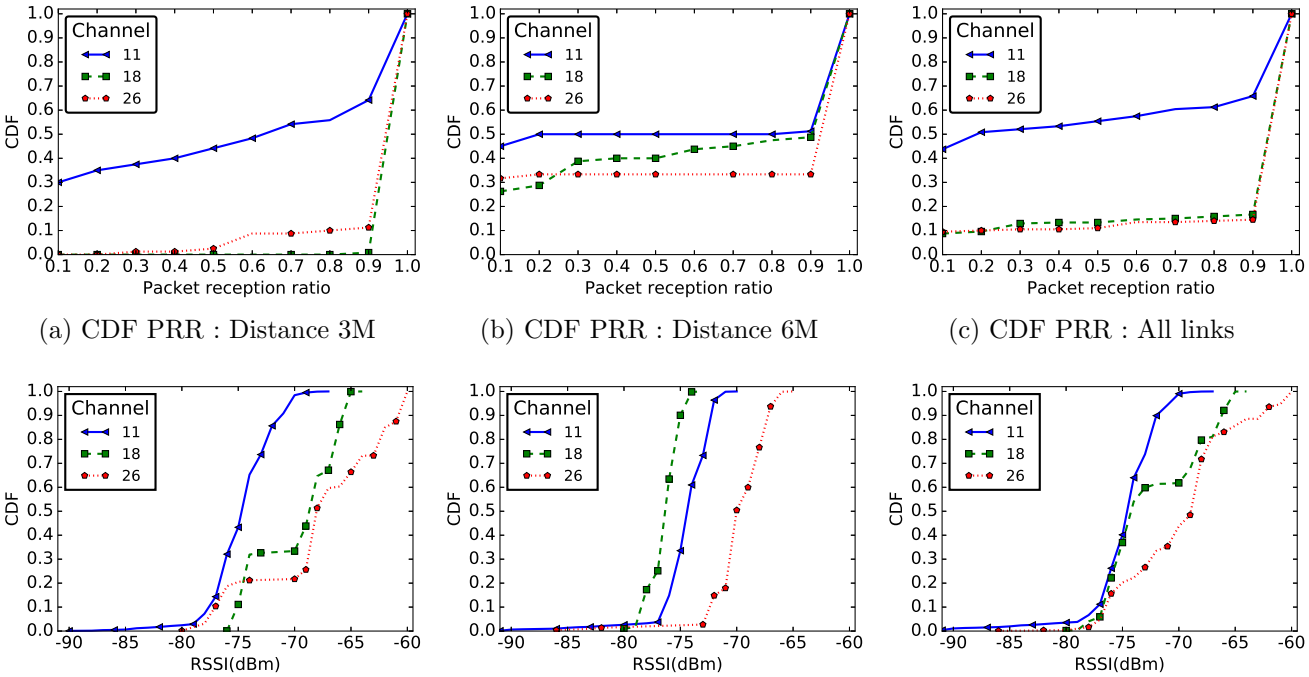
	2 bytes	**42 bytes**	**82 bytes**
$AVG_{\lvert PRR_{AB}-PRR_{BA}\rvert}$	**0.21**	0.05	0.05
$MAX_{\lvert PRR_{AB}-PRR_{BA}\rvert}$	**0.9**	0.50	0.55
$AVG_{\lvert \overline{RSSI}_{AB}-\overline{RSSI}_{BA}\rvert}$	**1.94**	1.68	1.67
$MAX_{\lvert \overline{RSSI}_{AB}-\overline{RSSI}_{BA}\rvert}$	4.47	**5**	4.70
$\%Unidirectional$	**18**	6	8

Table 3.2: Link asymmetry for different packet sizes.

3.5.3 The impact of communication channel

Because the ISM band used in IEEE 802.15.4 is shared by other technologies like IEEE 802.11 and Bluetooth [86, 89], interference from other networks influence the link properties in a WSN.

In order to test the impact of the communication frequency on the link quality, we conducted a set of experiments using 5 of the 16 channels (numbered from 11 to 26) defined by the IEEE 802.15.4 standard. For these experiments, 4 sensors were linearly deployed on the ground surface with a distance of 3 m between two consecutive sensors. Sensors are configured to use a transmission power of 0 dBm. For each channel, experiments were conducted for various application payloads, but results presented in this section consider an application payload of 82 bytes (since, in the previous section, we obtained better link quality with this application payload).



(a) CDF PRR : Distance 3M (b) CDF PRR : Distance 6M (c) CDF PRR : All links

(d) CDF RSSI : Distance 3M (e) CDF RSSI : Distance 6M (f) CDF RSSI : All links

Figure 3.10: CDF of PRR and RSSI for distances 3 m, 6 m and for all links in the network, for different communication channels.

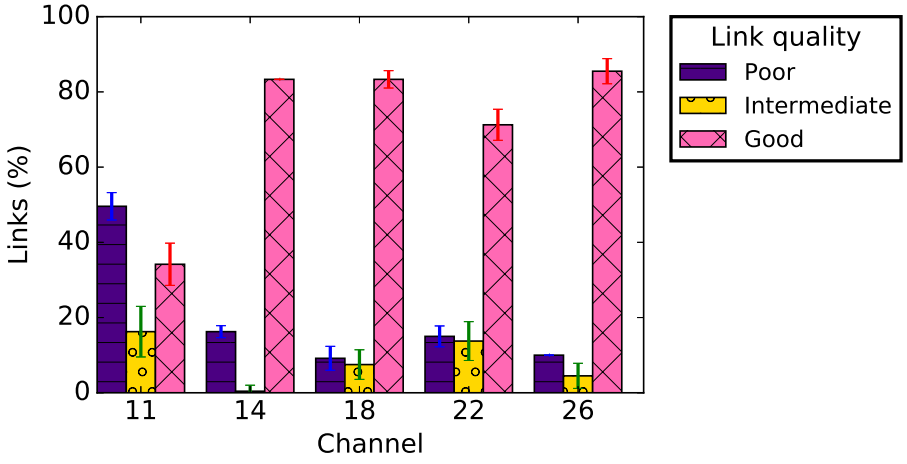


Figure 3.11: Link quality classification for different communication channels.

Figure 3.10 presents the distribution of the PRR and RSSI on communication channels 11,

18 and 26. In this figure, we consider the cases of links of distance 3 m only, 6 m only, as well as aggregated results for all the links in the network.

These results show that the quality of links is better on channels 18 and 26 compared to channel 11. Figure 3.11 also shows that, on channel 11, we observe more poor quality links, while more good quality links are observed on other channels. The poor performance of links on channel 11 can also be observed in Table 3.3, which presents the impact of communication channel on links asymmetry. Even if the PRR measured on forward and backward links are almost equal on channels 11 and 18, the values of measured RSSI are very different in both links. This table shows that almost half of the links in the network are unidirectional when communication channel 11 is used. This lower quality of the links on channel 11 can be related to interference caused by IEEE 802.11 [86, 130, 131, 132] networks deployed on the university campus.

We also observe that the links are better on channel 26 compared to other channels. This is because channel 26 is the only one that does not overlap with the spectrum used by IEEE 802.11 networks [86]. In this section, it is important to note that, for each experiment, all sensors in the network communicate on the same channel. Better performance can be obtained if more advanced MAC layer techniques, such as Time-Slotted Channel Hopping, are used [87, 88].

#Channel	**11**	**14**	**18**	**22**	**26**
$AVG_{\lvert PRR_{AB}-PRR_{BA}\rvert}$	0.0900	0.01625	**0.0925**	0.0780	0.0505
$MAX_{\lvert PRR_{AB}-PRR_{BA}\rvert}$	0.90	0.10	**1**	0.75	0.55
$AVG_{\lvert\overline{RSSI}_{AB}-\overline{RSSI}_{BA}\rvert}$	**2.35**	1.65	1.3	1.81	1.67
$MAX_{\lvert\overline{RSSI}_{AB}-\overline{RSSI}_{BA}\rvert}$	**11.67**	5	3.90	10.43	4.70
$\%Unidirectional$	**40.83**	13.33	9.11	9	8

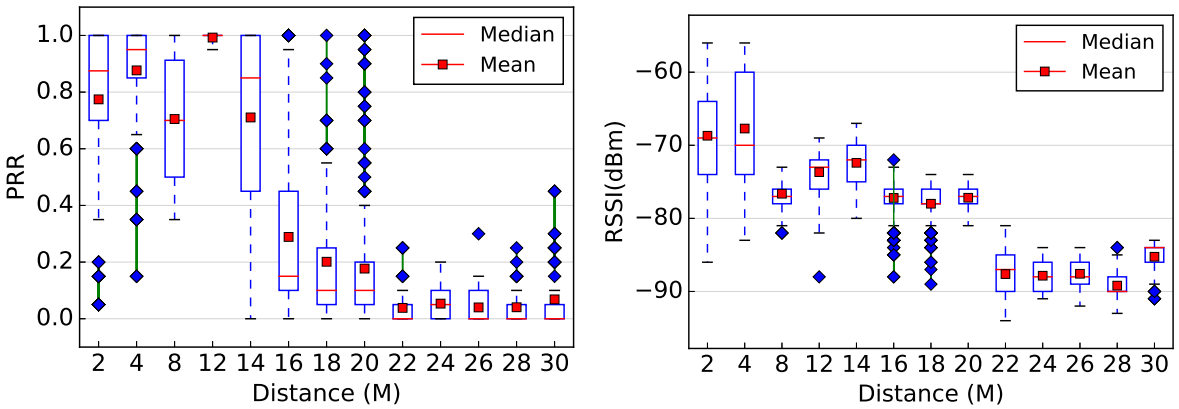
Table 3.3: Link asymmetry for different communication channels.

3.5.4 The correlation between link quality and distance

To show the correlation between the link quality and the distance separating communicating sensors, we linearly deployed a set of 5 sensors at ground level. By varying the distance between two consecutive sensors, we performed several experiments. In all experiments, sensors are configured to use channel number 26 and 0 dBm as transmission power. Figure 3.12 shows the PRR and RSSI as a function of the distance. The red points show the mean PRR (Figure 3.12a) or RSSI (Figure 3.12b). The red line in the box represents the median, the bottom edge of the box represents the 25th percentile, and the top edge represents the 75th percentile.

Intuitively, one might think that, the lower the distance between two sensors, the better the radio link quality between those sensors. The theoretical model described in Section 3.3, as well as many others available in the literature [84], implicitly assume this theoretical result, particularly for the received power level. Also, many hardware-free distance estimation solutions proposed in the literature consider this theoretical result [124, 133, 134]. Simulation tools for WSN also widely use similar models to measure the received power between two sensors.

However, results presented in Figure 3.12a for PRR, and particularly those presented in Figure 3.12b for RSSI, show that such assumptions do not hold in practice. Indeed, Figure 3.12a shows that the measured PRR is better at 12 m than for shorter distances. This may be due to the road shape or other environment factor. Also, Figure 3.12b shows that the RSSI is not a strictly decreasing function of the distance. Therefore, such considerations must be taken into account by protocols designers during theoretical or experimental performance evaluation.



(a) PRR values distribution

(b) RSSI values distribution

Figure 3.12: Correlation between the PRR/RSSI and the distance (please not that the x-axis is not linear).

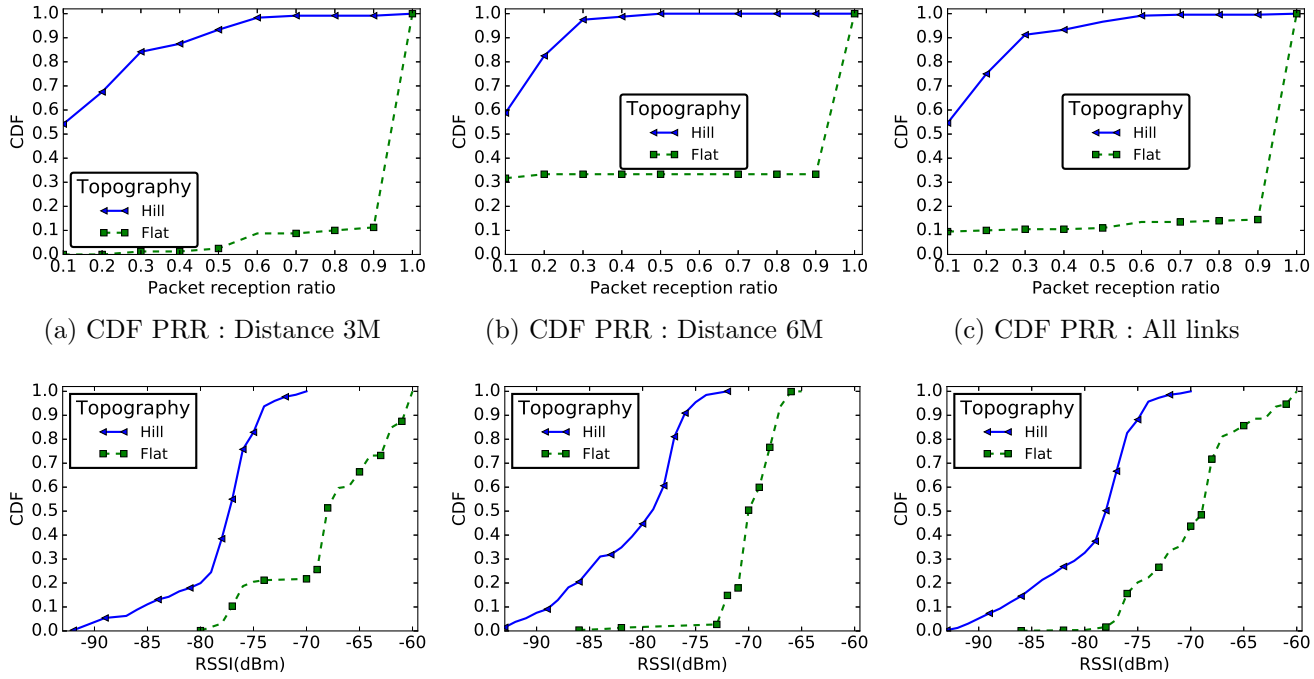
3.5.5 The impact of topography

The deployment environment can have an important impact on a radio link. In our case, the shape of the road, networks nearby and the weather are some environment factors that may influence the wireless link properties. We therefore performed experiments in another environment in addition to the relatively flat area used so far. In this second environment, the sensors are deployed on a hill with a steep slope. In this deployment, the sensors are configured to use channel number 26 and 0 dBm as output power for transmissions. The results presented in this section are for an application payload of 82 bytes.

	Flat	**Hill**
$AVG_{\lvert PRR_{AB}-PRR_{BA}\rvert}$	0.0505	**0.1**
$MAX_{\lvert PRR_{AB}-PRR_{BA}\rvert}$	**0.55**	0.4
$AVG_{\lvert \overline{RSSI}_{AB}-\overline{RSSI}_{BA}\rvert}$	1.67	**4.3**
$MAX_{\lvert \overline{RSSI}_{AB}-\overline{RSSI}_{BA}\rvert}$	4.7	**14.75**
$\%Unidirectional$	8	**45.83**

Table 3.4: Link asymmetry for flat and hill environments

Figure 3.13 shows the CDF of PRR and RSSI for both deployment areas. Based on these results, we observe links with very poor quality for the hill deployment (labeled “Hill”) compared to links on a flat area (labeled “Flat”), whether we consider the PRR or the RSSI metrics.



(a) CDF PRR : Distance 3M (b) CDF PRR : Distance 6M (c) CDF PRR : All links

(d) CDF RSSI : Distance 3M (e) CDF RSSI : Distance 6M (f) CDF RSSI : All links

Figure 3.13: CDF of PRR and RSSI for links with distances 3 m, 6 m, as well as for all the links in the network in the case of a hill and a flat area deployment.

The results presented in Figure 3.14 complement these findings. Indeed, while we observe 81.5% of good links on a flat area, we observe 70% of poor links on the hill deployment. As in previous sections, the majority of links in the network are either good or bad, with few intermediate links. Table 3.4 presents the link asymmetry metrics for the two different environments: the link asymmetry is also much more pronounced on the hill than on the flat area, with more than 45% unidirectional links. Our tests did not show a significant difference between the PRR and the RSSI values of uphill and downhill links, meaning that the radiation pattern of the CC2420 module antenna is not the reason for these results. We believe that the shape of the deployment environment is the main cause for this poor link quality, resulting in more signal reflections.

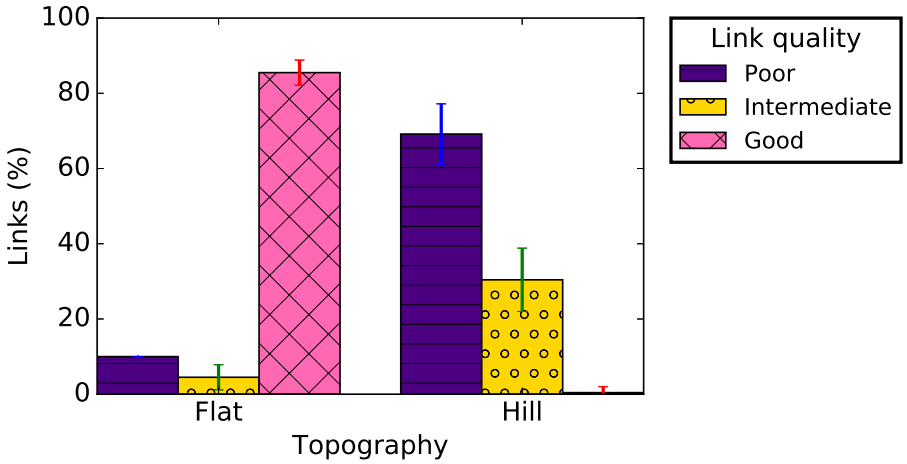


Figure 3.14: Link quality classification for different deployment environments.

3.6 Comparison to theoretical results

To strengthen the experimental results presented in this thesis, we compare the results obtained in our experiments with those produced by the theoretical model described in Section 3.3. In order to compare the empirical and theoretical results, we begin by estimating the path-loss exponent (ψ) and the shadowing standard deviation (ϕ) from the empirical data [84]. The idea is to find the values of ψ and ϕ that minimize the MSE (Mean Square Error) between the simplified model and the empirical power measurements. Using the methodology described in [84], we obtained the values shown in Table 3.5 in the case of the flat and hill deployments.

To strengthen the experimental results presented in this paper, we compare the results obtained in our experiments with those produced by the theoretical model described in Section 3.3. In order to compare the empirical and theoretical results, we begin by estimating the path-loss exponent (α) and the shadowing standard deviation (σ) from the empirical data [84]. The obtained values are shown in Table 3.5 in the case of the flat and hill deployments.

	ψ	ϕ
Flat deployment	3.56	6.08
Hill deployment	5.05	9.7

Table 3.5: Path-loss exponent and shadowing standard deviation for flat and hill deployment.

We use these values of ψ and ϕ as an input for the theoretical model described in Section 3.3, and then we generate theoretical values for the PRR and RSSI metrics. To check whether the empirical and theoretical distributions are comparable, we use the Mann-Whitney-Wilcoxon (MWW) test [135], a non-parametric test which does not make any assumption on the shape of the distributions. This test allow to check if two distinct sets of data come from the same distribution. The MWW test allows us to verify the null-hypothesis H_0: 'values from two independent sample sets (i.e. empirical and theoretical data in this case) come from the same distribution. In the

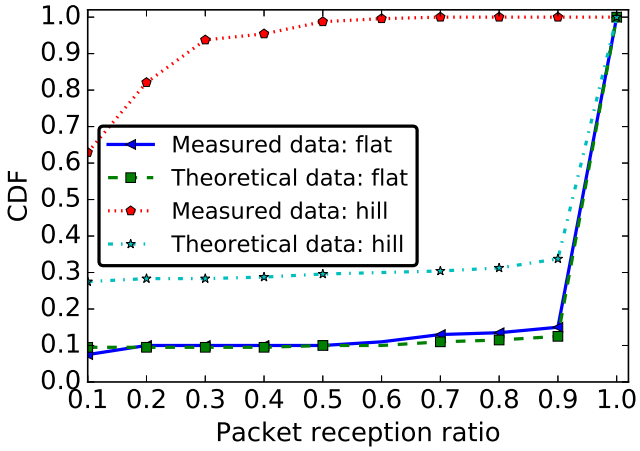


Figure 3.15: PRR Validation : Empirical and Theoretical CDF

case of the PRR metric, the MWW test validates H_0 in the case of the flat area deployment, but rejected it for the hill deployment. Meaning, if we consider a flat environment, the theoretical

model is close to the experimentation, while in the hill environment there is no correlation between theory and practise. This can also be observed in Figure 3.15, which shows the empirical and theoretical CDF for the flat and the hill deployments. This indicates that the poor radio link quality on the hill deployment is the consequence of a phenomenon that is not taken into account in the theoretical model. We also tested H0 for the PRR distribution on every individual link in our deployment: the null-hypothesis was rejected for 15% of links in the flat area deployment and for 37% of the links in the hill deployment.

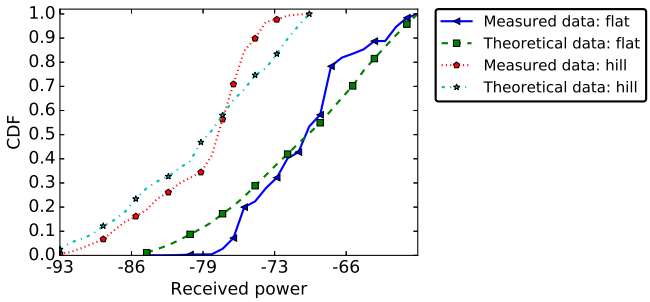


Figure 3.16: RSSI Validation: Empirical and Theoretical CDF.

Regarding the RSSI metric, the MWW test rejected the null-hypothesis H_0 for both deployment topographies (flat and hill) (See Figure 3.16). Meaning, there is no correlation between theory and practise results, whatever the considered environment is. This result is in line with the one presented in [93], where the authors conducted multiple statistical tests in order to check whether the empirical RSSI measured on a link can be approximated by a known probability distribution function. For all the probability distribution functions verified in their paper, the null-hypothesis was rejected. This can be explained by the fact that the theoretical model for RSSI considers the received power strength decreases with distance, a property we already showed not to be true in Section 3.5.4. In the per-link analysis, H_0 was rejected for 98% of the links in the flat area deployment and for 68% of the links in the hill deployment.

3.7 Recommendations for WSN designers

Because WSN at ground level have a lot of applications, including traffic monitoring, smart parking or bridges monitoring, we believe that engineering recommendations can be drawn from this study. Leading this experimental campaign allowed us to gain a significant insight regarding the deployment of wireless sensors at ground level, as well as regarding the properties of WSN links in general. Based on this experience, we formulate six recommendations for researchers and engineers interested in similar deployments.

Recommendation #1: Account for asymmetric links.

A classical and simple assumption when evaluating the communication protocols in WSN is that the links are symmetric: if sensor S_1 hears sensor S_2, S_2 also hears S_1. However, our results indicate that this property is not always true, especially when sensors are deployed at ground

level. Indeed, depending on the settings, up to 40% of the links in the network are unidirectional. This raises several important protocol problems. First of all, this means that solutions based on acknowledgment (ACK) messages might be inefficient, since the ACKs would be lost despite a successful message reception. Second, routing protocols need to take into account that ingress and egress routes might have very different properties.

> *Recommendation #2: Remember that the correlation between link quality and distance is weak.*

The physics of radio wave propagation predict a monotonic decrease of the received power strength with the distance. This is also a common assumption in analytical models and simulation tools. Nevertheless, our results confirm the findings of previous experimental campaigns, which demonstrate that this property is not true in practice. Moreover, the deviations from this predicted behavior seem to be exacerbated by a deployment at ground level. This also indicates the need for a denser deployment, which would allow to cope with the poor quality of certain links. At a protocol level, opportunistic links are not uncommon and could be exploited.

> *Recommendation #3: The communication channel 26 must be preferable and frequency hopping is better.*

Many studies on WSN do not consider the communication channel as a parameter, assuming the same quality on all transmission frequencies. In most field tests, designers evaluate their protocols on their testbed using only one channel. The results presented in Figures 3.10 and 3.11 show that communication performance varies depending on the channel used by the sensors for communication. In the particular case of IEEE 802.15.4, our results indicate that communication on channel 26, which is not shared with other technologies, is probably a good idea. However, the WSN performance should be tested on several communication channels in order to find the one with the most reduced interference. To be robust, frequency hopping on several channels will lead to a better robustness.

> *Recommendation #4: At ground level, high density or redundant deployment will be preferable.*

The results presented in this chapter, and particularly those presented Section 3.5.1, show that the link quality is worse at ground level when compared to a more classical deployment at height. Thus, in a WSN with sensors deployed at ground level, special measures must be taken not only during the design of the communication protocols, but also during the network deployment. If we assume that a link is usable only when its PRR is greater than or equal to 90% (a good link in Figure 3.4), for all pairs of sensors separated with a distance of 6 m, only 70% of the links will be valid in a WSN with sensors deployed at the ground level, compared to 100% for a deployment at height (Figure 3.6b). During network deployment at ground level, a higher sensor density

should be therefore preferred, in order to achieve a higher redundancy in the network and reduce the distance between the sensors, thus ensuring a certain reliability of the network. Results presented in this work suggest that the sensors density will depend on particular properties of the deployment area. Indeed, Figure 3.13 shows that, with sensors deployed at ground level, the link quality is even worse when the sensors are deployed on a hill compared to a flat area. However, increasing the sensor density in the network will also increase the contention and the collision at the medium access control layer. To deal with this issue, one will have to choose appropriate MAC layer protocols.

Recommendation #5: Very small messages should be avoided.

In several WSN applications, the data captured by a sensor is encoded by just few bits, e.g. a temperature value or a vehicle detection information. It might be tempting to transmit this data directly to a sink, in small packets with a payload of just a few bytes. Our experiments showed that the link quality is poor at ground level, no matter the packet size. However, results presented in Figure 3.8 show that messages with a very small application payload are more exposed. While we do not understand this phenomenon, aggregating multiple values in a single larger message not only will increase the packet reception probability, but will also reduce the contention and then the collision probability.

Recommendation #6: During the sensors deployment, the environment constraints must be taken into account.

The topography of the area where sensors are deployed, the networks and objects nearby, and the weather are some environment constraints that might have a negative impact on the radio link properties. In this work, we evaluate the impact of the topography of the environment by comparing the radio link properties considering a flat and a hill area. On the hill area, we do not measure the slope of the area. Nevertheless, Figure 3.13 shows a very poor radio link properties on the hill. Moreover, results presented in Section 3.5.1.3 show that, in the same network, the link properties also depend on the location of the sensor. Thus, to guarantee a reliable communication, during the sensors deployment or during communication protocols design, the environment characteristics must be taken into account.

3.8 Conclusion

Thanks to their potential applications, WSN with sensors deployed at ground level will play a bigger role in the future. In this chapter, we studied the properties of the radio link in a WSN with sensors deployed in such conditions. Our goal was not only to study the impact of the ground surface proximity on the radio link, but also to understand the impact of parameters, such as the message size, the communication channel frequency or the topography of the environment, on the quality of links between sensors deployed in such conditions.

For this, we conducted an extensive experimental campaign, comparing radio links at ground level with radio links in a network deployed above the ground. Our results indicate the feasibility of a WSN with sensors deployed at the ground level using off-the-shelf hardware. However, the link quality in this network is relatively poor compared with a deployment at height when considering the absolute values of the measured parameters (RSSI and PRR), the temporal behavior of the links, the spatial or asymmetrical properties of links.

We evaluated the impact of several parameters, such as the message size, the communication channel frequency, or the topography of the deployment area in the case of the ground level deployment. Our results showed a deteriorated link quality when the message size is too small, when the communication channels shared by other technologies, e.g. WiFi, are used, or when sensors are deployed on a hill.

Finally, we summarized our results as recommendations for WSN designers who might be interested in deploying ground level networks, for example in the case of large infrastructure monitoring. Some intriguing results obtained in our experiments, such as the poor PRR for packets with a very small application payload, or the poor link quality in the case of a hill deployment, open interesting research perspectives. Source codes and experiment data are available online through the link *"https://github.com/rdomga/WirelessLinkCharacterization"*.

Existing infrastructure-free WSN architectures for vehicular traffic monitoring at intersection consider a separation distance of 100 m between two sensors deployed per lane. While these architectures consider underground sensors, we considered in our experiment surface-mounted sensors, a more optimistic solution from a radio propagation point of view. Results presented in this chapter show a communication distance of 10-15 m, very far from the 100 m considered by the existing solutions. Based on this observation, we proposed in the next chapter WARIM, a reliable (in terms of network communications) WSN-based architecture for vehicular traffic monitoring at intersections. The communication protocols, as well as network mechanisms proposed for WARIM, must take into account the link properties highlighted in this chapter.

Chapter 4

WARIM: A Wireless sensor network Architecture for a Reliable Intersection Monitoring

As view in the previous chapters, WSN is a very promising candidate for ITS application because of its low-cost, network capacity, and connectivity properties. On one hand, existing infrastructure-based WSN architectures for vehicular traffic monitoring at an intersection present one main drawback: they need a road-side support to deploy the gateway, which increases the total cost of the architecture. On the other hand, existing infrastructure-free WSN architectures deploy two sensors per lane with a separation distance of 70-100 m between the two sensors. Experimental results presented in Chapter 3 show that it is not possible for nodes deployed in such conditions to communicate. Based on these experimental results, we propose, in this chapter, an infrastructure-free WSN-based architecture for vehicular traffic monitoring at an intersection. Our goal is to propose an architecture in which nodes can self-organize and can reliably communicate to report to the light controller data concerning vehicular traffic on each lane of the intersection.

This chapter is organized as follow. It starts by describing the problem statement in Section 4.1. Motivations for proposing a new WSN architecture for vehicular traffic monitoring at an intersection are presented in Section 4.2. Then, we present WARIM, the new WSN architecture in Section 4.3. The topology of the WSN formed by sensors deployed on a given lane is presented in Section 4.4. Challenges concerning the new architecture are discussed in Section 4.5, before the conclusion of this chapter in Section 4.6.

4.1 Problem statement

An ATLMS requires vehicular flow information, collected at the intersection level. The main drawback that can negatively impact the target application is the unreliable nature of the radio links of the network, which implies unreliable data gathering from the monitoring sensors. Indeed, results presented in Chapter 3 show an unreliable link at ground level. This may cause the loss of a higher number of packets transmitted by sensors to the intersection controller. Direct consequences of this may be the low accuracy of vehicle queue length estimation at a lane level

and the incapacity to detect and localize unauthorized stops around the intersection. Taking into account the poor radio link properties at ground level, our main goal in this chapter is to propose a reliable WSN architecture for vehicular traffic monitoring at an intersection.

A direct consequence of unreliable radio links is that the ATLMS may have a vision of the vehicular traffic far from the reality, and it can then produce suboptimal light plans. Thus, these factors must be considered when designing a WSN for vehicular traffic monitoring at an intersection. Moreover, the financial cost of the solution (hardware purchasing, deployment and maintenance) and the network lifetime must be taken into account. Indeed, if the financial cost of the solution is acceptable, this will encourage its implementation in many municipalities worldwide, and particularly in cities with reduced budget. The monitoring solution will also function as long as possible in order to reduce the frequency of maintenance and thus, vehicular traffic interruption.

Guided by the experimental results presented in Chapter 3, we propose WARIM (*Wireless sensor network Architecture for a Reliable Intersection Monitoring*), a lightweight WSN architecture for intersection monitoring. Our initial motivation is to provide a solution with a reasonable financial cost, in order to be accessible to most municipalities, especially in emerging countries. This leads us to consider low cost technologies and open standard communication protocols (e.g. IEEE 802.15.4). The hardware used in Chapter 3 for experimentations implements these technologies. Moreover, to reduce deployment and maintenance costs, we have to limit the human intervention, thus self-configuration and self-organization mechanisms must be developed. WARIM should also exhibit a long lifetime, thus the energy is a main concern, especially because nodes are equipped with batteries of limited capacity. Finally, to reduce deployment/maintenance time/cost and road deterioration, surface-mounted sensors must be used [12]. All this raises an important challenge from the networking point of view, since as shown in Chapter 3, communication between devices at ground level is characterized by a high degree of unreliability. Before presenting our solution, the next section summarizes our motivations to proposing a new WSN architecture for vehicular traffic monitoring at an intersection.

4.2 Motivations

A new WSN architecture for vehicular traffic monitoring at an intersection is proposed in this chapter. Our choice to propose a new WSN-based architecture is motivated among others by the low financial cost and flexibility of sensors, the drawbacks of existing results presented in Section 2.7 and the experimental results obtained in Chapter 3. These motivations are summarized in the current section.

4.2.1 The low-cost and flexibility of WSN

In Chapter 2, we presented different vehicles detection technologies used in ITS. Results presented in this chapter show that the magnetometer detector and video-image processor are the two mature vehicles detection technologies which allow to measure a high quantity of data (vehicle

count, length, speed, classification). Thanks to their low-cost, we chose in this thesis to use sensor nodes with magnetometer as vehicles detector and a radio module implementing IEEE 802.15.4 standard. This allows to keep reasonable the total financial cost of the hardware per intersection, particularly when many sensors must be deployed at each intersection.

Another advantage to use WSN is that the surface-mounted sensors are now available [12]. Compared to inductive loop or underground sensors, surface-mounted sensors are easy to install and significantly limit the road surface deterioration. The ease of installation property reduces the deployment time, and thus the financial cost of a WSN architecture based on this type of sensors. Moreover, by reducing the time required for sensors deployment, the vehicular traffic interruption duration is also reduced. Thus, surface-mounted sensors is an interesting choice when ease of deployment and the financial cost of the solution are criteria of a WSN architecture for vehicular traffic monitoring.

A multipurpose WSN can be developed by adding sensing modalities to the existing WSN [56]. Temperature sensors can be added to measure the road temperature, humidity sensors can be added to measure rain and accelerometer can be added to monitor structure of bridges and pavement. Gas sensor can also be added to measure the pollution in the city. This multifunction sensor nodes further extends the possibility of a more advanced ITS usage. By collecting all these data, this might allow to better organize or anticipate maintenance operations on the road infrastructure. It is difficult to provide such services when video-image processors are used.

The final advantage of the WSN is that the data collected can be easily made available on a central server for city-scale services. The primary purpose of data collection at an intersection is for intelligent traffic lights management. If data collected at all intersections at the whole city are stored on a central server, they can be used to deliver other more interesting services. For instance, the central server can interact with neighboring intersections in order to reduce the waiting time of vehicles at intersections. The data collected at the central server can also be used to offer services like routes guidance or vehicles tracking.

4.2.2 Existing solutions are not satisfactory

In Chapter 2, we presented existing infrastructure-free WSN-based architectures for vehicular traffic monitoring at intersections. In this thesis, we are interested in infrastructure-free architecture, i.e architecture which requires only surface-mounted sensors and not some other repeaters deployed along the roadway that can be used for data relay or data aggregation. Concerning infrastructure-free architectures discussed in the literature, [8, 10, 11] propose to deploy two sensors using the IEEE 802.15.4 standard per lane, separated by a distance of around 70-100 m. Experimental results presented in Chapter 3 show that, at ground level, it is not possible to establish a communication link between nodes separated by such a distance.

To solve this problem, a LPWAN technology, such as LoRa [40] or SigFox [41], can be used. These technologies allow low power and long range wireless communications. Thus, by deploying such technologies at intersections, a one-hop network in which each sensor deployed on the road may be form. For at least two reasons, we chose to use the IEEE 802.15.4 standard rather than

a LPWAN technology. Firstly, the IEEE 802.15.4 standard is nowadays a mature and low-cost technology compared to LPWAN. Secondly, the data rate is generally low (from 0.3 kbps to 27 kbps) with the LPWAN technologies, compared to 250 kbps of the IEEE 802.15.4 standard. Thus in case of dense network or high vehicular traffic or multipurpose WSN, the network capacity of LPWAN technologies may rapidly be exceed, which may cause low Packet Delivery Ratio (PDR) of LPWAN compared to IEEE 802.15.4. The non-guaranteed delay because of the low data rate of LPWAN is also an important drawback for the real-time vehicular traffic monitoring considering in this thesis. To address the problem of short range communication of the IEEE 802.15.4 standard, some authors [12, 47, 51] proposed to use gateways deployed along the road (infrastructure-based architectures). The problem of these gateways is that not only they obstruct the road, but they also increase the financial cost of the final solution.

4.2.3 Communications are possible at ground level

Results presented in Chapter 3 show that, when nodes are deployed at ground level, the ground has a negative impact on link quality. This limits the maximum sensors communication range in such a deployment. However, results also presented in Chapter 3 show that, at short distances in the order of 10-15 m, stable communication links can be established. In WSN, a sensor can sense and transmit data. A sensor can also forward a message coming from a neighbor node. The forwarding capability of nodes in a network allows nodes very far away to communicate using a multi-hop paradigm. By exploiting the forwarding capability of nodes in WSN, we propose in this chapter WARIM, a new WSN architecture for vehicular traffic monitoring at an intersection. Because of the low radio link quality, a dense deployment will help the network robustness, since redundancy offers several radio links opportunities.

4.3 A multi-hop, new WSN architecture for vehicular traffic monitoring at an Intersection

Motivated by the elements presented in the previous section, and considering the recommendations we made in Chapter 3, we propose, a new WSN architecture for vehicle count and queue length estimation for traffic light control at an intersection.

4.3.1 Proposed architecture

Based on the experimental results presented in the Chapter 3, we propose an architecture where sensors are deployed on the road surface, and they are able to communicate through a multi-hop WSN to report vehicle detection to the controller. In our proposition, WARIM (Fig. 4.1), several types of nodes are deployed at the intersection:

- **sensor:** deployed at ground level. It uses a magnetometer to detect vehicles and uses its radio module implementing the IEEE 802.15.4 standard to transmit this information to the sink;

- **sink:** collects the messages coming from sensors deployed on a given direction. Based on the collected data, the sink extracts and sends to the intersection controller information like the number of vehicles per lane and their waiting time;
- **intersection controller:** implements the traffic light management algorithm. Based on the data received from the sinks, the implemented algorithm produces the light plans for each direction.

The distance between two sensors and the number of sensors on a given lane is defined taking into account the communication technology, the radio link properties and the application requirements. For instance, if nodes communicate with radios implementing the IEEE 802.15.4 standard, the results presented in Chapter 3 show that the distance between two consecutive nodes will be in the order of $5 - 15\ m$, which is much smaller than the one proposed in related works. This short distance allows us to achieve two important goals. Firstly, communication between nodes is now reliable, even in the presence of vehicular traffic. Secondly, the dense deployment proposed in WARIM achieves a more accurate per-lane estimation. By using a small distance between sensors, WARIM provides an accurate measure of the vehicle queue length at lane-level, optimizing the functioning of the traffic light control algorithm.

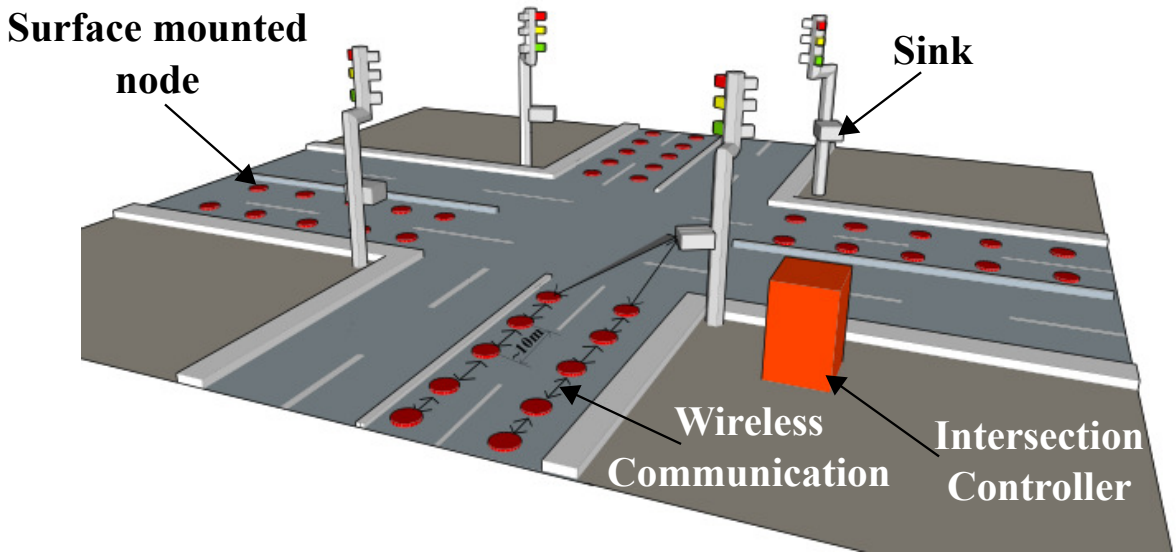


Figure 4.1: WARIM: Wireless sensor network Architecture for a Reliable Intersection Monitoring.

The number of sensors deployed on each lane depends on the desired accuracy and the node communication capabilities. The higher the number of deployed sensors, the better the accuracy, but a dense deployment also increases the WSN congestion, the collisions, the interferences and the financial cost of the monitoring system. This trade-off needs to be well understood before the deployment phase. The only requirement imposed by WARIM is that the deployment of the sensor nodes results in a connected network, considering a given level of accuracy in mind. If higher accuracy is required, a denser deployment should be used. WARIM uses only surface-mounted sensors without extra infrastructure, therefore highly reducing the deployment and maintenance costs, and also the vehicular traffic interruption during the deployment.

In Section 4.1, we highlighted the unreliability of the radio link as the main drawback of existing infrastructure-free WSN-based architectures for vehicular traffic monitoring at intersections. As consequences of this drawback, we cited the low accuracy of vehicle queue length at lane level and the incapacity to detect and to localize unauthorized vehicles stop that can have a negative impact on an ATLMS. Deploying several sensors per lane allows WARIM to deal with these drawbacks:

- WARIM defines the distance between consecutive sensors in order to guarantee link reliability;
- WARIM is able to detect when a vehicle changes lanes, and then improve the accuracy of the vehicle queue length estimation at lane level;
- WARIM is able to manage multiple lanes, but the Chapters 5, 6 and 7 focus on a single lane;
- finally, if sensors positions are known, many sensors per lane not only increase the chance to detect unauthorized stops, but also to localize them.

4.3.2 Discussions

The sensors we propose to use can be installed on the road surface with minor impact on the existing road infrastructure (i.e without civil works), therefore highly reducing the deployment and maintenance costs per sensor. But, one could argue deploying many nodes on a lane could significantly increase the financial cost of the solution proposed in this thesis compared to other detection technologies.

On one hand, one camera can be used to simultaneously measure the vehicle queue length on four to eight lanes of an intersection [64]. On the other hand, we propose in our architecture to install many (depending on the application requirements) sensors on each lane. When two or three sensors are deployed per lane, the magnetic sensors still present a low price compared to video-image processing technology [64]. But, by increasing the number of sensors on each lane, this increases the cost of our architecture and then might reduces the advantage cost of magnetic sensors. Nevertheless, for two reasons, we argue that, compared to camera, WSN using magnetometer as vehicles detectors is the best solution for vehicular traffic monitoring at an intersection.

1. Firstly, while a camera is limited by the maximum distance at which vehicles can be counted, with WARIM, the vehicles queue length that can be measured is limited by the number of sensors available, or the financial budget allocated for sensors purchasing, and the sensors communication capabilities. Indeed, the maximum distance at which two vehicles can be distinguished (and then counted) by a camera depends on the camera mounting height, the vehicle height and the separation distance between two consecutive vehicles. Note that, this distance is proportional to the separation distance and the mounting height, and is inversely

proportional to the vehicle height [14]. For instance, if we considered a camera mounted at a height of 10 m, an average vehicle height of 1.5 m and a separation distance (between two consecutive vehicles) of 4.5 m, the maximum distance at which the camera will be able to count vehicles on a lane is 27 m. Note that, by reducing the separation distance between two consecutive vehicles like in high congestion situation, this maximum distance is also reduced. Thus, the flexibility of WSN technology allows it to be implemented in different environments, taking into account the financial budget for sensors purchasing and their communication capabilities.

2. Secondly, even if ATLMS is the principal target application, the data collected at an intersection can have many other uses, including vehicular traffic monitoring in the whole city or vehicular tracking. The goal of traffic monitoring in the city is to have an idea on the traffic in the main points of the city at each time. Such information can then be used for road construction/maintenance planning and for car trip planning in the city. Vehicle tracking, classification, identification and re-identification are other services which may interest authorities. Concerning these services, for at least two reasons, we argue that magnetometer is the best choice compared to camera:

   - since each vehicle has a magnetic signature [16, 94], this information, which can be gathered by low cost sensors and processed by simple algorithms, is more suitable for vehicle tracking than other technologies like camera;
   - providing other services at the city scale requires the transmission of data from the intersection to the control center. With magnetometer, the only information transmitted to the control center is the vehicle magnetic signature. This information, encoded in some bytes, can be transmitted by a low-power, low-data rate and long-range technology like LoRa [40]. It might not be possible with camera which might require the transmission of large images size.

Moreover, as explained in Section 4.2.1, multi-purpose nodes can be used by WARIM. Thus, by equipping nodes not only with magnetometer to detect vehicles, but also by temperature, humidity or gas sensors, WARIM can be used to monitor the road state condition and the pollution in the city. Such parameters can be useful for road maintenance. For all these reasons, WSN is a good vehicles detection technology choice for vehicular traffic monitoring at an intersection.

4.4 From vehicular traffic monitoring to Linear Wireless Sensor Networks

In Fig 4.1, a set of nodes are deployed on different lanes of an intersection. The topology and architecture of WSN generally depend on the target application and the geographical area where nodes are deployed. If we consider nodes deployed on a particular lane of an intersection, they will form a WSN with a linear topology, or a LWSN.

The applications of LWSN are diverse, e.g. monitoring of large infrastructure such as bridges [57] and dams [58], and border control [59]. Thus, solutions proposed in this thesis can be applied not only to our use-case (vehicular traffic monitoring at an intersection), but also to all these applications. In the following section, we present some challenges of WARIM, the new infrastructure-free WSN architecture proposed in this chapter for vehicular traffic monitoring at an intersection.

4.5 Challenges

As mentioned above, WARIM is designed with the characteristic of ground level wireless links in mind, but it still has to tackle some important challenges in order to optimize networking operations. It is essential to propose a low cost solution that can be adopted by municipalities limited by financial budget. So, it would be interesting to develop an autonomous system both in terms of energy requirements and communication protocols, minimizing human intervention, which dominates the costs related to the deployment and maintenance of WSN. Since nodes are limited in their battery capacity, they are limited in their operation time. Thus, it would be interesting to take this into account by nodes deployment and communication protocols. Some challenges related to WARIM concern the network lifetime, self-organization and self-configuration, and data collection.

4.5.1 Network lifetime

In standard WSN, sensor nodes typically operate on a small capacity battery and are thus limited in their active lifetime [6, 7]. It may be infeasible or expensive to change batteries in sensor nodes once a wireless sensor network is deployed. One characteristic of a multi-hop WSN with single sink is the convergecast traffic, i.e. the traffic in the network is oriented toward the sink. In such a network, we have a high traffic intensity in the area close to the sink. Thus, sensors closest to the sink have a higher transmission/reception activity and tend to die early, causing network failure. Indeed, it is well known in the literature that the radio consumes most of the node energy for messages transmission and reception [110]. An exciting problem in WSN is to equally distribute traffic between nodes in the network in order to balance their energy consumption. This problem can be addressed through suitable deployment (the number of nodes deployed in each area should be proportional to the traffic in that area) or by designing energy-efficient communication protocols. In WARIM, this problem must be considered, particularly in the case of a WSN with linear topology.

4.5.2 Self-organization/Self-configuration

Network deployment and configuration are activities that increase the cost of WSN. In order to reduce deployment and configuration time, self-configuration protocols for different WSN applications are proposed in the literature [60]. One of the desired properties of WSN is the self-organization of nodes and protocols in the network. Indeed, using the wireless communication

channel, nodes can communicate to discover their neighborhood and to build a partial view of the network topology. Because of the instability of the wireless channel, the neighborhood of a node or the network topology is time-variant. This must be taken into account when building and maintaining the network topology. Developing self-configuration protocols is another important challenge in WSN. A configuration parameter of a sensor node for traffic monitoring can be its location: in WARIM, the location can be defined as the lane on which the node is deployed and its relative position with respect to the light controller. By allowing nodes in the network to run self-configuration protocols, this reduces the deployment time and subsequently the financial cost of WARIM. Self-configuration protocols also reduce the configuration error since the protocols ran by nodes can be validated before implementation in the physical network.

4.5.3 Data collection

Due to energy constraints and due to the environment, the communication range of wireless sensor nodes is generally limited. As shown by the results presented in Chapter 3, this is particularly true in vehicular traffic monitoring and for WARIM, where nodes are deployed at ground level. To send data to the controller, a multi-hop approach is required, which makes routing a key service for each node in the network. Numerous routing protocols have been proposed in the literature for WSN [63]. For energy-efficient routing protocols, battery level of nodes and topology control are generally used for routing decisions.

Developing an energy-efficient and reliable channel access mechanism is an important challenge in WSN, and particularly in WARIM. Indeed, nodes in WARIM use an open standard like IEEE 802.15.4. This standard uses frequencies in the non protected ISM bands which is shared by other technologies like Bluetooth or ZigBee. Thus, the channel access is a real challenge since communications by nodes in WARIM may be disrupted by communications in other networks deployed in the neighborhood. As we previously mentioned, many sensors must be deployed in the area close to the light controller in order to balance the traffic between nodes in the network. Deploying many nodes in this area makes the communications more complex, since this increases the collision probability. Nevertheless, we believe that the relative position of each node to the controller and the physical topology of the network could be combined and exploited to develop energy-efficient and reliable channel access solution for nodes in WARIM.

4.6 Conclusion

In this chapter, we deal with infrastructure-free vehicular traffic monitoring at road intersections using WSN technology. Existing infrastructure-free summarized in Section 2.7 are inaccurate in terms of the vehicles queue length estimation, and the proposed deployments are not suitable for network communications. In this chapter, first, we present our motivations for proposing a new infrastructure-free WSN-based architecture for vehicular traffic monitoring at an intersection. These motivations concern the flexibility and low-cost of WSN technology, the drawbacks of the existing architecture and the experimental results presented in Chapter 3.

We then propose a multi-hop WSN architecture, WARIM, to accurately measure the vehicles queue length on lanes of an intersection. It is an architecture in which several nodes are deployed on each lane, and where the maximum distance between two consecutive nodes on a lane is defined taking into account the communication technology and the deployment environment. In the proposed architecture, the maximum vehicle queue length that can be measured depends on the application, the environment and the financial budget available. The data collected at the intersection by WARIM can be used to offer other services at the city level. Some examples of these services are vehicular traffic monitoring at the city scale, vehicle classification, vehicular tracking and communications between neighboring intersections. Finally, we discuss some challenges that can be addressed to allow WARIM to satisfy its autonomous, energy-efficient, low-cost and lightweight properties. These challenges include: *i)* the network lifetime, *ii)* the self-organization and self-configuration of nodes and *iii)* data collection mechanisms.

Chapter 5

Sensor Deployment in Wireless Sensor Networks with Linear Topology using Virtual Node concept

The topology and architecture of WSN generally depend on the target application and the geographical area where sensors are deployed. A LWSN is a special case, where the physical topology of the network is a line. Considering WARIM, the infrastructure-free and WSN-based architecture proposed in Chapter 4 for vehicular traffic monitoring at an intersection, nodes deployed on a lane of an intersection form a LWSN.

In this chapter, we consider a LWSN in which a message generated by a sensor is forwarded in multi-hop until the sink. In WSN in general, sensor nodes typically operate on a small capacity battery and are thus limited in their active lifetime. It may be infeasible or expensive to change batteries in sensor nodes once a wireless sensor network is deployed. It is therefore important to design energy-efficient communication protocols and to optimize deployment strategies. In this chapter, we are interested in a simple, energy-efficient and realistic deployment strategy, which can allow a network to operate for a long time.

In a multi-hop wireless sensor network with a convergecast traffic, we observe a high traffic accumulation in the neighborhood of the sink. This area constitutes the bottleneck of the network, since the sensors deployed within it rapidly exhaust their battery. This problem can be addressed through suitable deployment, by designing energy-efficient protocols or by considering heterogeneous energy distribution. In this chapter, we are interested in simple, smart and realistic deployment strategies, which can allow a network to achieve the maximum operation time. Our goal is to propose for LWSN a deployment strategy that, while allowing the network to function as long as possible, guarantees a communication path between each node and the sink. The deployment strategy will also take into account messages received due to the overhearing phenomenon. We look for designing a solution that can be implemented in a relatively short time and by an operator who has not received any complex training.

In the remainder of this chapter, we start by presenting, in Section 5.1, the problem statement of this chapter. In Section 5.2, we present some related work on the problem of nodes deployment in a LWSN. In Section 5.3, we formalize the virtual node approach and Section 5.4 describes a

greedy algorithm for sensor deployment. The objective function of our greedy algorithm is to minimize the maximum number of operations (transmission or reception) per sensor. Section 5.5 presents analytical performance results. In Section 5.6, we discuss how our approach can be compared to the related work, while in Section 5.7 we consider networks in which physical sensors are heterogeneous in terms of battery capacity. Finally, Section 5.8 concludes this chapter.

5.1 Problem statement

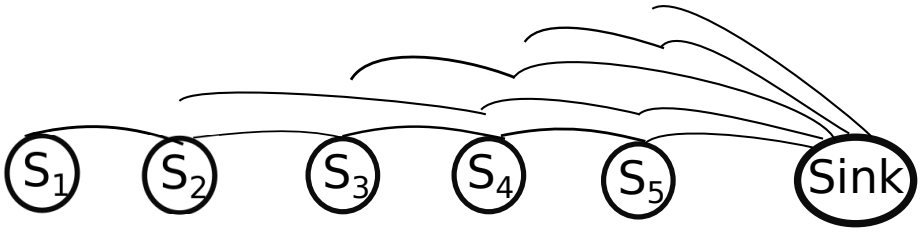


Figure 5.1: A LWSN of five sensors and one sink with convergecast traffic.

In a multi-hop LWSN with convergecast traffic (see Fig. 5.1), the amount of traffic that a sensor has to forward significantly increases as the distance to the sink becomes smaller [96, 97]. Moreover, sensors are more exposed to the overhearing phenomenon: a sensor can receive a message addressed to another one. Thus, sensors closest to the sink, because of their high traffic load and their high overhearing probability, tend to exhaust their battery early, causing network failure. This problem can be addressed through suitable deployment, by designing energy-efficient protocols or by considering non homogeneous energy distribution. This chapter deals with simple and energy-efficient sensors deployment in LWSN.

One efficient solution to the previous mentioned problem is to relocate sinks [99, 100, 101]. The idea behind these solutions is to regularly change the sink position in order to equilibrate the sensors' energy consumption. Usually, authors find the best way to relocate sinks by determining their optimal locations and the duration of their sojourn time. In this thesis, we are interested by a LWSN with static sensors and sink.

Many sensors deployment solutions are proposed in the literature [102, 103, 104, 109]. In these solutions, message receptions are not always taken into account, although transmission and reception consume more or less the same energy. Moreover, deployment recommendations are usually expressed in terms of distances between consecutive sensors. For a LWSN with a large number of sensors, a solution demanding such an accuracy on the inter-node distances may be costly and difficult to manage.

In a LWSN, an event that occurs in an area can be detected by many sensors deployed in that area. We propose a new deployment approach in which the LWSN is divided into virtual nodes and where consecutive virtual nodes should be wirelessly connected. An event happening within the area covered by the virtual node can be detected by any sensor, no matter its position in the virtual node. We then propose a greedy algorithm for calculating the number of physical sensors to deploy for a virtual node, with as objective function to minimize the maximum traffic

per sensor. In our work, we assume that a transmission power which guarantees connectivity between consecutive virtual nodes is used.

The concept of virtual node or virtual sensor is not new in the field of WSN. It is particularly used in the context of WSN programming [105, 106, 107]. In ,[106], every component represents a processing task applied to a stream of data originated from physical sensors which are modeled as a virtual node. In [105], a virtual node corresponds either to a data stream received directly from sensors or to a data stream derived from other virtual sensors. In this thesis, the concept of virtual node is used in the context of sensors deployment to reduce the complexity of the deployment.

Our solution is different from the existing ones on at least two points. First, there is no need to keep a precise distance between sensors. Indeed, after dividing the network into virtual nodes (the number of virtual nodes is smaller than the number of physical nodes), the position of sensors within a virtual node is not predefined. Second, contrarily to our solution, which considers also message receptions, the related work only account for message transmissions. This is not realistic because it is well known that most radio modules consume almost the same energy for receptions and transmissions [46]. It is important to note that message receptions mentioned here include receptions due to the overhearing phenomenon. Results presented in this chapter show that, by properly selecting the number of sensors, a simple greedy deployment can improve the network lifetime by up to 40%, when compared to the uniform deployment.

We also propose an alternative solution, which uses heterogeneous sensors. More precisely, we consider a network in which the battery capacities of sensors are not uniform, i.e the battery of a sensor is proportional to its traffic load. Results show that this solution significantly improves the network lifetime.

5.2 Related work

In a multi-hop WSN with a convergecast communication model, the sensors close to the sink have a high transmission activity. Thus, they constitute the bottleneck of the network, since sensors are constrained in terms of battery (limited capacity), memory (limited buffer size) and communication capabilities (high contention). Many energy-efficient solutions have been proposed at all layers of the communication stack, ranging from hardware design, MAC and routing protocols, and data aggregation [98]. In this thesis, we are interested in a sensor deployment strategy which could prolong the network lifetime.

Sensor deployment strategies can be designed with different objectives in mind. For example, the objective of the deployment can be to guarantee a balanced energy consumption of sensors in the network. To address this problem, one can deploy more sensors close to the sink and make them transmit at lower power levels [102, 103, 104]. In this way, sensors close to the sink transmit more messages, but consume less energy per message. When all sensors use the same transmission power and sensor density is high close to the sink, a message can be forwarded by

a significant number of sensors. Load-balancing can then be achieved by adopting duty-cycle scheduling [108].

The problem of sensors deployment is widely addressed in the literature [115, 116, 117, 118, 119, 120]. In [120], authors propose a hybrid differential evolution and simulated annealing (DESA) algorithm for clustering and choice of cluster heads. They prolong the network lifetime by appropriately selecting the cluster head and affecting sensors to clusters. [115] addresses the optimal deployment problem defined as the determination of the minimum possible number of sensors aiming to achieve the targeted partial connected coverage, the lowest financial cost and the highest lifetime. In [119], authors propose a novel node deployment strategy based on quasi-random method of low-discrepancy sequences to increase the lifetime and the coverage of the network. While most of the previous work consider 2-D or 3-D topologies, this paper focuses on LWSN, i.e. a WSN with 1-D topology. Rather than considering a hierarchical network (like in clustered networks), we are interested by flat WSN in which all the messages generated by sensors are forwarded in a multi-hop manner toward the sink deployed at a network edge.

In [97], authors propose an analytic description of the traffic over a linear, randomly deployed WSN. They evaluate the effect of the number of sensors and their distribution on network traffic. They propose a non uniform deployment obtained through an increased network density closer to the sink. Their results show that, given a number N of sensors, such a deployment can significantly reduce the maximum traffic load per node compared to a uniform deployment, and hence improve the network lifetime. However, the authors do not propose a deployment strategy or some concrete recommendations for sensors deployment.

In [109], the problem of sensor deployment is addressed when all sensors have the same transmission power level. Given the required lifetime of a sensor network, the energy constraint of sensors, and the area to be covered, authors study the problem of finding the minimum number of sensors needed to build such a network and the corresponding deployment scheme. In [102] , the authors assume that a sensor can select the transmission power among a set of power levels. Given the energy constraint of sensors, the problem of sensors deployment is addressed with the goal of maximizing the covered area and the network lifetime. They formalize the deployment problem as a Mixed-Integer Linear Programming (MILP) problem. Their results show that, by properly selecting the number of sensors, the distance between them, and the corresponding transmission power, the WSN lifetime can be improved by up to 30%.

Despite their performance in terms of network lifetime, a drawback of the solutions proposed by [102, 109] is the complexity of their implementation, particularly in large LWSN. Indeed, the deployment recommendations are expressed in terms of distances between consecutive sensors in the network. This is not practical if the deployment must be done by a human considering a network with a large number of sensors. In this chapter, we focus on a solution that is simple to implement, i.e. it requires reduced deployment effort, and can at the same time improve significantly the network lifetime compared to a uniform deployment. Another important contribution with respect to the state of the art is that, because of the broadcast nature of the

Parameter	Description
N	Number of sensors
S_i	Sensor deployed at position $i, 1 \leq i \leq N$
K	Number of virtual nodes
B_k	The K^{th} virtual node, $1 \leq k \leq K$
λ	Avg. number of messages generated per virtual node per time slot
Δ	Time slot duration
p	Probability for the virtual node B_i to receive messages from virtual node B_{i-2} or B_{i+2}
α	Probability for a sensor to receive messages destined to another sensor
E_{sensor}	The sensor battery capacity
P	The average transmission power
E_i	Energy consumed per sensor per time slot in the virtual node B_i
τ	The time required to transmit one packet
T_i	The total number of transmissions from a virtual node B_i: generated and relayed messages
n_i	Number of sensors in the virtual node B_i
$\overline{R}_i^{ns}$	Avg. number of receptions per sensor in virtual node B_i when the radio of nodes is always on
$\overline{R}_i^{os}$	Avg. number of receptions per sensor in virtual node B_i when there is no overhearing (perfect scheduling)
$\overline{R}_i^{ps}$	Avg. number of receptions per sensor in virtual node B_i when the overhearing is partial (practical scheduling)
$\overline{O}_i$	Avg. number of operations per sensor per time slot in the virtual node B_i

Table 5.1: Model Notations.

wireless communication channel, the proposed solution takes into account messages received due to the overhearing phenomenon.

5.3 Problem description

Given a set of homogeneous sensors in terms of battery capacity, our goal is to design a simple, energy-efficient and realistic sensor deployment that maximizes the network lifetime. In Section 5.7, we investigate the case where sensors are heterogeneous.

In many WSNs applications, the same event can be detected by several sensors if the network is dense. For instance, if a WSN is deployed to monitor the vehicular traffic on a highway or on lanes of an intersection, a vehicle might be detected at the same time by different sensors. This depends on the average vehicle length and the distance between two consecutive nodes. Another example is the monitoring of the status of an infrastructure like a bridge. In such application, a damage on an area of the bridge could be detected by many sensors. Thus, rather than considering a flat network like in related work [102, 109], we consider a network divided into virtual nodes wirelessly connected. A virtual node represents the set of sensors deployed in a given area which can measure the same physical phenomenon. In the related work, the deployment recommendations are expressed in terms of distance between consecutive sensors. In

our approach, once the virtual nodes are defined, the positions of sensors in a virtual nodes are not predefined. In the next section, our hypothesis are further detailed. Tab. 5.1 presents the list of important notations used in our model.

5.3.1 Assumptions

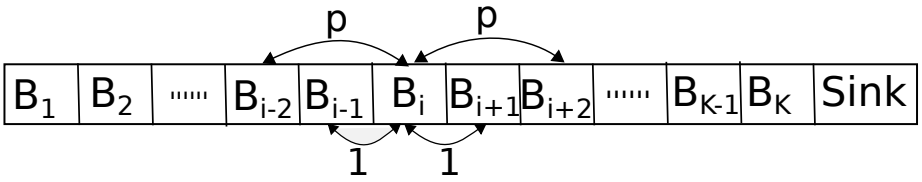


Figure 5.2: Network Model.

We consider a set of N sensors $\{S_1, \cdots, S_N\}$ deployed in a $1D$ space to monitor a phenomenon. We assume that the area to be monitored is covered by K virtual nodes $\{B_1, \cdots, B_K\}$, as presented in Fig. 5.2. We also assume that the sink is a single node which corresponds to the virtual node B_{K+1}. In our work, we assume that consecutive virtual nodes are connected, in order to guarantee that a message generated in a virtual node can be received by the sink. However, note that this assumption does not imply that physical nodes use homogeneous power levels, since the physical distance between nodes is not necessarily uniform, and several physical nodes can form a virtual node.

An event happening in an area can be detected by many sensors. We consider that, for each event detected by a virtual node, only one message will be generated, as a result of a cooperation algorithm executed by sensors of the same virtual node. This cooperation algorithm also equally distributes to the nodes in B_i the traffic coming from B_j, $j < i$. We also assume that sensors can run a ranking algorithm (see Chapter 6) to determine the relative position of each sensor to the sink. Such an algorithm allows sensors to compute the virtual node index in which they are deployed.

5.3.2 Connectivity model and traffic pattern

We assume a multi-hop LWSN with a convergecast communication model, i.e data from all sensors are forwarded toward the sink located at position $K + 1$. Unlike a distributed peer-to-peer wireless network, the traffic load is highly asymmetric, i.e. sensors closer to the sink have a heavier relay load. We assume a wireless connectivity between consecutive virtual nodes, B_i and B_{i+1}. We also assume that a sensor in the virtual node B_i can communicate with sensors in B_{i+2} or B_{i-2} with probability p (see Fig. 5.2). We consider a shortest path routing protocol [111]. Thus, a packet transmitted by B_i and received by B_{i+1} and B_{i+2} is forwarded by B_{i+2}. We denote by λ the average number of messages generated from a virtual node per time unit (Δ). Typically, λ is application dependent and represents the average number of events detected by a virtual node during a time slot. We denote by T_i the total number of transmissions (messages

generated or relayed) of virtual node B_i per time unit. This traffic (T_i) comes from three different sources:

- λ: messages generated by the virtual node B_i itself
- $(1-p)\cdot T_{i-1}$: messages relayed by B_i from B_{i-1}
- $p\cdot T_{i-2}$: messages relayed by B_i from B_{i-2}.

Thus, the traffic load of virtual node B_i is given by Eq. 5.1.

$$T_i = \begin{cases} \lambda, \text{ if } i = 1 \\ (2-p)\cdot\lambda, \text{ if } i = 2 \\ p\cdot T_{i-2} + (1-p)\cdot T_{i-1} + \lambda, \text{ if } 2 < i \leq K \end{cases} \tag{5.1}$$

Eq. 5.1 is a recurrence equation whose solution is given by Eq. 5.2.

$$T_i = \beta + \gamma\cdot(-p)^i + \frac{\lambda\cdot i}{p+1}, \forall i, 1 \leq i \leq K \tag{5.2}$$

In Eq. 5.2, β and γ are constants calculated from T_1 and T_2 (see Eq. 5.1). Thus, $\beta = \frac{p\cdot\lambda}{(p+1)^2}$ and $\gamma = -\frac{p\cdot\lambda}{(p+1)^2}$. The traffic load in the virtual node $B_{i,1\leq i\leq K}$ is thus given by Eq. 5.3.

$$T_i = \frac{\lambda}{p+1}\cdot\left[\frac{p}{p+1} + \frac{(-p)^{i+1}}{p+1} + i\right], \forall i, 1 \leq i \leq K \tag{5.3}$$

Thanks to the scheduling algorithm used, we assume a uniform distribution of traffic between sensors in a given virtual node. The average number of transmissions per sensor in the virtual node B_i, which contains n_i sensors, is:

$$\overline{T}_i = \frac{T_i}{n_i}, \forall i, 1 \leq i \leq K \tag{5.4}$$

5.3.3 Number of operations per sensor

In a WSN, a sensor consumes energy during data sensing and processing, and through its radio for data transmission and reception. It is well known that the radio generally consumes most of the energy of the sensor [110]. In this chapter, we denote as energy-consuming operations only the transmission and the reception of a packet. Thus, the number of operations executed by a sensor is the total number of messages sent and received by that sensor. The total number of operations per virtual node is equal to the number of operations of all the sensors which form the virtual node. But, since our goal is to maximize the sensor lifetime, in this section, we focus on the number of operations per sensor. We have already seen that the average number of transmissions per sensor deployed in a virtual node is given by Eq. 5.4.

A sensor may receive a message destined to another sensor: it is the overhearing phenomenon [112]. Therefore, the number of receptions observed by a sensor depends on the scheduling algorithm used. We assume in this chapter that the goal of the scheduling algorithm is to determine the time at which a sensor must switch its radio on or off. Thus, a perfect scheduling is when a sensor switches its radio on only when it has a message to send or to receive. In order to model the number of receptions per sensor per time period, different cases are considered in the following:

- No scheduling, where sensors keep their radio always on and ready to transmit/receive
- Optimal Scheduling, where the overhearing is eliminated by a perfect scheduling.
- Practical Scheduling, where the overhearing is partially, but not totally, eliminated.

5.3.3.1 Case 1: No scheduling

We assume here that each sensor maintains its radio on all the time. In such a situation, a sensor will receive all transmissions in its communication range. This includes transmissions from neighboring virtual nodes and from the virtual node to which the sensor belongs. For a virtual node B_i, $1 \leq i \leq K$, there are two sources of messages:

We assume here that each sensor maintains its radio on all the time. In such a situation, a sensor will receive all transmissions in its communication range. This includes transmissions from neighboring virtual nodes and from the virtual node to which the sensor belongs. For a virtual node B_i, $1 \leq i \leq K$, there are two sources of messages:

- Receptions from neighboring virtual nodes: this concerns transmissions from virtual nodes $B_{i-2}, B_{i-1}, B_{i+1}$ and B_{i+2}. This quantity is equal to $p \cdot T_{i-2} + T_{i-1} + T_{i+1} + p \cdot T_{i+2}$, since we assume that a sensor in a virtual node B_i can always receive transmissions from virtual nodes B_{i1} and B_{i+1}, and it can communicate with virtual nodes B_{i2} and B_{i+2} with a probability p. We define $T_i = 0$ for $i < 1$ and $i > K$
- Messages transmitted by other sensors in the same virtual node: $T_i - \frac{T_i}{n_i}$. Indeed, for a traffic load T_i in virtual node B_i, the average number of transmissions for a given sensor S_j is given by Eq.5.4. Then, the remaining messages $\left(T_i - \frac{T_i}{n_i}\right)$ will be transmitted by other sensors. Sensor S_j will also receive these messages, since we assume each sensor continuously maintains its radio on.

The total number of receptions is then given by Eq. 5.5.

$$\overline{R}_i^{ns} = p \cdot T_{i-2} + T_{i-1} + T_{i+1} + p \cdot T_{i+2} + T_i - \frac{T_i}{n_i} \tag{5.5}$$

Under the assumption of the radio always on, the average number of operations per sensor per time unit in a virtual node B_i is the sum of transmissions (Eq. 5.4) and receptions (Eq. 5.5) per sensor and is expressed by Eq. 5.6.

$$\begin{aligned} \overline{O}_i &= \overline{T}_i + \overline{R}_i^{ns} = \frac{T_i}{n_i} + p \cdot T_{i-2} + T_{i-1} + T_{i+1} + p \cdot T_{i+2} + T_i - \frac{T_i}{n_i} \\ &= p \cdot T_{i-2} + T_{i-1} + T_{i+1} + p \cdot T_{i+2} + T_i \end{aligned} \tag{5.6}$$

From Eq. 5.6, an interesting property appears: the average number of operations per sensor per time unit is independent of the number of sensors deployed in the virtual node. Thus, a simple solution is to consider the same number of sensors per virtual node; this corresponds to a uniform deployment. Under the assumption of the radio always on, the optimal solution is to deploy one sensor in each virtual node, since it is useless to deploy more sensors in a virtual node.

5.3.3.2 Case 2: Optimal scheduling

We assume in this section an optimal scheduling algorithm where a sensor is able to switch its radio on only when there is a message to forward, or only when it has a new message to transmit. If such a scheduling algorithm is used, there will be no overhearing and the messages received by a sensor will be only those it needs to relay for the previous virtual nodes. The traffic relayed by sensors in the virtual node B_i is $T_i - \lambda$. Since we assume a uniform distribution of this traffic among sensors in the virtual node B_i, the average number of receptions per sensor is given by Eq. 5.7.

$$\overline{R}_i^{os} = \frac{T_i - \lambda}{n_i} \tag{5.7}$$

Thus, the average number of operations per sensor per time unit is the sum of transmissions (Eq. 5.4) and receptions (Eq. 5.7) and is given by Eq. 5.8.

$$\overline{O}_i = \overline{T}_i + \overline{R}_i^{os} = \frac{T_i}{n_i} + \frac{T_i - \lambda}{n_i} = \frac{2 \cdot T_i - \lambda}{n_i} \tag{5.8}$$

From Eq. 5.8, we observe that, when the number of sensors n_i increases, $\overline{O}_i$ decreases. Thus by increasing the number of sensors deployed in a virtual node, the network functioning time increases too.

From Eq. 5.8, we observe that, when the number of sensors n_i increases, $\overline{O}_i$ decreases. Thus by increasing the number of sensors deployed in a virtual node, the network functioning time increases too.

5.3.3.3 Case 3: Practical scheduling

In a wireless network, because of multiple constraints (environment, hardware, communication technology, etc), scheduling algorithms [112, 113, 114] are not perfect. A sensor may receive a message addressed to another one: this is the overhearing phenomenon. Thus, if we consider a particular sensorS_j deployed in the virtual node B_i, in addition to the messages transmitted by previous virtual nodes and relayed by S_j, S_j may receive other transmissions in its communication range. We model this effect by a parameter α. This parameter represents the ratio of messages transmitted in the communication area of a sensor S_j, not addressed to S_j, but received by S_j. In the following, α will be denoted as the overhearing ratio. To calculate the total number of messages that S_j can receive due to the overhearing phenomenon, the following components must be considered:

In a wireless network, because of multiple constraints (environment, hardware, communication technology, etc), scheduling algorithms [112, 113, 114] cannot be perfect. A sensor may receive a message addressed to another one: this is the overhearing phenomenon. Thus, if we consider a particular sensor S_j deployed in the virtual node B_i, in addition to the messages transmitted by previous virtual nodes and relayed by S_j, S_j may receive other transmissions in its communication range. We model this effect by a parameter α. This parameter represents the ratio of messages transmitted in the communication area of a sensor S_j, not addressed to S_j, but received by

S_j. In the following, α will be called the overhearing ratio. To calculate the total number of messages that S_j can receive due to the overhearing phenomenon, the following components must be considered:

- Messages transmitted (generated or relayed) by sensors in the virtual node B_{i-1}, but relayed by sensors in the virtual node B_{i+1}

$$\overline{R}_i^{rsnr} = p \cdot T_{i-1} \tag{5.9}$$

- Messages transmitted (generated or relayed) by previous virtual nodes, relayed by sensors in the virtual node B_i, but a different node from S_j

$$\overline{R}_i^{rsr} = (T_i - \lambda) - \frac{T_i - \lambda}{n_i} = \frac{(T_i - \lambda) \cdot (n_i - 1)}{n_i} \tag{5.10}$$

- Messages transmitted by other sensors in the virtual node B_i:

$$\overline{R}_i^{rsc} = T_i - \frac{T_i}{n_i} = T_i \cdot \frac{n_i - 1}{n_i} \tag{5.11}$$

- Messages transmitted by sensors in virtual nodes B_{i+1} and B_{i+2}:

$$\overline{R}_i^{rsf} = T_{i+1} + p \cdot T_{i+2} \tag{5.12}$$

Thus, the number of messages received by sensor S_j due to overhearing is defined by Eq. 5.13.

$$\begin{aligned} \overline{R}_i^{o} &= \alpha \cdot (\overline{R}_i^{rsr} + \overline{R}_i^{rsnr} + \overline{R}_i^{rsf} + \overline{R}_i^{rsc}) \\ &= \alpha \cdot \left[\frac{(T_i - \lambda) \cdot (n_i - 1)}{n_i} + p \cdot T_{i-1} + T_{i+1} + p \cdot T_{i+2} + T_i \cdot \frac{n_i - 1}{n_i} \right] \\ &= \alpha \cdot \left[\frac{(2 \cdot T_i - \lambda) \cdot (n_i - 1)}{n_i} + p \cdot T_{i-1} + T_{i+1} + p \cdot T_{i+2} \right] \end{aligned} \tag{5.13}$$

To obtain the total number of receptions of S_j, we also consider the messages transmitted from previous virtual nodes and relayed by S_j (see Eq. 5.7). Eq. 5.14 gives the average number of receptions per sensor deployed in a virtual node B_i.

$$\begin{aligned} \overline{R}_i^{ps} &= \overline{R}_i^{os} + \overline{R}_i^{o} \\ &= \frac{T_i - \lambda}{n_i} + \alpha \cdot \left[\frac{(2 \cdot T_i - \lambda) \cdot (n_i - 1)}{n_i} + p \cdot T_{i-1} + T_{i+1} + p \cdot T_{i+2} \right] \end{aligned} \tag{5.14}$$

The average number of operations per sensor per time unit is then given by the Eq. 5.15.

$$\begin{aligned} \overline{O}_i &= \overline{T}_i + \overline{R}_i^{ps} \\ &= \frac{T_i}{n_i} + \frac{T_i - \lambda}{n_i} + \alpha \cdot \left[\frac{(2 \cdot T_i - \lambda) \cdot (n_i - 1)}{n_i} + p \cdot T_{i-1} + T_{i+1} + p \cdot T_{i+2} \right] \\ &= \frac{(2 \cdot T_i - \lambda)}{n_i} + \alpha \cdot \left[\frac{(2 \cdot T_i - \lambda) \cdot (n_i - 1)}{n_i} + p \cdot T_{i-1} + T_{i+1} + p \cdot T_{i+2} \right] \end{aligned} \tag{5.15}$$

Note that, the optimal scheduling corresponds to the case $\alpha = 0$, whereas no scheduling scenario corresponds to the case $\alpha = 1$. It is also important to note that $\overline{O}_i$ decreases when n_i increases. But when n_i is very large: $\overline{O}_i \approx \overline{O}_i^* = \alpha \cdot (2 \cdot T_i - \lambda + p \cdot T_{i-1} + T_{i+1} + p \cdot T_{i+2})$. Thus, if $n_i = n_i^*$ ($\overline{O}_i \approx \overline{O}_i^*$) then, it is useless to deploy more than n_i^* sensors in the virtual node B_i, since this will not improve the lifetime of sensors in B_i.

5.3.4 Problem formulation

Given N sensors that form K virtual nodes, our goal is to find the number n_i of sensors in virtual node B_i, $1 \leq i \leq K$, such that $\sum_{i=1}^{K} n_i = N$, and which minimizes the maximum number of operations executed by a sensor. Formally, the problem can be formulated as:

$$minimize \quad \max_{1 \leq i \leq K} \overline{O}_i \tag{5.16}$$

subject to the following constraints:

$$\sum_{i=1}^{K} n_i = N \tag{5.17}$$

$$n_i \geq 1, i = 1, \cdots, N \tag{5.18}$$

Eq. 5.16-5.18 define a mixed-integer nonlinear programming problem. In the following section, given the number N of sensors, we propose a greedy algorithm to calculate, for each virtual node B_i, the value of n_i.

5.4 Deployment algorithm

We assume in this work that a sensor is declared dead when it exhausts its battery. We assume a uniform load distribution within a virtual node. This means that all sensors in the same virtual node will die approximately at the same time. Thus, we define the network lifetime as the time until all sensors in a virtual node die (or, equivalently, until the first sensor dies). The network lifetime is maximized by reducing as much as possible the traffic load per sensor. The problem addressed in this chapter can be formulated as follows: given N sensors forming K virtual nodes, how many sensors should be deployed in each virtual node in order to maximize the network lifetime? Hereafter, the average number of operations per sensor per time unit in the virtual node B_i is denoted $\overline{O_i}$ (see Eq. 5.15). We propose in Algorithm 5.1 a greedy approach to compute the number of sensors to deploy in each virtual node. Recall that, in each virtual node, we will have at least one sensor, for connectivity and sensing purposes. That is why the number N of sensors is greater than or equal to the number K of virtual nodes. Initially, one sensor is assigned to each virtual node (Line 4) and the remaining sensors are iteratively deployed. At each iteration, one sensor is added to the virtual node which has the highest $\overline{O}_i$. If more than one virtual node have the maximum number of operations per sensor per time unit, the one with the highest index is selected (Lines 10-11): it is the one which is closest to the sink. Indeed, the traffic is highest as the virtual node is closer to the sink. The consequence of adding a sensor to a virtual node B_i is the reduction of the average number of operations per sensor per time unit in that virtual node. Since our objective is to propose a deployment which minimizes the maximum number of operations per sensor, the problem can be decomposed in a series of sub-problems, which are optimally solved iteratively by the greedy Algorithm 5.1, resulting in an optimal general solution.

Algorithm 5.1 Deployment algorithm

Input: $N, K, \lambda, \alpha, p, N \geq K$
Output: $n_i, 1 \leq i \leq K, n_i \geq 1, \sum_{i=1}^{K} n_i = N$

for Each virtual node B_i **do**
  calculate T_i using Eq. 5.3
end for
$n_i \leftarrow 1, 1 \leq i \leq K$
$remaining\_sensors \leftarrow N - K$
while $remaining\_sensors \neq 0$ **do**
  for Each virtual node B_i **do**
    calculate $\overline{O_i}$ using Eq. 5.15
  end for
  $I \leftarrow \{i | \overline{O}_i = \max_j \overline{O}_j\}$
  $i \leftarrow max\{I\}$
  $n_i \leftarrow n_i + 1$
  $remaining\_sensors \leftarrow remaining\_sensors - 1$
end while

At the beginning of this algorithm, the traffic of each virtual nodes is calculated. We assume that, for a virtual node B_i, T_i is calculated in $O(K)$ basic CPU operations (addition, subtraction, etc.). Thus, calculating this value for all virtual nodes is done in $O(K^2)$. In the iterative part of the algorithm, at each iteration, the average number of operations per sensor per virtual node is calculated. This processing can be done in $O(K)$. Since the goal of the iterative part of the deployment algorithm is to find the virtual node in which each of the NK (since initially one sensor is deployed per virtual node) virtual nodes will be deployed, the iterative part of our algorithm can be executed in $O(N \cdot K)$. Thus, the result of the deployment is computed in $O(K^2 + N \cdot K)$. Even for a large network with a high number of sensors, it is possible to rapidly obtain the number of sensors which forms each virtual node, since this algorithm can be executed on a computer with good performance, and not on the nodes themselves.

5.5 Analytical results

In this section, we present analytical results concerning the virtual node-based approach proposed in this thesis. These results concern the greedy deployment described by Algorithm 5.1 and a simple uniform deployment. In the latter one, the same number of sensors are deployed in all virtual nodes. Firstly (see Section 5.5.2), we compare the two deployments in terms of sensors distribution for different parameters. Secondly (see Section 5.5.3), the two deployments are compared in terms of network lifetime. In this thesis, we define the network lifetime as the time until the first sensor exhausts its energy. The next section describes the evaluation parameters.

5.5.1 Evaluation parameters

Without loss of generality, we assume that the average output power is $P_t = 61.9\ mW$, which is the maximum transmission power of the radio used by *Tmote Sky* [102]. We also assume that

each sensor has the same battery capacity of $E_{node} = 5\ Ah$. If we assume a packet size of 128 Bytes and a data rate of 250 $kbps$, the time τ to transmit one packet is equal to 4.096 ms. Thus, E_i, the average energy consumed by a sensor in the virtual node B_i per time unit is defined by Eq. 5.19.

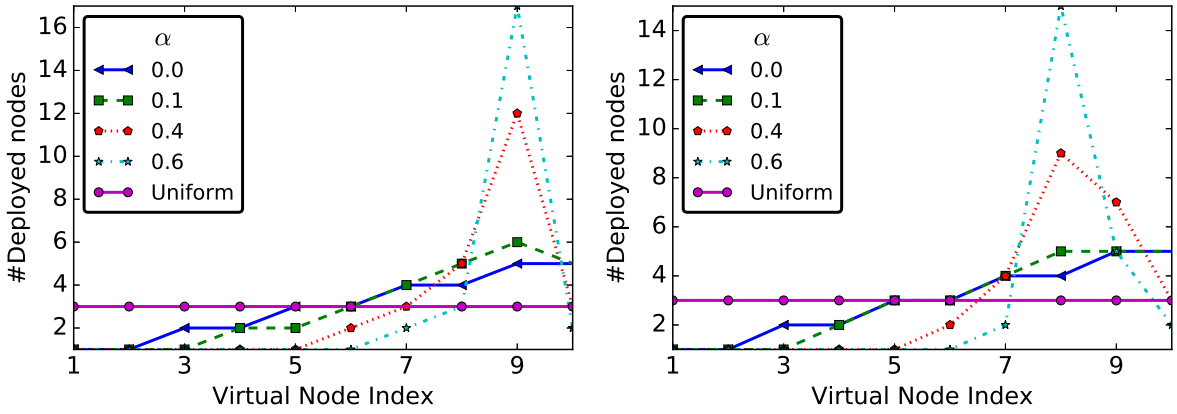
$$E_i = \overline{O}_i \cdot \tau \cdot P \tag{5.19}$$

Tab. 5.2 summarizes the values of parameters we use to obtain numerical results.

Parameter	**Description**	**Value**
E_{node}	The sensor battery capacity	5 Ah
P_t	Average transmission power	61.9 mW
τ	Time to transmit one message	4.096 ms
Δ	Time slot duration	5 s
λ	Avg. number of messages generated per virtual node per time slot	1 message

Table 5.2: Evaluation parameters.

5.5.2 Spatial nodes distribution



(a) $p = 0.0$ (b) $p = 0.5$

Figure 5.3: Number of nodes deployed in each virtual node: $N = 30, K = 10$.

Fig. 5.3 presents the sensors distribution obtained by a uniform deployment and the greedy Algorithm 5.1. It gives the number of sensors forming each virtual node as a function of the virtual node position and the overhearing ratio (α) when we consider $N = 30$ sensors to deploy in a network of $K = 10$ virtual nodes. Note that, when the number of virtual nodes is 10, the sink is deployed alone in a particular virtual node (not represented in Fig. 5.3) at the position $K + 1$. Fig. 5.3a and 5.3b are for $p = 0$ and $p = 0.5$ respectively (p being the probability of a sensor in the virtual node B_i to receive messages from B_{i-2} or B_{i+2}).

With the greedy deployment proposed by Algorithm 5.1, we observe that there is always a virtual node B_{i*} which receives the highest number of sensors. The position of this virtual node

depends on p and is independent of α. When $p = 0.0$, B_9 is the bottleneck of the network (highest traffic) and receives the maximum number of sensors (Figure 5.3a). Indeed, unlike sensors in B_{10} which receive messages only from the left side, sensors in B_9 receive messages from both the left and the right side. When $p = 0.5$ (Fig. 5.3b), the virtual node B_8, which receives messages from "two-hop" virtual nodes, becomes the bottleneck of the network and is assigned the maximum number of sensors. For other virtual nodes which are farther from the sink than B_{i*}, the number of sensors deployed in a virtual node increases linearly as a function of the virtual node index.

When only transmissions are considered, like in most of the related work, the last virtual node of the network will always receive the maximum number of sensors. Since we consider transmissions and receptions of messages, our results show a different trend. In Fig. 5.3, we also observe that, when the overhearing ratio (α) increases, the number of sensors deployed in the virtual node B_{i*} increases too. Indeed, the traffic is highest in this virtual node and a high value of α results in a high number of receptions.

In summary, when the overhearing ratio is not taken into account ($\alpha = 0$), the number of sensors per virtual node is almost a linear increasing function of the virtual node index. When the overhearing is taken into account ($\alpha > 0$), the number of sensors deployed per virtual node increases until a virtual node B_{i*} which receives the highest number of sensors, and then drastically decreases. The position of B_{i*} depends on p. If $p = 0, B_{i*} = B_{K-1}$ and when $p > 0, B_{i*} = B_{K-2}$.

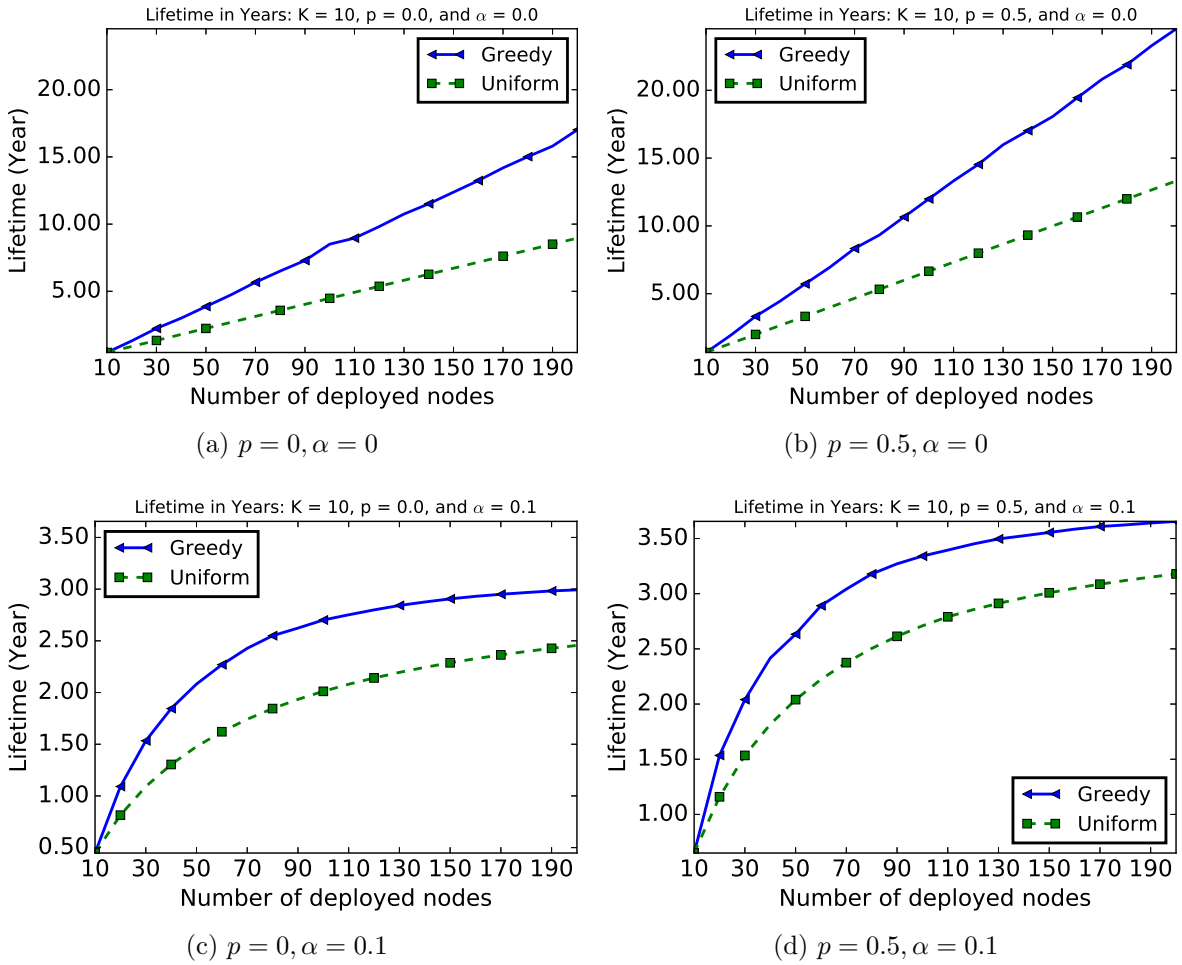
5.5.3 Lifetime: Comparison with uniform deployment

In a uniform deployment, the same number of sensors is deployed in all virtual nodes. We consider as our baseline the network lifetime $LT^{(uniform)}$ obtained with such a deployment. Since we assume that the traffic is balanced between sensors in a virtual node, the lifetime of a virtual node is equal to the average lifetime of sensors deployed in that virtual node. We denote by $E_{max} = \max\{E_i | i = 1, \cdots, K\}$ the maximum energy consumed by a sensor per time unit in the network. Recall that E_i is the energy consumption rate per sensor in the virtual node B_i (see Eq. 5.19). If E_{node} is the initial energy of a sensor, for a given deployment, the lifetime of the network will be equal to $\frac{E_{node}}{E_{max}} \cdot \Delta$.

5.5.3.1 The lifetime of the greedy and uniform deployment

Fig. 5.4 presents the lifetime of the greedy and the uniform deployment when $p = 0$ and $\alpha = 0$ (Fig. 5.4a), $p = 0.5$ and $\alpha = 0$ (Fig. 5.4b), $p = 0$ and $\alpha = 0.1$ (Fig. 5.4c) and $p = 0.5$ and $\alpha = 0.1$ (Fig. 5.4d). The results presented in this figure show that the greedy deployment, in terms of lifetime, outperforms the uniform one. Indeed, in the greedy deployment, more sensors are deployed in virtual nodes with higher load while, in the uniform deployment, the same number of sensors are deployed in all virtual nodes, no matter their relative position to the sink.

When overhearing is not taken into account (Fig. 5.4a and 5.4b), we observe a significant increase in lifetime (more than 20 years) compared to a network with an overhearing ratio of 0.1 (about 3.5 years). This means that, if we are able to use a perfect activity scheduling which



(a) $p = 0, \alpha = 0$

(b) $p = 0.5, \alpha = 0$

(c) $p = 0, \alpha = 0.1$

(d) $p = 0.5, \alpha = 0.1$

Figure 5.4: Impact of the number of nodes on the lifetime for different values of α and p.

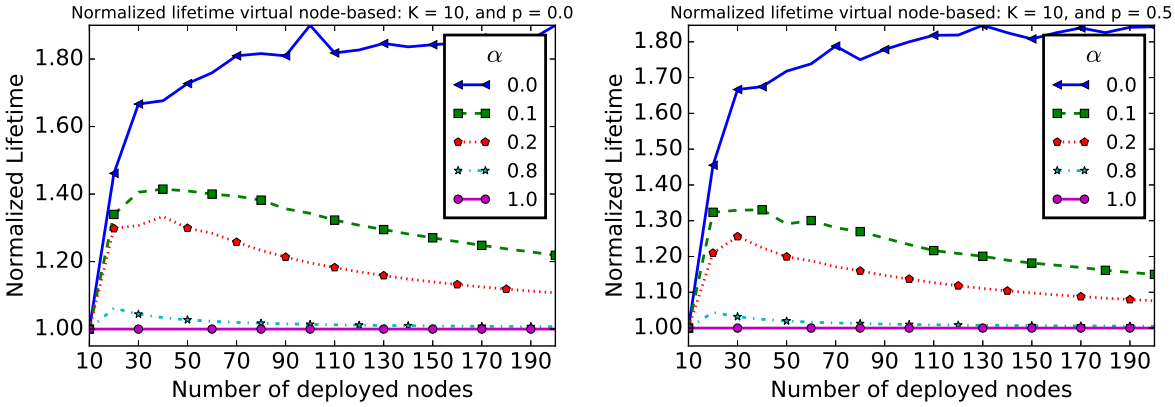
reduces the overhearing, the network lifetime can increase significantly. Of course, this perfect scheduling would most likely require some supplementary control traffic which is not considered in our study.

In Fig. 5.4, the lifetime increases with the number of sensors. Increasing the total number of sensors in the network means to increase the number of sensors deployed per virtual node, both for the greedy and the uniform deployment. A direct consequence of increasing the number of sensors per virtual node is the reduction of the average number of transmissions per sensor, and thus prolonging the network lifetime. When the value of p increases, the lifetime increases too. Indeed, since we assume a shortest-path routing, a high value of p means a high probability for sensors in virtual node B_i to relay traffic from B_{i-2}. Thus, by increasing the value of p, the amount of traffic relayed by B_i from B_{i-1} is reduced (See Eq. 5.1).

When $\alpha = 0$, the average number of operations per sensor per time unit tends to 0 (see Eq. 5.15) when the number of sensors deployed in the virtual node is very large. That is why, in Fig. 5.4a-5.4b, the lifetime is a strictly increasing linear function of the number of sensors. On the other hand, when overhearing is taken into account ($\alpha > 0$), the number of operations per node per time unit Δ in the virtual node B_i when the number of sensors is very large tends to $\alpha \cdot (2 \cdot T_i - \lambda + p \cdot T_{i-1} + T_{i+1} + p \cdot T_{i+2})$. Thus, when this limit is reached, it is useless to deploy

more sensors. That is why, in Fig. 5.4c-5.4d, we observe a threshold which is asymptotically reached starting from a given number of sensors. We can also notice in this figure that the gap between the lifetime of the greedy and the uniform deployment is reduced when more sensors are deployed. Indeed, the greedy deployment rapidly reaches the maximum lifetime compared to the uniform deployment.

5.5.3.2 The lifetime gain of the greedy deployment



(a) $p = 0.0$ (b) $p = 0.5$

Figure 5.5: Normalized lifetime, impact of the number of sensors

To illustrate how much the network lifetime could be extended when using the greedy Algorithm 5.1 rather than a simple uniform deployment, we normalize the lifetime of the greedy deployment by that of the uniform one. Thus, we use as metric the ratio r expressed by Eq. 5.20.

$$r = \frac{LT^{(greedy)}}{LT^{(uniform)}} \tag{5.20}$$

Fig. 5.5 shows, for different values of α, the normalized lifetime of the network as a function of the number of sensors when the number of virtual nodes (K) is equal to 10. We consider the cases $p = 0$ and $p = 0.5$ in Fig 5.5a and 5.5b, respectively.

The normalized lifetime decreases when α increases. When there is no overhearing ($\alpha = 0$), the normalized lifetime increases continuously and when $N \geq 70$, we observe an improvement of the lifetime greater than 80% compared to a uniform deployment. When there is no scheduling ($\alpha = 1$), i.e. all sensors always keep their radio on, the solution obtained with the greedy algorithm is equivalent to that given by a uniform deployment. This observation is consistent with our model and hypothesis, since Eq. 5.6, which expresses the average number of operations per node per time unit in a virtual node, shows that, when $\alpha = 1$, $\overline{O}_i$ is independent of the number of sensors deployed in the virtual node B_i.

When $0 < \alpha < 1$, we observe, in Fig. 5.5, a significant decrease of the performance compared to the case $\alpha = 0$. Nevertheless, the lifetime of the uniform deployment can be improved by up to 40% (Fig. 5.5a, $p = 0$) or 30% (Fig. 5.5b, $p = 0.5$), depending on the overhearing ratio.

As expected, increasing the overhearing ratio (α) also increases the number of receptions and internal interferences in the network, and then reduces the lifetime of the network. We highlight that, from a given number of sensors, the performance of the greedy deployment compared to the uniform deployment starts decreasing. Indeed, as presented in the Fig. 5.4, from a given number of sensors, our greedy deployment reaches its maximum lifetime, while the lifetime of the uniform deployment still increases, and thus, the gap between the greedy and the uniform deployment is reduced.

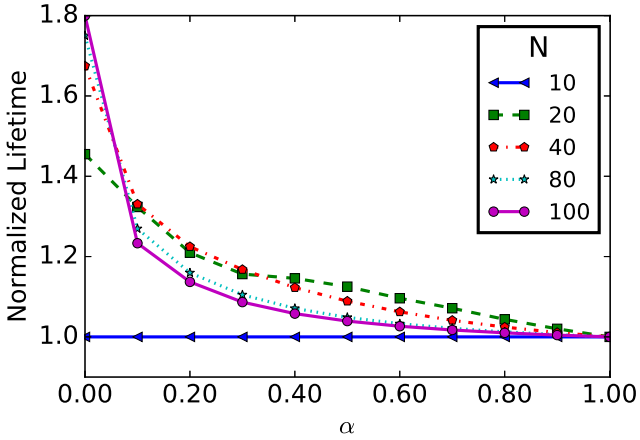


Figure 5.6: Impact of α: $K = 10$ and $p = 0.5$

Fig. 5.6 presents the normalized network lifetime as a function of α and the number of deployed sensors. When the overhearing phenomenon is taken into account (i.e. $\alpha > 0$), the gain (compared a deployment with $\alpha = 0$) is reduced. Thus, when designing a deployment strategy or when evaluating the performance of communication protocols, it is important to take into account all messages that can be received by sensors. For values of α close to 1, the greedy deployment is equivalent (or close) to a uniform deployment in terms of network lifetime.

5.5.3.3 Impact of the network length

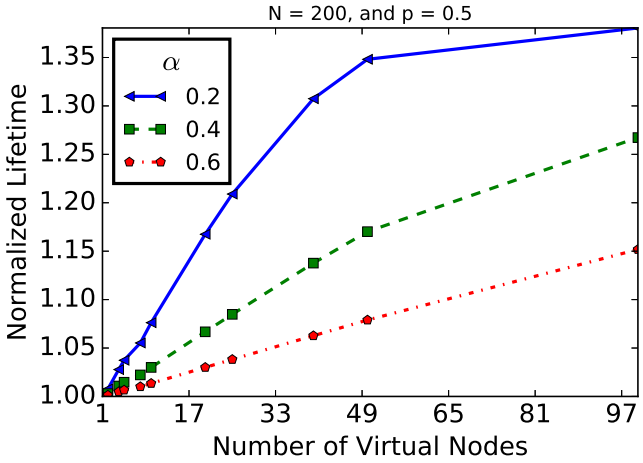
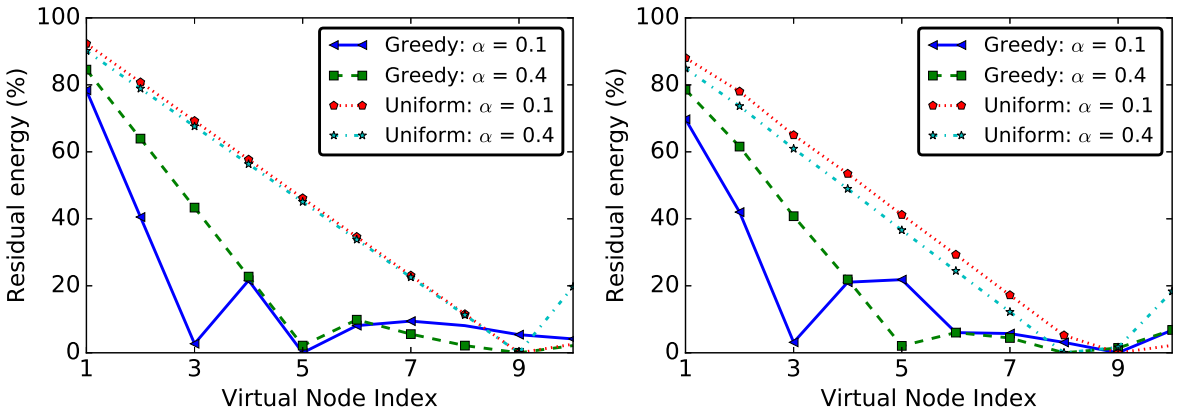


Figure 5.7: Impact of the network length (K) for a network of $N = 200$ sensors.

Fig. 5.7 shows the impact of the network length, in terms of the number of virtual nodes. For a given number of sensors ($N = 200$), this figure shows the impact of the network length (K) and the overhearing ratio (α) on the deployment proposed in this paper. It is important to note in this figure that we consider only values of K less than N and multiple of N ($K < N$). For $K = N$, we will have 1 sensor per virtual node and a normalized lifetime of 1. Results presented in this figure show that the gap between the greedy deployment and the uniform one increases with K ($K < N$). Indeed, by increasing the K value, the overall network traffic in the network increases too, and particularly in virtual nodes close to the sink. However, considering a uniform deployment, increasing the network length while keeping a constant number of sensors (as it is the case here) reduces the number of sensors to form each virtual node. A direct consequence is a negative impact on the network lifetime of the uniform deployment. It is not the case for the greedy deployment, since this algorithm deploys sensors in virtual nodes taking into account network traffic in each area.

5.5.4 Residual energy



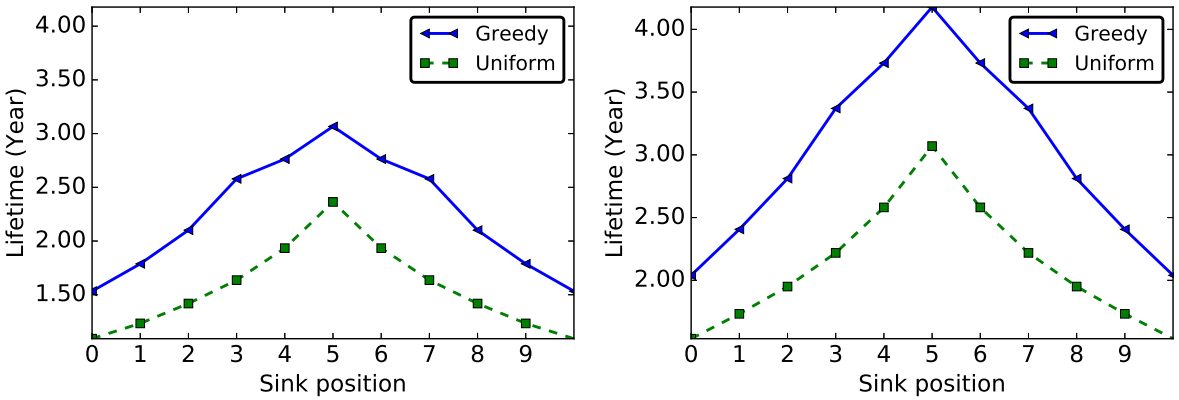
(a) $p = 0.0$ (b) $p = 0.5$

Figure 5.8: Residual energy per sensor as function of the virtual node: $K = 10, N = 30$.

Considering a LWSN organized in $K = 10$ virtual nodes, we present in this section the residual energy of sensors per virtual nodes at the end of the network functioning. In Fig. 5.8, results are given for $N = 30, p = 0$ (Fig. 5.8a), $p = 0.5$ (Fig. 5.8b) and for different values of α. As in the previous results, note in this figure that the sink is deployed at position $K + 1$ (i.e. virtual node index 11). An important observation concerning the results presented in this figure is the highest residual energy of sensors for the uniform deployment. While the same number of sensors are deployed in each virtual nodes with the uniform deployment, our greedy deployment deploys more sensors in the area close to the sink. Thus, for virtual nodes far from the sink, the traffic per sensor is lower for the uniform deployment, since it deployed more sensors compared to the greedy deployment proposed in this chapter. This results in lower energy consumption for these sensors in the case of the uniform deployment compared to the greedy one.

5.5.5 Impact of the sink position

So far, we assumed that the sink is deployed at position $K+1$. In this section, we evaluate the impact of the sink position on the network lifetime considering the uniform deployment as well as the greedy one proposed in this paper. We vary the sink position from 0 to $K+1$.



(a) $p = 0.0$ (b) $p = 0.5$

Figure 5.9: Lifetime of the uniform and greedy deployments as function of the sink position: $K = 10, \alpha = 0.1, N = 30$.

Figure 5.9 shows the lifetime of the network considering the uniform and greedy deployments as function of the sink position. We consider a network formed of 10 virtual nodes, 30 sensors and $\alpha = 0.1$. Results are shown for $p = 0$ (Fig. 5.9a) and $p = 0.5$ (Fig. 5.9b). From the network traffic point a view, the sink deployed at position 0 or $K+1$ corresponds to the configuration considered in the previous section and gives the same network lifetime. But, by deploying the sink at a different position than the network edges, the among of traffic forwarded by the closest neighbors of the sink is reduced. And, as shown in Fig. 5.9, the optimal position is the middle of the network.

5.6 Comparison to related work

In this section, we compare the deployment strategy proposed in this thesis to the one proposed in [102]. In the following, the solution proposed in this thesis will be denoted as virtual node-based and the one proposed in [102] as distance-based. Since these two approaches are quite different, it is not easy to provide a fair comparison. In the following section, we start by describing the solution proposed by [102]. We assume in this section that the sink is in the virtual node B_{K+1}.

5.6.1 Distance-based deployment

In [102], the deployment is formulated as a Mixed-Integer Linear Programming (MILP) problem. In their formulation, the authors denote by d_i the distance between sensors S_i and S_{i+1}, and R_j the communication range when the transmission power P_j is used. For each sensor S_i, m binary

variables $(x_{i1}, x_{i2}, \cdots, x_{ij}, \cdots, x_{im})$ are defined to denote the assignment of power level for S_i. Thus :

$$\sum_{j=1}^{m} x_{ij} = 1 \tag{5.21}$$

$$d_i = \sum_{j=1}^{m} x_{ij} \cdot R_j \tag{5.22}$$

The authors assume that S_i relays all the messages from S_{i-1}. If μ is the number of messages generated by a sensor per time unit, the number of operations (transmissions only, since in [102] receptions are not taken into account) per sensor per time unit is defined by Eq. 5.23.

$$\overline{O}_i^{(r)} = \mu \cdot i \tag{5.23}$$

If we consider receptions and the overhearing phenomenon, as in this chapter, the number of operations per sensor per time unit is now equal to:

$$\overline{O}_i^{(r)} = \mu \cdot (i - 1 + i + \alpha \cdot (i + 1)) = \mu \cdot (2 \cdot i + \alpha \cdot (i + 1) - 1) \tag{5.24}$$

The energy consumed by S_i per time slot is then equal to:

$$E_i = \overline{O}_i^{(r)} \cdot \tau \cdot \left(\sum_{j=1}^{m} x_{ij} \cdot P_j \right) \tag{5.25}$$

The objective function in [102] is given by:

$$minimize \ \ max\{E_i | i = 1, \cdots, N\} \tag{5.26}$$

subject to the following constraint:

$$\sum_{i=1}^{N} x_{ij} \cdot d_i \geq L \tag{5.27}$$

where L is the initial network length, a parameter defined in [102]. We use the CPLEX software package to find the optimal solution to this problem. It is important to note here that, in the solution provided by the solver, the total area covered by sensors is larger than L (see Eq. 5.27). In the following, we denote by $L_{simulated}$ the (final) length of the network provided by the solver.

In our work, we form a network of virtual nodes and we propose a greedy algorithm which calculates the number of sensors to deploy in each virtual node. The MILP problem proposed by [102] and described in this section (Eq. 5.26-5.27) considers a flat network and expresses the deployment results in terms of distance between consecutive nodes.

Parameter	Distance-based	Virtual node-based
Communication range	R_{min} to R_{max}	R_{max}
Number of virtual nodes	–	$K = \lceil \frac{L_{simulate}}{R_{max}} \rceil$
Messages generation rate	μ per node	$\mu \cdot N \cdot \frac{R_{max}}{L_{simulated}}$ per virtual node

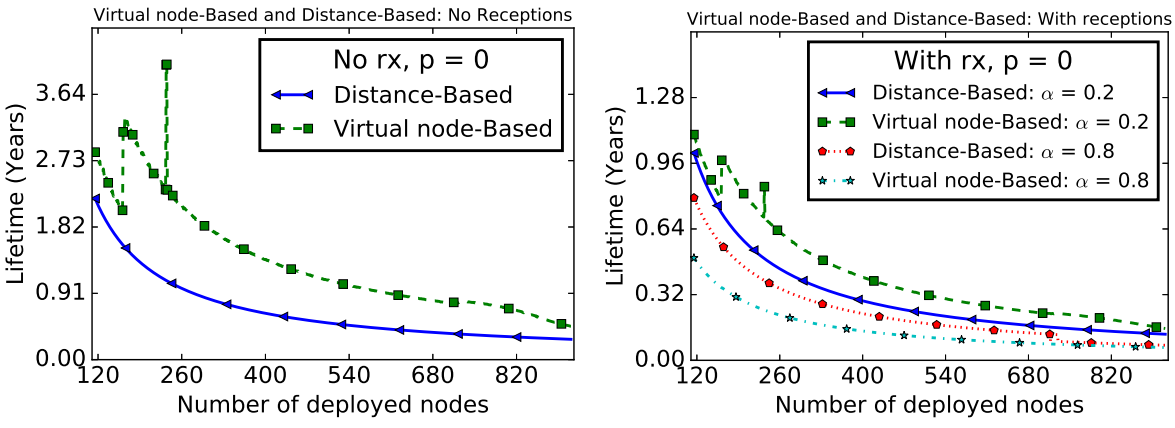
Table 5.3: Virtual node-based and distance-based comparison parameters.

5.6.2 Comparison framework

To provide a fair comparison, we use the same parameter values as the ones used in [102] to solve the problem described by Eq. 5.26 and 5.27. In our virtual node-based approach, we assume as the average transmission power P_t, the maximum power level used in the distance-based approach. Thus, the area covered by a virtual node is of length R_{max}, the transmission range when the maximum transmission power is used. Since [102] assumes that S_i relays all traffic from S_{i-1}, we set $p = 0$, i.e communications are not allowed between virtual nodes B_i and B_{i+2}.

In the virtual node-based and the distance-based approaches, traffic models are different. In our evaluation, we guarantee the same average messages generation per sensor in both approaches. If, for the distance-based approach, μ is the average number of messages generated per sensor per time slot, the total number of messages generated in the network of length $L_{simulate}$ per time slot is $\mu \cdot N$, where N is the total number of sensors. Thus, in our virtual node-based approach, we assume that λ, the average number of messages generated per virtual node per time slot is equal to $\mu \cdot N \cdot \frac{R_{max}}{L_{simulated}}$. Comparison parameters are summarized in Table 5.3.

5.6.3 Comparison results



(a) No Receptions

(b) With Receptions

Figure 5.10: Network lifetime of Distance-Based and Virtual node-Based deployment: $L = 5$ km.

Fig. 5.10 presents the lifetime of the network for distance-based and virtual node-based deployments. In Fig. 5.10a, only transmissions are considered, while Fig. 5.10b also takes into account receptions with and overhearing ratio of 0.2 and 0.8. In the evaluations, we assume

that $\Delta = 1$ minute and $\mu = 1/\Delta$. As it might be expected, considering messages reception (Fig. 5.10b), the lifetime is divided by more than two compared to the model in which only transmissions are taken into account (Fig. 5.10a). When receptions are not considered (Fig 5.10a), the virtual node-based approach gives better performance compared to the distance-based one. This advantage of the virtual node-based approach is also noticed for low and medium overhearing ($\alpha = 0.2$ in Fig. 5.10b). In the virtual node-based approach, all sensors use the maximum power level. Thus, sensors are more exposed to the overhearing phenomenon. That is why, in Fig 5.10b, the distance-based approach gives better performance when $\alpha = 0.8$.

When the number of sensors in the network increases, the lifetime of the two approaches decreases. Indeed, deploying more sensors in the network has an impact on the two approaches. For the distance-based solution, this increases the traffic in the network (since each node generates messages) and the network length. When the network length increases, the number of virtual nodes increases too. Moreover, increasing the sensors density allows more sensors in the distance-based approach to use lower power level. That is why we observe a reduction of the gap between the two approaches in Fig. 5.10.

In Fig. 5.10, we observe a significant increase of the performance of our virtual node-based approach at some particular values of N. This is an artifact of the discrete power level model used in [102], where 6 power levels are possible. By increasing the number of sensors, smaller power levels are used with the distance-based approach. Thus, when a sufficient number of sensors is deployed, all sensors use the next lowest power level [102]. At these points, the resulting network length $L_{simulated}$ is almost equal to L, which reduces the number of virtual nodes and favors our solution.

5.7 Investigating the case of heterogeneous sensors

In the previous sections, we assumed we have a large number of homogeneous sensors. We divided the network into virtual nodes and then we proposed a greedy algorithm to compute the number of sensors to deploy in each virtual node in order to extend the network lifetime. In this section, the network is still divided into virtual nodes, but only one sensor is deployed into each virtual node. We assume that sensors are heterogeneous in terms of battery capacity, and our goal is to find the battery capacity to assign to a given virtual node, taking into account its traffic, in order to extend the network lifetime.

5.7.1 Problem formulation

We assume that we have an energy budget E_{net}, and our goal is to distribute this energy to all sensors, in order to maximize the network lifetime. We assume in this section that one sensor is deployed in each virtual node (i.e. $K = N$). If we replace $n_i = 1$ in Eq. 5.15, we will have a new model of the average number of operations per sensor per time unit, defined by the Eq. 5.28:

$$\overline{O}_i = (2 \cdot T_i - \lambda) + \alpha \cdot \left[p \cdot T_{i-1} + T_{i+1} + p \cdot T_{i+2} \right] \tag{5.28}$$

We denote by ξ_i the energy assigned to virtual node B_i. Since E_i (Eq.5.19) is the energy consumption rate of the virtual node B_i, the lifetime of this virtual node, $LT_i^{(energy)}$, is defined by Eq. 5.29:

$$LT_i^{(energy)} = \frac{\xi_i}{E_i} \tag{5.29}$$

If we define the network lifetime (LT) as the time until the first virtual node exhausts its energy, it will be defined by Eq. 5.30:

$$LT^{(energy)} = min\ \{LT_i^{(energy)} | i = 1, \cdots, K\} \tag{5.30}$$

Our goal is to find $\xi_{i,1\leq i\leq K}$ which maximizes $LT^{(energy)}$. Formally, the problem can be presented as:

$$maximize\ LT^{(energy)} \tag{5.31}$$

subject to the following constraints:

$$\sum_{i=1}^{i=K} \xi_i = E_{net} \tag{5.32}$$

$$\xi_i > 0 \tag{5.33}$$

Eq. 5.31-5.33 define a linear optimization problem. We find the optimal solution to this problem using the CPLEX software package.

Intuitively, the energy budget E_{net} could be distributed among virtual nodes proportionally to their traffic. Thus, alternatively, we propose Eq. 5.34 to model the energy assigned to virtual node B_i.

$$\xi_i = \frac{\overline{O}_i}{\sum_{j=1}^{j=K} \overline{O}_j} \cdot E_{net} \tag{5.34}$$

In Eq. 5.34, $\overline{O}_i$ is the average number of operations per sensor per time unit in the virtual node B_i. Fig. 5.11 presents the optimal solution (provided by CPLEX software) to the problem described by Eq. 5.31-5.33 and the solution given by Eq. 5.34. This figure shows that the two solutions are almost identical.

Recall that $E_i = \overline{O}_i \cdot \tau \cdot P$. Thus, $LT_i^{(energy)}$ is defined by Eq. 5.35:

$$LT_i^{(energy)} = \frac{\xi_i}{E_i} = \frac{\overline{O}_i}{\sum_{j=1}^{j=K} \overline{O}_j} \cdot E_{net} \cdot \frac{1}{\overline{O}_i \cdot \tau \cdot P} = \frac{E_{net}}{(\sum_{j=1}^{j=K} \overline{O}_j) \cdot \tau \cdot P} \tag{5.35}$$

Note in Eq. 5.35 that all sensors have the same lifetime and thus, will exhaust their battery at the same moment.

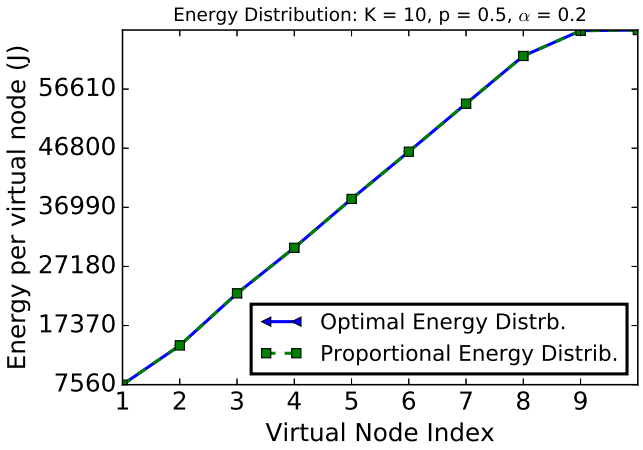


Figure 5.11: Optimal and proportional energy distribution per virtual node for an energy budget $E_{net} = 408240$ J.

5.7.2 Uniform vs proportional energy distribution

For a uniform energy distribution, the same amount of energy is assigned to each virtual node. We define $\xi_i^{(u)}$, the energy assigned to each node, given by Eq. 5.36.

$$\xi_i^{(u)} = \frac{E_{net}}{K} \tag{5.36}$$

And $LT_i^{(uniform)}$, the lifetime of the virtual node B_i for a uniform energy distribution, is defined by Eq. 5.37.

$$LT_i^{(uniform)} = \frac{\xi_i^{(u)}}{E_i} = \frac{E_{net}}{K} \cdot \frac{1}{\overline{O}_i \cdot \tau \cdot P} = \frac{E_{net}}{K \cdot \overline{O}_i \cdot \tau \cdot P} \tag{5.37}$$

If the lifetime of the optimal energy distribution is normalized by the lifetime given by the uniform energy distribution, we obtain:

$$\frac{LT_i^{(energy)}}{LT_i^{(uniform)}} = \frac{E_{net}}{(\sum_{j=1}^{j=K} \overline{O}_j) \cdot \tau \cdot P} \cdot \frac{K \cdot \overline{O}_i \cdot \tau \cdot P}{E_{net}} = \frac{\overline{O}_i \cdot K}{\sum_{j=1}^{j=K} \overline{O}_j} \tag{5.38}$$

Thus, if different energy capacities can be assigned to sensors, the optimal capacity assignment (Eq. 5.34) can improve the lifetime of the virtual node B_i by a factor of $\frac{\overline{O}_i \cdot K}{\sum_{j=1}^{j=K} \overline{O}_j}$ compared to a uniform energy distribution. If we denote by $\overline{O}_{max} = \max\{\overline{O}_i | i = 1, \cdots, K\}$, the maximum number of operations per sensor per time unit in the network, we obtain: $\overline{O}_{max} \cdot K > \sum_{j=1}^{j=K} \overline{O}_j$ and $\frac{\overline{O}_{max} \cdot K}{\sum_{j=1}^{j=K} \overline{O}_j} > 1$.

5.7.3 Deployments based on heterogeneous and homogeneous sensors

Our goal here is to compare the deployment based on energy distribution solution proposed in this section to the one based on virtual nodes described by Algorithm 5.1. Recall that, given N sensors, this greedy algorithm finds the number of sensors to deploy in each virtual node. To allow a fair comparison, we assume that the energy unit is equal to E_{node}, and an energy budget

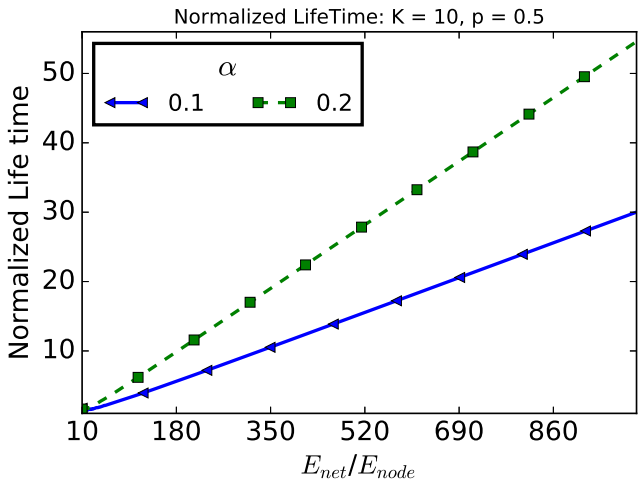


Figure 5.12: Energy distribution vs Virtual nodes-based: Impact of the energy budget when $K = 10, p = 0.5$.

of E_{net} can be discretized to have $N = \frac{E_{net}}{E_{node}}$ homogeneous sensors. Thus, given the energy budget, we can fairly compare the energy distribution deployment to the sensors distribution one by discretizing the energy budget and implementing the two solutions.

Fig. 5.12 presents $\frac{LT^{(energy)}}{LT^{(greedy)}}$ as a function of $\frac{E_{net}}{E_{node}}$ (recall that $LT^{(greedy)}$ is the network lifetime obtained by Algorithm 5.1). This figure shows that the energy distribution solution outperforms the sensor distribution one and the gap increases with the ratio $\frac{E_{net}}{E_{node}}$. It is also important to note that, by deploying a single sensor in each virtual node, the routing problem is simplified and there is no load balancing problem of the traffic per virtual node as in the case of the solution proposed in the previous sections. However, it may not be economically or practically feasible for industrials to produce sensors with different battery capacities. Moreover, a sensor failure might conduct to a network partition since a single sensor is deployed per virtual node.

5.8 Conclusion

In a multi-hop wireless network with a convergecast communication model, we observe a high traffic intensity in the area close to the sink. In WARIM, the WSN architecture proposed in this thesis for vehicular traffic monitoring at an intersection, the network formed by nodes deployed on each lane has this characteristic.

In this chapter, we studied the problem of sensors deployment to maximize the lifetime of WSN with linear topology. InWARIM, nodes deployed on a lane form a network with such a topology. To simplify the deployment, we divided the network into virtual nodes. By considering messages transmission and/or reception as energy operations consuming, we proposed an analytical model of the number of operations per virtual node. Unlike existing works, our model considers the message reception cost. It also takes into account messages received due to the overhearing phenomenon.

Considering homogeneous sensors in terms of battery capacity, we proposed a greedy algorithm to compute the number of sensors to deploy in each virtual node. Performance evaluation

shows that, depending on the number of sensors, the simple virtual node-based deployment proposed can improve the network lifetime by up to 40%, when compared to the uniform deployment. Moreover, it outperforms the distance-based approach proposed in the literature [102] when a scheduling algorithm which reduces the overhearing phenomenon is used.

Finally, we studied an alternative solution. We considered a network in which the batteries of sensors are nonuniform, i.e. the battery of a sensor is proportional to its traffic. Results show that, by properly selecting the battery to assign to each sensor, such an approach significantly improves the network lifetime.

Chapter 6

Sensor Ranking in Wireless Sensor Networks with Linear Topology

In WARIM, the new WSN architecture proposed in this thesis for vehicular traffic monitoring at an intersection, sensors deployed on the lane surface collect data concerning vehicular traffic at lane level. These sensors then relay the collected data in a multi-hop manner toward the traffic light controller. The vehicles queue length on each lane and their average waiting time are some parameters taken as input by the TLMS. The closeness of the values of these parameters to the ground truth may have an impact on the performance of the TLMS. To determine the values of these parameters, it is important to know the position of each sensor related to the light controller. Indeed, if each sensor inserts its position in each message it generates, the traffic light controller will be able to determine the values of these parameters. Moreover, knowing the sensors position at the intersection, the controller will also be able to detect unauthorized stop. Using such sensors location, a geographic or a gradient-based routing protocol can also be designed to collect efficiently vehicular traffic data.

To obtain the location of each sensor at the intersection, a simple solution is to equip each of them with a GPS. But this solution, not only increases the cost of WARIM, but is also a huge energy consumption source. In this chapter, considering sensors deployed on a lane to form a LWSN, our goal is to propose a mechanism allowing to get the position of each sensor relative to the lights controller. Since one of the goals of WARIM is to be as autonomous as possible, the mechanism proposed in this chapter must limit the human intervention, i.e. must function without important manually configured parameters.

In the remainder of this chapter, we start by motivating, describing and formulating the sensor ranking problem in LWSN in Section 6.1. We continue with a discussion of the related work in Section 6.2. In Section 6.3, we discuss the usage of a single anchor to rank sensors, an idea used in our iterative algorithm, presented in Section 6.4. Using a simulator developed during this thesis (see Appendix A), we evaluate, in Section 6.5, the performance of the ranking algorithm proposed in this chapter. This evaluation considers different channel models (an unit disk graph model in which some links are missing and a channel model built from empirical data) and sensors deployment strategies (uniform, random and the virtual nodes-based deployment proposed in Chapter 5). Finally, we conclude this chapter in Section 6.6.

6.1 Problem statement

6.1.1 Context

Sensor location is needed in LWSN not only for network operations (e.g. geographic routing and data collection mechanisms), but also at the application layer. For instance, in WARIM, when a vehicle sends a message, it is important to insert its location in the message in order to allow to the lights controller to identify the area where this message comes from. This allows to estimate the vehicle queue length and to localize incidents. Another example is when a problem is detected by one of the large set of sensors monitoring a 10 km long bridge, sensor location information is required for an intervention. While precise coordinates (in the range of cm) might be required in some scenarios, a relative *ranking* information is often enough for most LWSN applications. Considering a deployed LWSN, a relative sensor's rank or position (the terms rank and position are interchangeably used in this chapter) is defined as the number of sensors closer to the traffic light controller compared with that sensor. In this chapter, we are interested by a distributed algorithm for sensor ranking in LWSN. Indeed, manually assigning coordinates, or ranks, during the sensor deployment phase is an error prone, time consuming operation, also requiring qualified workforce. We define the sensors ranking as the process of finding the relative (to the sink) rank of each sensor.

Theoretically, the RSSI is a decreasing function of the distance, an assumption largely used in WSN, especially for sensor ranking in a LWSN [126, 127, 128, 131, 130, 132]. Nevertheless, multiple studies show that, because of the hardware quality or constraints related to the environment (e.g. obstacles and climate conditions), the RSSI by itself does not provide enough information for correct range estimation. Rather than using information provided by the hardware, [134] uses neighborhood information to rank sensors in a LWSN thanks to two anchors. Inspired by this work, we show in this thesis that it is possible to rank sensors in a LWSN knowing only the first sensor of the network as anchor, relaxing the two-anchors constraint. Based on this result, we propose an iterative algorithm for sensor ranking in a LWSN where, in each iteration, a virtual anchor is selected, allowing to correctly rank more and more sensors in the network.

6.1.2 Notations

We consider a network consisting of N sensors, $S_1, S_2, \ldots, S_N$, deployed along a line in this order, i.e S_i is the i$^{\text{th}}$ sensor of the network. We denote by $\mathcal{G} = (\mathcal{S}, \mathcal{L})$ the directed graph associated to the network. The vertices $\mathcal{S}$ of this graph correspond to the sensor and there is an edge $(S_i, S_j) \in \mathcal{L}$ whenever S_j can receive messages from S_i. We denote by $\Gamma(S_j)$ the one-hop neighbors of a sensor S_j, i.e all the sensors S_i such that $(S_i, S_j) \in \mathcal{L}$, while $h(S_i, S_j)$ is the number of hops between sensors S_i and S_j. We denote by $x_i \in \mathbb{R}$ the estimated one-dimensional coordinate of S_i. We assume that sensor S_1 is manually configured by an operator with the coordinate $x_1 = 0$. As in Chapter 5, S_N is the sensor closest to the light controller. Tab. 6.1 presents the list of important notations used in our algorithm.

Parameter	Description
N	The number of sensors
S_i	The sensor deployed at position i
$S_{(i)}$	The sensor deployed at position i from the point of view of the ranking algorithm
$\Gamma(S_i)$	The set of one-hop neighbors of S_i
$\mathcal{G}$	The directed graph associated to the network
$\mathcal{L}$	The vertices of $\mathcal{G}$
$\mathcal{S}$	The edges of $\mathcal{G}$
$h(S_i, S_j)$	The number of hops between S_i and S_j
x_i	The estimated one-dimensional coordinate of S_i
R	The sensor communication range
r	The ratio of correctly ranked pairs of sensors
C	The total number of pairs of sensors
C_{ok}	The number of well ranked pair of sensors

Table 6.1: Ranking model notations.

6.1.3 Problem

The inputs of our problem are the network connectivity graph $\mathcal{G}$ and S_1. Given these inputs and a sensor $S_i, i \in [2, N]$, we define the rank (or the position) of S_i as the number of sensors deployed between S_1 and S_i, including S_1 and S_i. The question is whether we can correctly rank each sensor in the network with minimum pre-configured parameters. Therefore, the output of the problem is a sequence of sensors, $S_{(2)}, \ldots, S_{(i)}, \ldots, S_{(N)}$. $S_{(i)}, i \in [2, N]$ being, from the point of view of the ranking algorithm, the i$^{\text{th}}$ sensor of the network. An ideal algorithm for this problem would produce a sensor sequence which corresponds to the ground-truth, i.e. $S_{(2)} = S_2, \ldots, S_{(i)} = S_i, \ldots, S_{(N)} = S_N$

6.1.4 Evaluation metric

In general, the evaluation metric for localization algorithms is the localization error, i.e. the distance between the real coordinates and the estimated ones. However, our focus is on sensor ranking, a less precise, but usually sufficient localization information. For performance evaluation, we use the ratio of correctly ranked pairs of sensors. For any pair of sensors (S_i, S_j), the sensors are deployed in a given order in the network: either S_i is located (relatively to the light controller) before S_j, in which case $i > j$, or the opposite. In the output sequence, the pair (S_i, S_j) is well ranked if the order of the two sensors is the same as in the real deployment. Therefore, the ratio of correctly ranked pairs of sensors, r, is defined by Eq. 6.1:

$$r = \frac{C_{ok}}{C} \tag{6.1}$$

In Eq. 6.1, C_{ok} is the number of well ranked pair of sensors and C is the total number of pairs of sensors in the network. A high value of r means an output sequence of good quality, with $r = 1$ for a perfect order. Because of the unpredictable and unstable nature of links in WSN, two

problems can appear: the ranking of two sensors can be reversed, or a sensor might be entirely absent from the output sequence. The proposed metric is interesting since it takes into account these two problems.

6.2 Related work

Sensor localization is a widely studied subject in the field of WSN. Most localization solutions use ranging techniques based on measurements provided by the hardware. As measured metrics, we can have the angle of arrival (AoA), the time difference of arrival (TDOA), the time of flight (ToF) and the RSSI [121] as well as the phase difference between transmitter and receiver [122]. A second class of localization approaches uses network topology [124] or neighborhood correlation [125] information to estimate the distance between sensors.

6.2.1 RSSI-based solution

Regarding sensor ranking in LWSN, most of the existing solutions make the assumption that the strength of a radio signal decreases with the distance between the two communicating sensors. This implies that the closest neighbor should always receives messages with the highest RSSI value. Theoretically, as shown in Chapter 3 (see Figure 3.1) and in the related work [84], this assumption is true.

In all the RSSI-based solutions, each sensor ranks its neighbors in the decreasing order of the measured RSSI. The basic algorithm is, starting from an initial sensor, to build a sequence containing all sensors. In [126, 127], the number of sensors are known and the authors assume that no information is available concerning the first and the last sensor in the network. Thus in [126], a centralized solution is proposed in which many permutations are checked. A permutation is valid only if, starting from the first or the last sensor of the permutation, the same sequence is obtained, but in the reverse order. The worst-case running time of this algorithm is $\Omega(n!)$, making it impractical even in medium-sized sensor networks. But in [127], some mechanisms are integrated in the algorithm in order to reduce the number of permutations checked. In [128], the first sensor is known and authors also propose mechanisms to handle exceptions. An exception occurs when the output sequence does not contain all the sensors.

These simple RSSI-based approaches do not work very well in practice, where the radio signal between two sensors propagates over multiple paths, either constructive or destructive. Therefore, the idea of frequency diversity is exploited in [132], where sensors transmit and measure the RSSI on 16 different frequency channels. The authors propose to build a probability tree using the maximum RSSI measured on all these frequencies, the sensor ranking being given by the path with the maximum joint probability. However, the evaluation of this solution on a testbed, during a working day, shows that surrounding WiFi networks have a significant negative impact on its performance.

6.2.2 Neighborhood-based solution

Instead of using a metric provided by the hardware and based on the physical layer, a neighborhood-based solution is proposed in [134]. In this solution, the position of a sensor is estimated as the centroid relative to its neighbors, the idea being to bring physically connected neighbor sensors closer to one another in terms of virtual coordinates. The same idea, illustrated in Figure 6.1, is also exploited in [136], for the case of 2-dimensional networks. Assuming a simple 1-dimensional topology with only five sensors, as in Figure 6.1, the authors consider that the positions and the coordinates of sensors S_1 and S_5 are known: $x_1 = 0$ and $x_5 = d, d > 0 (N = 5)$. The coordinates of sensors S_2, S_3 and S_4 are initially unknown, and set to 0. Sensors begin by discovering their neighbors (and their respective coordinates), building a connectivity graph, as in Figure 6.1a. Once the neighbors are known, each of the sensors S_2, S_3 and S_4 sets its coordinate to the average of its neighbors coordinates. We thus have $x_2 = 0, x_3 = d/4$ and $x_4 = d/3$, as shown in Figure 6.1b. These new coordinates are announced to the neighbors and, after a second computation, we obtain $x_1 < x_2 < x_3 < x_4 < x_5$, as shown in Figure 6.1c, which gives the correct rank of sensors in the network. In [136], sensors initialize their coordinates with random values instead of 0. Using fixed or random initial coordinates does not affect the ranking performance of the centroid algorithm, but only its convergence time. As explained, in [134], sensors S_1 (the first

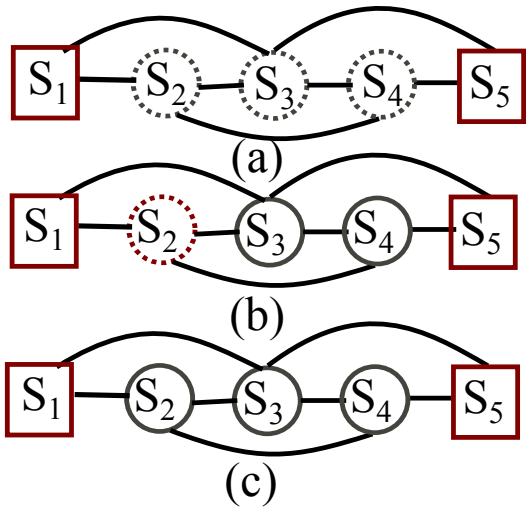


Figure 6.1: The centroid-based algorithm proposed in [134]: (a) After neighbors discovery: the coordinates of sensors $S_2 - S_4$ are unknown (b) Sensors S_3 and S_4 calculate their coordinates (c) Sensor S_2 calculates its coordinate.

sensor) and S_N (the last sensor) of the network have well-known coordinates and they are used as anchors. Moreover, all sensors are synchronized by a global clock, used to compute the beginning of each iteration. In each iteration, all the sensors broadcast their coordinates to their neighbors. The two anchors check whether all their neighbors have non-null coordinates and, when this is true, they broadcast a STOP message, which will be forwarded over the whole network in the following iterations. When a sensor S_i receives a STOP message from both anchors, it stops

updating its coordinate. The coordinate of sensor S_i at iteration t, is given by:

$$x_i^t = \frac{\sum_{S_k \in \Gamma(S_i)} x_k^{t-1}}{|\Gamma(S_i)|} \tag{6.2}$$

Considering a unit disk graph (UDG) model for sensor connectivity, [134] proves that, after $h(S_1, S_N)+1$ iterations, their solution produces the correct ranking. Two important observations can be made. First of all, this solution only produces the correct ranking; converging to stable values for the virtual coordinates might require extra iterations. Second, this result is only true for LWSN having a network diameter greater than one. Indeed, if the network diameter is equal to one, all sensors will have the same neighborhood under the UDG assumption, and the algorithm will produce the same coordinate for all sensors, as we show in the next section.

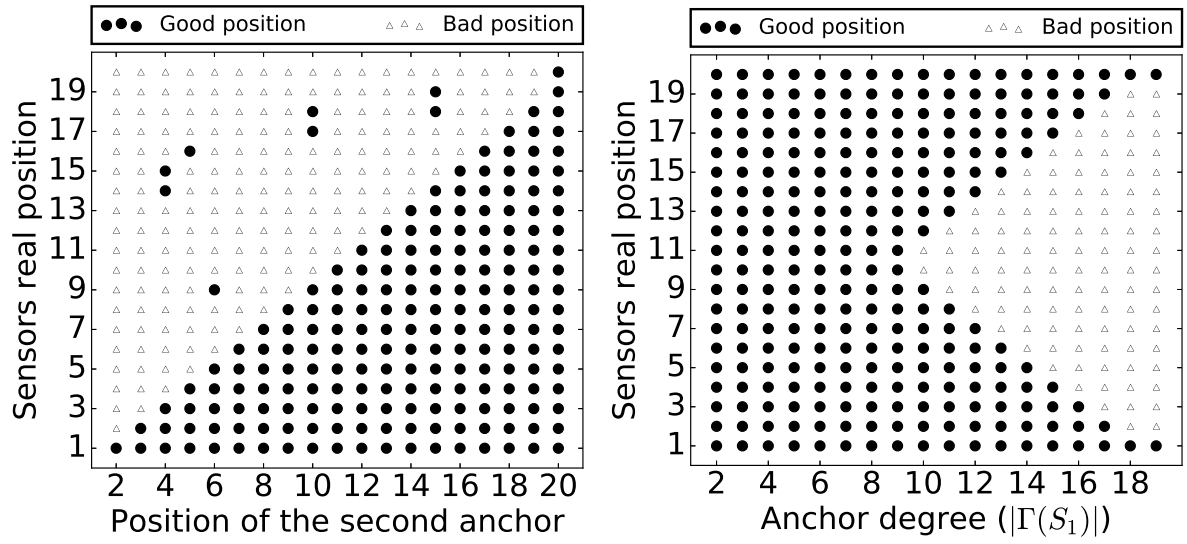
6.3 From two to one anchor solution

We are interested in an energy-efficient, lightweight and self-configured protocol for sensor ranking in a LWSN, where the number of parameters manually configured during network deployment is low. Therefore, we begin by investigating the consequences of an incorrectly configured second anchor in the solution proposed by [134]. We consider S_1 and S_k, $k \leq N$, the two anchors. However, we slightly modify the stopping condition, as described by Algorithm 6.1. In this modified algorithm, an anchor sends a STOP message when the coordinates of its neighbors converge to a stable value. As the coordinates of all non-anchor sensors are initialized to 0, it is easy to prove, by induction on t, that $\forall j, 1 < j < k, x_j^{t+1} \geq x_j^t$ and $x_j^t < x_k$, where S_k is the second anchor. This condition is sufficient to guarantee that this algorithm will converge and the anchors will send the STOP messages. Using a dedicated linear network simulator

Algorithm 6.1 Modified ranking algorithm for anchor S_a

1: Transmit the S_a coordinate (x_a)
2: **if** For All $S_i \in \Gamma(S_a), x_i^t - x_i^{t-1} < \epsilon$ **then**
3: Transmit $STOP_a$
4: **end if**

detailed in Appendix A, we evaluate the performance of the centroid-based approach when the second anchor is incorrectly defined and under varying communication range conditions. In a first step, we consider an UDG model, i.e two sensors are able to communicate if the distance between them is less than the communication range R. With this model, links between sensors are then symmetric. We consider a network of 20 sensors linearly and uniformly deployed with a fixed distance of 5 m between two consecutive sensors. Thus, sensors are deployed in the order $S_1, S_2, \ldots, S_{20}$. To take into account the sensor degree, we vary R from 10 m to 95 m (or $|\Gamma(S_1)|$ from 2 to 19). This allows us to have a network diameter range from 10 to 1 hop(s). For all simulations, S_1 is considered as the first anchor. The results are shown in Figure 6.2. Figure 6.2a corresponds to a communication range of 20 m (i.e. $|\Gamma(S_1)| = 4$). In this configuration, sensors from S_2 to S_{20}, one after another, are setup as the second anchor. On this figure, the



(a) (b)

Figure 6.2: Sensor ordering in a LWSN. The y-axis corresponds to the real sensor positions and a black point means that the estimated rank of the sensor is correct, while an unfilled triangle means a wrong estimated rank. (a) Impact of the position of the second anchor ($|\Gamma(S_1)| = 4$ or $R = 20\ m$) (b) Impact of sensor degree (sensor S_{20} is configured as the second anchor).

x-axis corresponds to the position of the second anchor, e.g. position 3 means that sensor S_3 is configured as the second anchor. The main message highlighted by this figure is that, when the sensor at the position k (sensor S_k) is configured as the second anchor, the estimated ranks of sensors $S_2, S_3, \ldots, S_{k-1}$ are always correct, and when the second anchor is located at the opposite position of the sink, all the sensors positions are correctly estimated.

In Figure 6.2b, to investigate the impact of sensor degree, we vary $|\Gamma(S_1)|$ from 2 ($R = 10\ m$) to 19 ($R = 95\ m$) sensors, as indicated by the x-axis. In this configuration, S_1 and S_{20} are the two anchors. When $|\Gamma(S_1)|$ is less than 10 sensors, any two sensors have a different one-hop neighborhood, and the centroid-based approach gives perfect ranking results. By increasing $|\Gamma(S_1)|$ (i.e the communication range), sensors at the network center start having all the other sensors in their neighborhood. When $|\Gamma(S_1)| = 19$, the network becomes a clique and all the sensors have the same neighborhood. In such a situation, these sensors have the same coordinates, and then the same rank. These sensors are the ones with a wrong estimated rank in Figure 6.2b.

The results presented in this section show that, in the case of a uniform sensors deployment, considering an UDG connectivity model and a network of size $N \geq 2 \cdot |\Gamma(S_1)| + 1$, regardless the sensor S_k designated as the second anchor, the estimated ranks of sensors $S_2, S_3, \ldots, S_{k-1}$ are always correct. In the next section, we describe a solution in which the sensors in the whole network can be iteratively ranked, without a pre-configured (precise knowledge of the) second anchor. However, the results also indicate that the proposed solution only functions in networks with a diameter of at least 3 hops. Nevertheless, we argue that LWSN generally respect this constraint in practice, and we evaluate the performance of the proposed algorithm in realistic scenarios in Section 6.5.

6.4 Iterative sensor ranking algorithm

In our quest of reducing the number of manually configured parameters in the network, we notice that, even when the second anchor is not perfectly chosen, a part of the sensors are still correctly ranked. Based on this observation, we propose, in this section, an iterative algorithm for sensor ranking in a LWSN with the objective to rank correctly all the sensors. This algorithm runs in two steps: in the first step, each sensor discovers its neighborhood, while the second step, executed only by some well-selected sensors, finds the ranks of all the sensors in the network. S_1 is the first sensor to execute the second step of the ranking algorithm, following by its farthest 1-hop neighbor (determined by the ranking algorithm) and so on.

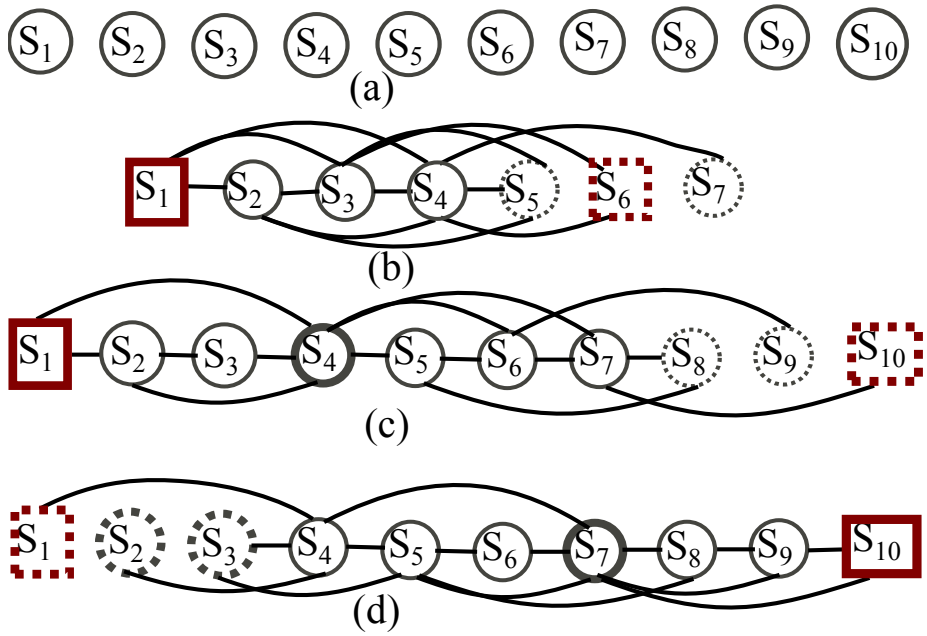


Figure 6.3: An example of a 10-sensors LWSN, $|\Gamma(S_1)| = 3$. (a) Sensor deployment; (b) Local view of the topology from S_1; (c) Local view of the topology from S_4; (d) Local view of the topology from S_7. Dotted lines are for two-hop neighbors and red rectangles are for anchor sensors.

6.4.1 Neighborhood discovery

In this first step of our solution, sensors discover their 1-hop and 2-hop neighbors. Basic *hello* messages can be broadcasted for 1-hop neighbor discovery. For 2-hop neighbors discovery, sensors exchange their list of one-hop neighbors. Considering a network of 10 sensors, uniformly deployed under an UDG channel model with $|\Gamma(S_1)| = 3$, as shown in Figure 6.3a, each sensor in the network builds a local view of the network. This local view is shown for S_1, S_4 and S_7 in Figures 6.3b, 6.3c and 6.3d, respectively.

Using this approach, each sensor in the network has a complete view of the connectivity between its 1-hop neighbors. Concerning the 2-hop neighbors, connectivity information is only partially available: the identity of the 2-hop neighbors is known, as well as their relations with one-hop neighbors; however, connectivity information between 2-hop neighbors is not known.

6.4.2 Node ranking

These local views of the topology, available at each sensor, are used for sensor ranking. As shown in Section 6.3, the sensors between two selected anchors are well ranked, as long as the distance between the two anchors is sufficient. Therefore, we propose to start with the first sensor, S_1, and rank the sensors in its 1-hop neighborhood. The two anchors are then iteratively modified, by selecting at each iteration sensors placed farther away, until ranks are computed for the entire network. Practically, S_1 chooses (based on a similarity metric defined in Section 6.5.2.2) one of its 2-hop neighbors as the second anchor, and it uses Algorithm 6.1 to rank all the sensors it has information on. Sensor S_1 then selects its farthest 1-hop neighbor to repeat this second step of our ranking algorithm, until reaching the end of the LWSN (all its 1-hop and 2-hop neighbors already have a rank). This iterative solution is detailed in Algorithm 6.2. This

Algorithm 6.2 Iterative algorithm

Input: $\mathcal{G} = (\mathcal{S}, \mathcal{L})$, S_1
Output: Sensor ordering

```
1:  S ← S_1
2:  stop ← 0
3:  while stop = 0 do
4:     G' ← Get-Local-Topology(G, S_i)
5:     S_frst ← S_j, a ranked node in G' with minimum rank
6:     if There are unranked two-hop neighbors then
7:        S_snd ← Select one as the second anchor
8:     else
9:        S_snd ← S_j ∈ Γ(S) with minimum one-hop neighbors
10:       stop ← 1
11:    end if
12:    Define-Anchors(G', S_frst, S_snd)
13:    Run-Centroid(G')
14:    Compute-Nodes-Rank(Γ(S_i))
15:    Transmit Γ(S)
16:    S ← Node-With-Max-Coordinate(Γ(S))
17:    if S does not have unranked neighbors then
18:       stop ← 1
19:    end if
20: end while
```

algorithm is executed by some selected sensors in the network, the sensor S_1 being the first one (Line 1). At each iteration, the currently selected sensor S builds a local view of the topology containing its 1-hop and 2-hop neighbors (Line 4). Examples of such topologies are presented in Figure 6.3. Sensor S then defines two sensors in its local topology as anchors (Lines 5 - 12), and it applies the centroid-based ranking algorithm on its local topology (Line 13). We note that, with the exception of the first iteration, sensor S is not one of the two anchors used in Algorithm 6.1. Running the algorithm on its local view of the topology, S updates the ranks of its 1-hop neighbors (Line 14) and shares this information with them (Line 15). Finally, after ranking all its 1-hop neighbors, S selects the one with the highest rank (i.e. the sensor situated

the farthest away) to execute the next iteration (Line 16). We note that, in an iteration, sensor S essentially does processing, sending only one message. This message allows the 1-hop neighbors of S to update their ranks, and to implicitly select which sensor will execute the next iteration. It is also important to note that the second step of our ranking algorithm ends with the sensor S which does not have neighbors (Lines 17-19).

The only question that remains to be answered regarding Algorithm 6.2 is the choice of the two anchors. At the beginning of the iteration, sensors in the 2-hop neighborhood of S can be divided into two groups: a group $\mathcal{R}$, containing sensors with ranks that are already defined (although not necessarily correct), and a second group $\mathcal{U}$, for sensors with undefined ranks. For the first anchor, S uses the sensor in $\mathcal{R}$ with the minimum rank, which allows updating the coordinates of all the 1-hop neighbors of S_i. At the beginning of the ranking algorithm, $\mathcal{R}$ is empty. In this situation, S_1 is setup as the first anchor. Regarding the second anchor, this sensor needs to be placed as far as possible from the first one. Therefore, a 2-hop neighbor in $\mathcal{U}$ is defined as the second anchor and, in Section 6.5, we propose a metric to guide this choice in order to obtain good ranking. Of course, when the ranking computation approaches the network edge, the choice of a two-hop neighbor might not be possible, and the set $\mathcal{U}$ might even be empty. In this situation, S selects, among its one-hop neighbors, the one with the minimum neighborhood size, which, in a uniform deployment, is the most likely to be the last sensor in the network.

6.4.3 Complexity analysis of the iterative ranking algorithm

In this section, we analyze the communication cost for each sensor. This cost is expressed in terms of the number of messages sent and received by a sensor. To simplify our analysis, we assume a UDG channel model and sensors uniformly deployed.

At the beginning of our ranking algorithm, each sensor discovers its 1-hop and 2-hop neighbors. A *hello* message, containing the *sender ID*, is used for 1-hop neighbors discovery. For 2-hop neighbors discovery, a particular message containing the *ID* list of the 1-hop neighbors is sent by each sensor. Thus, for this first step of our ranking algorithm, the average number of messages sent and received by each sensor is 2 and $2 \cdot |\Gamma(S_1)|$ respectively.

The second step of our ranking algorithm is ran by some selected sensors (named S in Algorithm 6.2), S_1 being the first one. In this step, the sensor S runs Algorithm 6.2. Thus, after executing the centroid Algorithm 6.1, it has to inform all its neighbors of their ranks, and also the next sensor selected to run the ranking algorithm. S can send this information in unicast or in broadcast mode. If the broadcast mode is chosen, S can send a message containing all its 1-hop neighbors and their ranks. This message will also allow to select the next sensor (denoted by S_0, the farthest 1-hop neighbor of S) to run the centroid algorithm. Thus, the number of messages received by each 1-hop neighbor is 2 (and 1 if the unicast mode where chosen). Indeed, the 1-hop neighbors on the right side of S will receive this message from S and also from S_0. Thus, in addition to the messages sent or received for 1-hop and 2-hop neighbors discovery, S experiences 1 transmission and all its neighbors 2 receptions.

We denote by $|\overline{\Gamma(S_1)}|$ the average number of neighbors per sensors. Thus, considering an individual sensor, the average communication cost (transmissions and receptions of messages) for neighborhood discovery and for the iterative ranking algorithm is $3 + 2 \cdot |\overline{\Gamma(S_1)}|$ for S and $4 + 2 \cdot |\overline{\Gamma(S_1)}|$ for other sensors. Moreover, the ranking algorithm is executed once just after the network deployment. Thus, we argue that it is worth paying the energy consumption cost induced by using this algorithm for sensors ranking in a LWSN. The overall communication cost for the ranking algorithm depends on the number of iterations of Algorithm 6.2, which depends on some network properties like the network diameter or the average sensor degree. These network properties depend, in turn, on the deployment scenario and environment. If we assume I iterations, the average number of transmissions and receptions in the whole network fro sensors ranking will be $(4 + 2 \cdot |\overline{\Gamma(S_1)}|) \cdot I$.

6.5 Performance evaluation

6.5.1 Evaluation overview

In this section, we evaluate the performance of the ranking algorithm proposed in this chapter. This algorithm is evaluated taking into account network and topological parameters. Concerning the topological parameters, we evaluate the ranking algorithm for different deployments: uniform, random and the virtual nodes-based deployment proposed in Chapter 5.

Considering the network parameters, so far we showed results for an UDG model, i.e. a network with stable, symmetric and uniform links. However, as it is shown in Chapter 3, in a real WSN deployment, the communication area of a sensor is never a perfect disk. We evaluate the performance of the proposed algorithm under more realistic network conditions. Thus, two channel models are considered in our evaluation. In the first model, we consider that a percentage of links are missing in the UDG model. The missing links are randomly selected. The second channel model is built from the empirical data collected during the experiments conducted during this thesis. These data correspond to the ones presented in Chapter 3. In this channel model, the link between two sensors is labeled by a weight representing the packet reception rate on that link. More details are given in the next sections concerning these two channel models. Nevertheless, we separately evaluate the performance of our ranking algorithm for the two channel models

The ranking algorithm is evaluated in terms of ratio of correctly ranked pairs of sensors (see Section 6.1 for more details) and ranking error. We define the ranking error as the shift of a sensor from its real position. We express the ranking error in terms of hops or/and distance, depending on the deployment. While for a uniform and random deployments we express the ranking error in number of hops, for the virtual nodes-based deployment proposed in Chapter 5, it is expressed in hops and distance. Expressing the ranking error in distance for the virtual nodes-based deployment also allows to understand the ranking error in terms of virtual sensors shift.

For each scenario, 1000 or 2000 topologies are generated. Thus, the ratio of correctly ranked pairs of sensors is an average of the results obtained for all the generated topologies. These results are shown with 95% confidence intervals.

6.5.2 The case of UDG-based channel model

In this section, our goal is to evaluate our ranking algorithm with only one fixed anchor considering a channel model built using the UDG assumption. We start by describing in the following section our evaluation methodology, and particularly the channel model.

6.5.2.1 Evaluation methodology

We consider networks with asymmetric links, i.e, between sensors S_i and S_j, we have two independent links $S_i \rightarrow S_j$ and $S_j \rightarrow S_i$. Sensors are linearly deployed (uniformly, randomly or using the virtual nodes-deployment proposed in Chapter 5) in a LWSN. We define a missing link as a link that exists in a UDG scenario, but not in the studied network. We evaluate the performance of our algorithm when a given percentage of links are missing. However, to ensure connectivity, we assume that the sensor deployment guarantees a stable and symmetric link between two consecutive sensors S_i and S_{i+1}. Such a condition can be satisfied during network deployment, by selecting the appropriate hardware or by adapting given parameters (i.e. transmission power) to the environment. Considering a given network size and a given missing link probability (the same for all the links, except direct physical neighbors, which are always connected), we simulate several topologies. For each topology, we apply our iterative algorithm, and finally we calculate the ratio of correctly ranked pairs of sensors, as described in Section 6.1.

6.5.2.2 Preliminary observations

For our first results, 2000 topologies are considered. For each topology, we consider 13 sensors uniformly deployed to form a LWSN, a distance of 5 m between consecutive sensors and $|\Gamma(S_1)| = 6$ (or $R = 30\ m$). Recall that $|\Gamma(S_1)|$ is the number of neighbors of S_1.

S_1 coverage area

The fact that we consider missing links creates new challenges compared with the theoretical UDG scenario. In the UDG case, the 1-hop neighborhood of sensor S_1 would be formed by sensors $S_2 - S_7$, the sensor S_7 being the farthest neighbor. However, this is no longer the case when missing links are allowed. Considering 2000 topologies, Figure 6.4 shows, for this scenario, the farthest 1-hop neighbor of the sensor S_1 as a function of the missing link probability. This Figure shows that, when some links are missing, the sensor coverage area, and implicitly the neighborhood size, evolve contrary to the assumption of the UDG model. A direct consequence is that, when the missing link probability increases, the number of iterations required by our sensor ranking algorithm increases, as less sensors are ranked in each iteration.

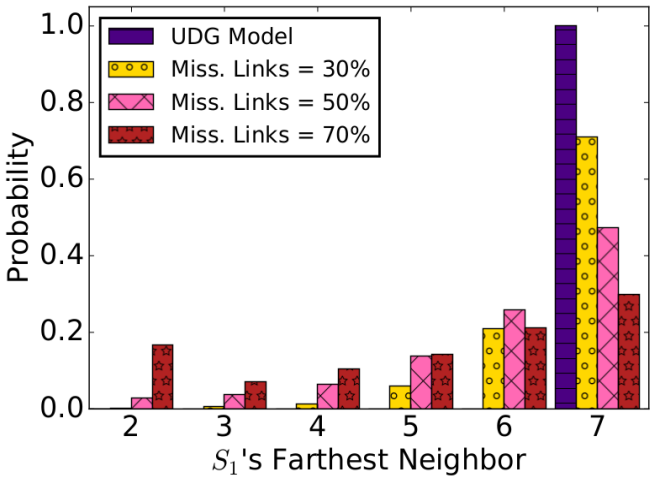


Figure 6.4: The coverage area of S_1, presented as the identifier of the farthest one-hop neighbor, for different missing link probabilities.

How to select the second anchor?

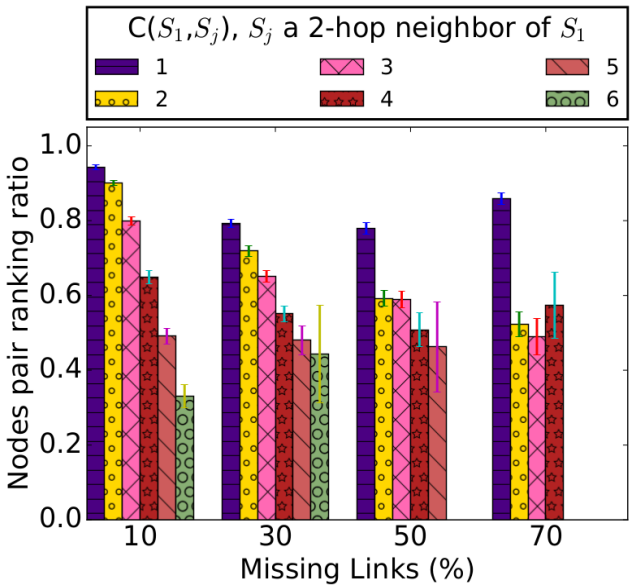


Figure 6.5: Sensors pair ranking ratio in the 1-hop neighborhood of sensor S_1, depending on the similarity between S_1 and its 2-hop neighbors selected as the second anchor, and the percentage of missing links in the UDG model.

Another consequence to consider a network with missing links can be noticed in the selection of the second anchor for each iteration, as described in Section 6.2. As explained, sensor S executing an iteration of the sensor ranking algorithm should ideally select as a second anchor the most distant unranked 2-hop neighbor. However, no information concerning the distance between sensors is available. Moreover, in realistic settings, with asymmetric links, even the definition of the 2-hop neighborhood needs to be clarified, as one sensor might hear another, but not the other way around.

Therefore, considering 2-hop neighbors S_i and S_j, we define a metric to guide the choice of

the second anchor among the 2-hop neighbors: $C(S_i, S_j)$, the similarity of sensors S_j related to S_i, defined as the number of sensors receiving messages from S_j and, at the same time, able to send messages to S_i. It can also be defined as the number of sensors having S_j in their 1- hop neighborhood, and at the same time being a 1-hop neighbor of S_i. Formally:

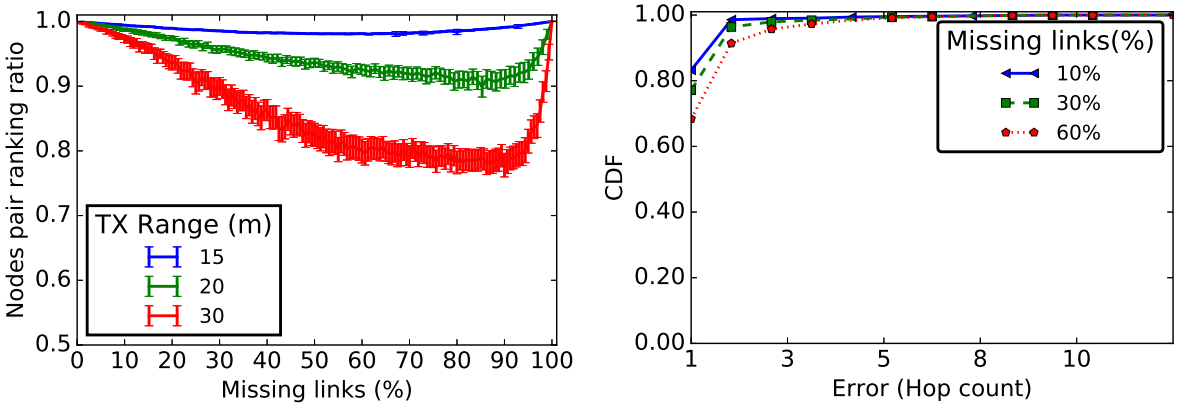
$$C(S_i, S_j) = |\{S_k \in \Gamma(S_i) \backslash S_j \in \Gamma(S_k)\}| \tag{6.3}$$

Figure 6.5 presents the ratio of correctly ranked pairs of sensors in the 1-hop neighborhood of the sensor S_1, depending on the choice of the second anchor in terms of similarity. This practically corresponds to the first iteration of the sensor ranking algorithm. The results show that the 2-hop neighbor with the lowest similarity gives the best ranking ratio. We use this insight in the following, to evaluate the overall performance of the iterative ranking algorithm.

6.5.2.3 Ranking performance

In this section, we evaluate the performance of our ranking algorithm for different deployments. For each deployment, 1000 different topologies are generated.

Uniform deployment



(a) Ranking ratio

(b) Ranking error

Figure 6.6: Ratio of the correctly ranked pairs of sensors as a function of the percentage of missing links for different transmission ranges (6.6a) and Ranking error for a transmission range of 20 m (6.6b): UDG-based channel model, uniform deployment.

We consider a network of 20 sensors uniformly deployed with a distance of 5 m between consecutive sensors. We vary the transmission range from 15 m to 30 m, i.e. $|\Gamma(S_1)|$ varies from 3 to 6. Figure 6.6a presents the performance of our ranking algorithm as a function of the percentage of missing links and the transmission range. Note that, since we assume there is always a symmetric link between two physically consecutive sensors, these links are not considered among the possible missing links. This means that, deleting all the other links in the network (i.e. 100% missing links) results in a backbone topology, in which each sensor has exactly two neighbors: its successor and predecessor in the physical deployment. That is why, in such a

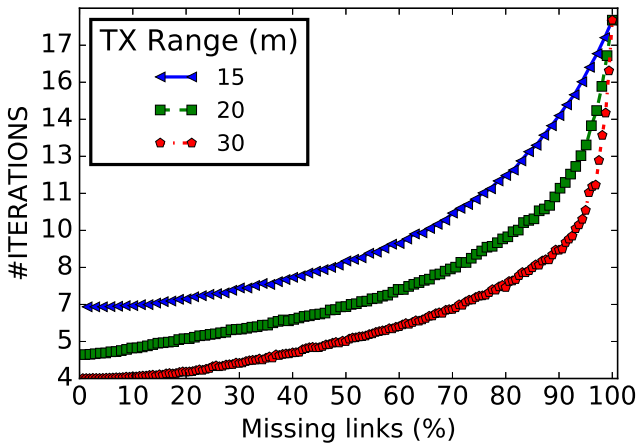


Figure 6.7: Number of iterations as a function of the percentage of missing links, for different transmission ranges: UDG-based channel model.

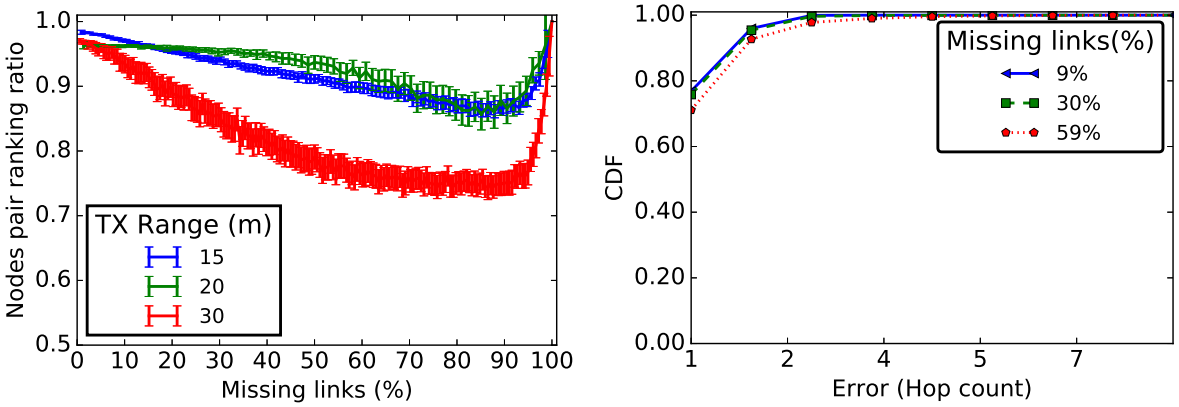
situation, 100% of pairs of sensors are well ordered, regardless the transmission range. This figure also highlights the good performance of the proposed algorithm: a correct ranking ratio of more than 0.95 is obtained for a transmission range of 15 m ($|\Gamma(S_1)| = 3$). We observe that, as the 1-hop neighborhood size increases (i.e. higher transmission range), the ranking performance decreases. However, in practice, a sensor with many 1-hop neighbors can prune some of them and use in the ranking algorithm only high quality links (based on the radio link PRR or RSSI). The trade-off of reducing the neighborhood size is an increase in the number of iterations required by the ranking algorithm (and thus the energy consumption), as it can be noticed in Figure 6.7. A direct consequence is therefore the increase of the overall energy cost required by the ranking algorithm; we note however that this algorithm is only executed after sensor deployment or re-deployment. Thus, we argue that it is worth paying the cost induced by link pruning in order to guarantee a good ranking performance.

For a transmission range of 20 m ($\Delta = 4$), Figure 6.6b presents the CDF of the ranking error for different percentage of missing links. This figure shows that, even when an error exists, the probability to observe a ranking error of only one hop is very high (80%). Thus, even with 30% of missing links, the ranking algorithm proposed in this chapter gives good performance in terms of ranking ratio and ranking error. Indeed, as shown by the Figure 6.6, when there are 30% of missing links, the ranking algorithm proposed in this thesis obtains a sensor pairs ranking ratio greater than 95% and a sensor position shift of one hop in case of ranking error.

Random deployment

We consider 20 sensors randomly deployed to form a LWSN. The sensors are deployed following a Poisson process with 0.2, the average number of sensors per unit length (m). Thus, the distance between two consecutive sensors follows an exponential distribution and 5 m is the average distance between two consecutive sensors. We vary the transmission range from 15 m to 30 m.

Figure 6.8a presents the sensors pair ranking ratio as a function of the missing links and the radio transmission range. As in the case of uniform sensors deployment, the sensors pair ranking



(a) Ranking ratio

(b) Ranking error, Transmission range = 20 m

Figure 6.8: UDG-based channel model, random deployment: (6.8a) Ratio of the correctly ranked pairs of sensors as a function of the percentage of missing links, for different transmission range (6.8b) Ranking error for a transmission range of 20 m.

ratio decreases as the transmission range increases. Even when one link over two is missing (50% of missing links), Figure 6.8a shows a ranking ratio greater than 95%, for a transmission range lower than 20 m.

Figure 6.8b presents, for a transmission range of 20 m, the CDF of the ranking error in hop count for different percentage of missing links. This figure shows that, as for the uniform deployment, when a ranking error exists, there is a high probability to have a sensor position shift of only one hop from its real position. Nevertheless, this probability is smaller compared to a uniform deployment. This is due to the non uniform sensors degree that can be observed in a network when a random deployment is considered.

Results presented in Figure 6.8 show that, even in case of a random deployment, the ranking algorithm proposed in this chapter leads to good results. The random deployment considered in this section produces network topologies in which sensors degree are not uniform. Non-uniform sensors density in all areas of the network, unreliable and asymmetric links are some properties of real WSN. These properties are generally observed in real networks because of the sensors deployment, or the spatio-temporal variation of links between sensors.

Virtual nodes-based deployment

In this section, we evaluate the performance of our ranking algorithm considering the virtual nodes-based deployment proposed in Chapter 5. If the previous results showed that our ranking algorithm is good enough to rank physical sensors, what about the use of such algorithm applied to a network in which these physical sensors are deployed using the concept of virtual nodes as defined in Chapter 5. This deployment assumes that sensors of the virtual node B_i are able to communicate with those of virtual nodes B_{i-1} and B_{i+1}. It also assumes that sensors of B_i can communicate with those of B_{i-2} or B_{i+2} with a probability p. Thus, it is important to evaluate our ranking algorithm considering topologies satisfying these assumptions. In this section, we build topologies which guarantee communication between sensors of the same virtual node, and

also between sensors of virtual node Bi and those of virtual nodes B_{i-1} and B_{i+1}. The topologies built also guarantee a communication between sensors of B_i and those of B_{i-2} and B_{i+2} with probability p. In case of communication between sensors of B_i and those of B_{i-2} and B_{i+2}, for a given sensor S_k of B_i, the number of sensors of B_{i-2} or B_{i+2} connected with S_k is randomly selected. Considering a LWSN organized in 10 virtual nodes ($K = 10$) covering 10 m each and $\alpha = 0.2$ (the probability for a sensor to overhear a message addressed to another one), we evaluate the performance of the ranking algorithm proposed in this chapter as a function of p and N (the number of deployed sensors).

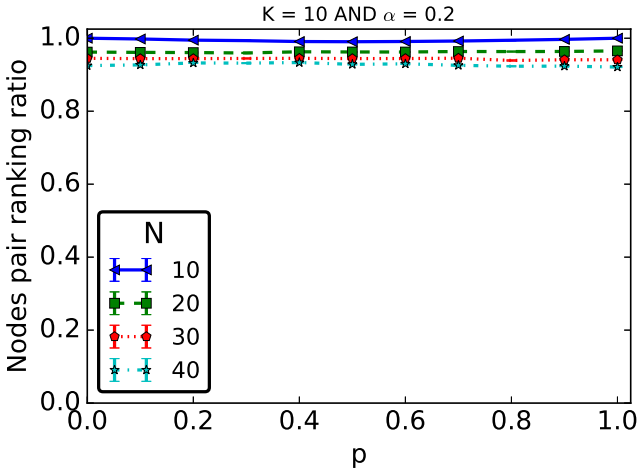
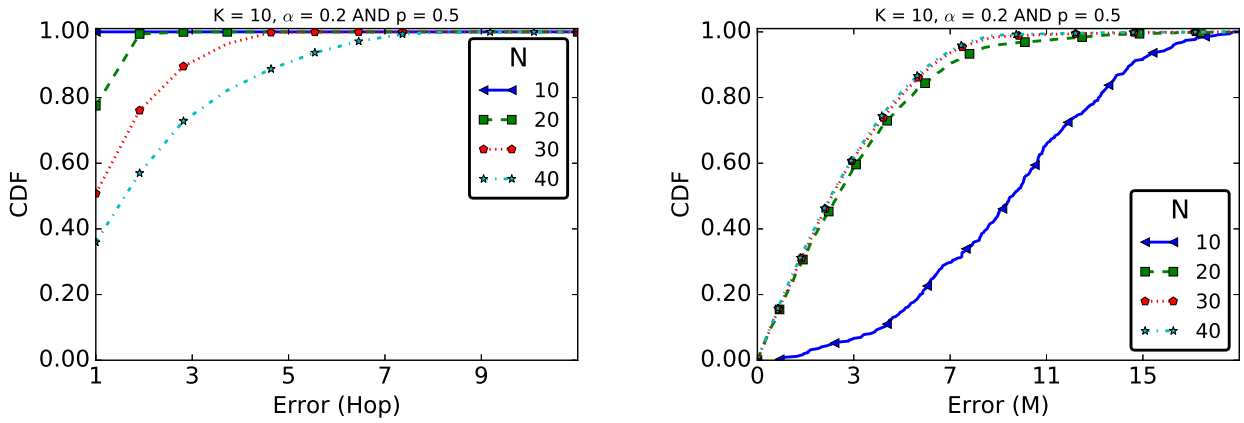


Figure 6.9: Sensor pairs ranking ratio as a function of p and N: UDG-based channel model, virtual nodes-based deployment, 10 virtual nodes, virtual node length 10 m, $\alpha = 0.2$.



(a) Ranking error in hop count

(b) Ranking error in meter

Figure 6.10: CDF of the ranking error in hop count (6.10a) and in meter (6.10b) for some values of N: UDG-based channel model, virtual nodes-based deployment, 10 virtual nodes, virtual node length 10 m, $\alpha = 0.2, p = 0.5$.

Figure 6.9 shows the sensors pairs ranking ratio as a function of N (the number of sensors) and p (the probability of communication between sensors of the virtual node B_i and those of B_{i-2} and B_{i+2}). Considering different network sizes allows to evaluate the impact of sensors density on our ranking algorithm. But, the sensor density is not uniform in all the area of the network.

Indeed, as shown in Chapter 5, increasing the number of sensors in the network particularly increases the sensor density in the virtual nodes close to the sink. Results presented in Figure 6.9 show that, no matter the value of p, as for uniform (see Figure 6.6a) and random (see Figure 6.8a) sensors deployments, we obtain a sensors pairs ranking ratio greater than 95% when we consider the virtual nodes-based deployment proposed in Chapter 5. The ranking algorithm gives a perfect results for $N = 10$ (i.e. one sensor per virtual node since we assume $K = 10$ virtual nodes). By increasing the total number of sensors (or sensor density), the sensors degree decreases as for the uniform and random deployments.

The good performance of our ranking algorithm is also highlighted by the ranking error presented in Figure 6.10. Indeed, this figure shows that, even in the case of a ranking error, we have good chance that the shifted sensors remain in their same virtual node. For $N = 20, 30$ or 40, even if we observe a ranking error of 3 or 5 hops (Figure 6.10a), we observe in Figure 6.10b that, we have 90% of chance to observe an error less than 10 m, the length of a virtual node. For $N = 10$, we have a perfect ranking as shown in Figure 6.9. This result is confirmed by the Figure 6.10a which shows the ranking error in hops. It is important to note that, event in the case of a perfect raking, due to the centroid algorithm, a sensor coordinate might shift from its real position. That is why in Figure 6.10b we observe a high probability to obtain a ranking error greater than 10 m.

In summary, results presented in this section show that, even by considering the virtual nodes-based deployment proposed in Chapter 5, the ranking algorithm proposed in this chapter gives very good results. And in case of ranking error, the chance that the falsely ranked sensors remain in the same virtual node is high. This is interesting since it reduces the impact of the ranking error on the communication protocols or network mechanisms working at virtual nodes level.

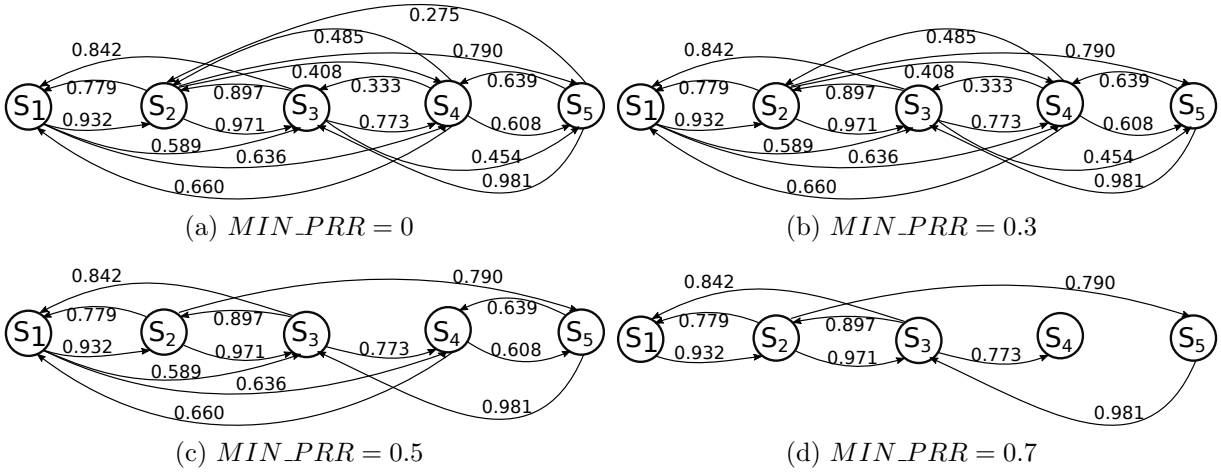
6.5.3 The case of a channel model from empirical data

In this thesis, we propose WARIM (see Chapter 4), a new realistic WSN architecture for vehicular traffic monitoring at an intersection. It is an architecture in which sensors are deployed at ground level and establish multi-hop communication paths to transmit vehicular traffic data collected at lane level to the sink. WARIM is proposed based on the results obtained after the extensive experimentation campaigns conducted in Chapter 3 with sensors deployed at ground level. In this section, using data on links quality collected during this experimentation campaign, we build an empirical channel model. Our goal is to evaluate the ranking algorithm proposed in this thesis considering a more realistic channel model representing the wireless connectivity at ground level.

6.5.3.1 Evaluation methodology

From a networking point of view, the link quality can be defined by the PRR. Considering links between the pairs of sensors separated by a distance d, we have a set of PRR values $S_{PRR}(d)$ collected during our experimentation campaigns. Our goal is to design a model allowing to define, for each link of distance d, an empirical PRR value. From the set $S_{PRR}(d)$ of PRR collected for the distance d, we calculate μ_d and σ_d, which are respectively, the average and standard deviation

of PRR values in $S_{PRR}(d)$. Thus, for each link of length d, the PRR value is generated from a normal distribution of mean μ_d and standard deviation σ_d. The ranking algorithm proposed in



(a) $MIN\_PRR = 0$ (b) $MIN\_PRR = 0.3$

(c) $MIN\_PRR = 0.5$ (d) $MIN\_PRR = 0.7$

Figure 6.11: Some example of topologies for different values of MIN_PRR.

this thesis is based on the sensors neighborhood. As for the theoretical channel model, we consider a network with asymmetric links, i.e, between sensors S_i and S_j, we have two independent links $S_i \rightarrow S_j$ and $S_j \rightarrow S_i$. The PRR values of these two links are randomly selected from the normal distribution previously defined. For each link, the PRR is a value less than or equal to 1 and greater than 0. The question now is: how does a sensor S_i define its neighbors? We define MIN_PRR, a threshold allowing to select valid links. A link is valid if its PRR is greater than MIN_PRR. For instance, when $MIN\_PRR = 0$, all links in the network are valid and when $MIN\_PRR = 0.5$ (or 50%), a link is valid if and only if its PRR is greater than 0.5. Figure 6.11 illustrates some topologies considering different values of the threshold (MIN_PRR). In this section, we evaluate the performance of our ranking algorithm as a function of this threshold.

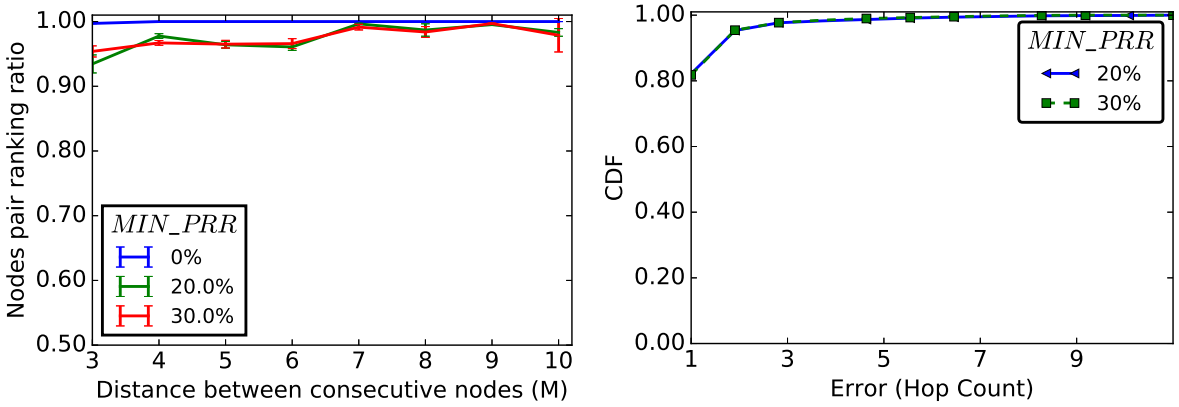
As for the theoretical channel model, we evaluate the performance of our ranking algorithm considering three different sensor deployments: uniform, random and the virtual nodes-based deployment proposed in Chapter 5, but considering the channel model based on the experiments of the Chapter 3. Considering a given deployment, we generate 1000 topologies. Results presented in this section are an average obtained from the results of each of the 1000 topologies. In this section, the connectivity metric defined by Eq. 6.3 is also used to select the second anchor.

6.5.3.2 Ranking performance

Considering the empirical channel model described in the previous section, we evaluate the ranking algorithm proposed in this thesis. For each of the three deployments considered in this chapter, we evaluate our ranking algorithm in terms of the sensors pair ranking ratio and ranking error.

Uniform deployment

This section presents the results of our ranking algorithm for a uniform sensors deployment. We consider 20 sensors uniformly deployed with a maximum transmission range of 30 m. Figure 6.12



(a) Ranking ratio

(b) Ranking error

Figure 6.12: Ratio of the correctly ranked pairs of sensors as a function of the distance between two consecutive sensors (6.12a) and CDF of the ranking error when the distance between two consecutive sensors is 5 m (6.12b): Empirical channel model, uniform deployment.

shows the sensors pair ranking ratio as a function of the distance (in meters) between consecutive sensors and MIN_PRR. Considering different separation distances (in meters) between two consecutive sensors allows to evaluate the impact of the sensors density, sensors degree and network diameter on our ranking algorithm. Recall that MIN_PRR is a threshold used by sensors to choose their neighbors. This parameter also allows to evaluate the impact of the sensors degree on our ranking algorithm. Indeed, by increasing the value of this parameter, this decreases the neighborhood size of sensors.

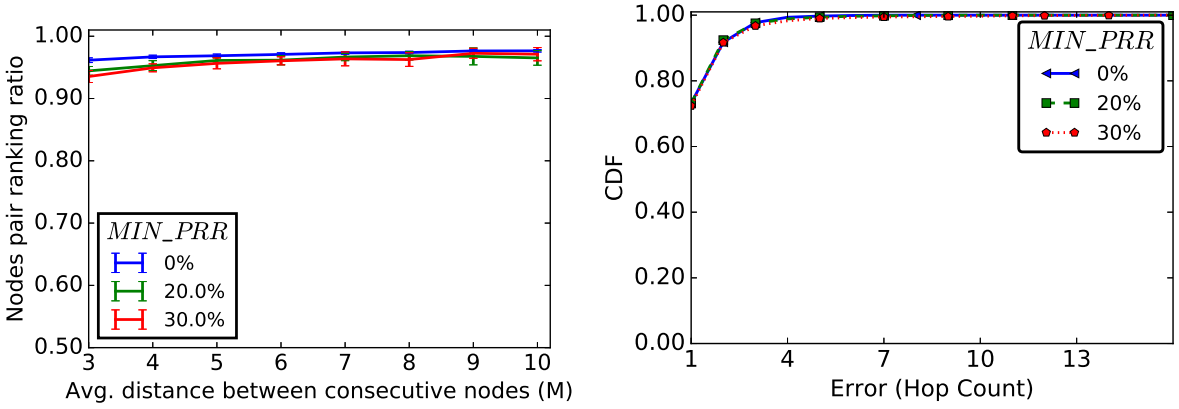
By observing this figure, the main conclusion is that, even with the empirical channel model, the ranking algorithm proposed in this thesis gives a good ranking ratio greater than 0.9. It is also important to note in this figure that, by increasing the distance between sensors, the ranking ratio increases too. Indeed, increasing the distance between sensors in a uniform sensors deployment reduces the sensors degree. Thus, this reduces the probability of ranking error. Also, our results highlight that week links, i.e. links with low PRR are useful.

In order to improve the data transmission reliability in a WSN, a sensor will always select the link with the highest PRR values to forward a message. For the ranking algorithm, we observe an intriguing result in the Figure 6.12a: the best ranking is obtained when all links are considered. Since the goal of the centroid approach is to bring close physically connected neighbors, the higher is the number of neighbors identified, the better will be the calculated position.

For a distance of 5 m between consecutive sensors, Figure 6.12b presents the CDF of the ranking error for three values of MIN_PRR. This figure shows that, in case of ranking error, no matter the MIN_PRR values, almost 70% of the sensors are shifted by one position only.

Random deployment

Figure 6.13 shows, for a random deployment, the sensors pair ranking ratio (Figure 6.13a) as a function of the average distance between two consecutive sensors. As for a uniform deployment, we consider 20 sensors. This figure also shows the ranking error (Figure 6.13b) when the average



(a) Ranking ratio

(b) Ranking error

Figure 6.13: Ratio of the correctly ranked pairs of sensors as a function of the average distance between two consecutive sensors (6.13a) and CDF of the ranking error when the average distance between two consecutive sensors is 5 m (6.13b): Empirical channel model, random deployment.

distance between sensors is 5 m. Results are presented for three different values of MIN_PRR: 0%, 20% and 30%. Even for a random deployment, we observe good performances with a sensors pair ranking ratio greater than 95%. This observation is valid for a dense deployment (average distance between consecutive sensors is equal to 3 m) and for a sparse deployment (average distance between consecutive sensors 10 m). As for the uniform deployment, we obtain the best performance when $MIN\_PRR = 0$, i.e., each sensor considers all other sensors identified in its neighborhood, no matter the PRR value. Moreover, Figure 6.13 shows that, even in the case of ranking error, we have a high probability to observe a sensor position shift of at most 2 hops. It is difficult to guarantee a strictly uniform sensors deployment in a WSN. The results presented in this section show that, even in a random deployment guaranteeing network connectivity, the ranking algorithm proposed in this thesis shows not only its efficiency in terms of sensors ranking in a LWSN, but also its robustness to deployment errors.

Virtual nodes-based deployment

In this section, we evaluate the performance of our ranking algorithm considering the virtual nodes-based deployment proposed in Chapter 5 under a channel model coming from our experimental study in Chapter 3. We consider a network divided into 10 virtual nodes covering 10 m each. To satisfy the connectivity assumptions of the virtual nodes-based deployment, all links between sensors in the same virtual node or between sensors of virtual nodes B_i and B_{i+1} or B_{i-1} are maintained. Thus, the PRR threshold (MIN_PRR) concerns only links between the virtual nodes B_i and B_{i+2} (and B_{i-2}). In our virtual nodes-based deployment proposed in Chapter 5, we defined by p the communication probability between sensors in virtual node B_i and those in virtual nodes B_{i+2} and B_{i-2}. Thus, in our simulations, the MIN_PRR values which gives the connection probability between B_i and B_{i+2} (or B_{i-2}) the closest to p is chosen.

Figure 6.14 shows the performance of the ranking algorithm proposed in this thesis as a function of N and p for $\alpha = 0.2$. As for the previous deployments, our centroid-based algorithm

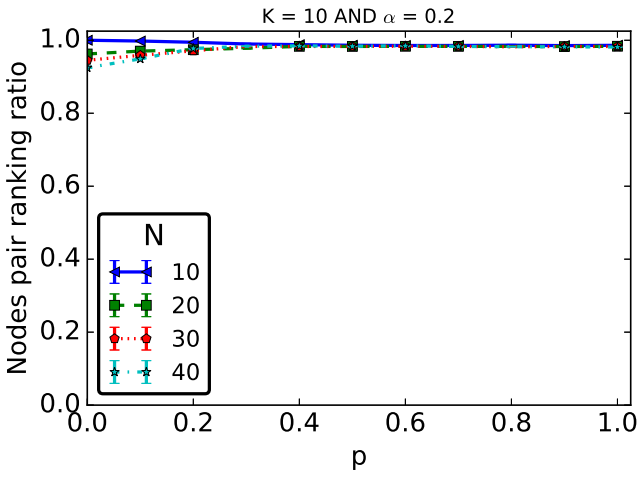
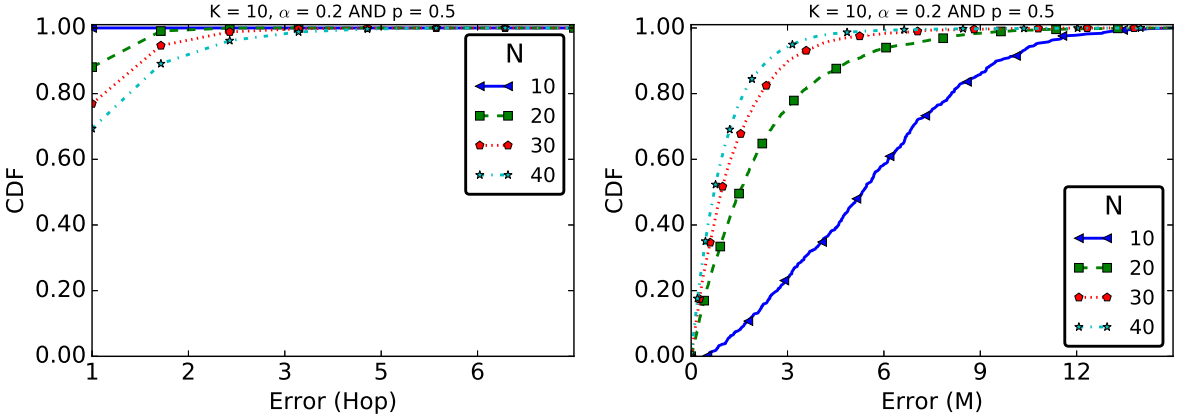


Figure 6.14: Sensor pairs ranking ratio as a function of p and N: Empirical channel model, virtual nodes-based deployment, 10 virtual nodes, virtual node length 10 m, $\alpha = 0.2$.



(a) Ranking error in hop count

(b) Ranking error in meter

Figure 6.15: CDF of the ranking error in hop count (6.10a) and in meter (6.10b) for some values of N: Empirical channel model, virtual nodes-based deployment, 10 virtual nodes, virtual node length 10 m, $\alpha = 0.2, p = 0.5$.

gives a sensors pairs ranking ratio greater than 95% no matter the values of p or N. Figure 6.15 shows the CDF of the ranking error in hop count (Figure 6.15a) and in meters (Figure 6.15b) when $p = 0.5$. As for previous deployments, Figure 6.15a shows that, in the case of a ranking error, the probability to have a sensor position shift of at most 2 hops is very high (0.9). Recall that in the virtual nodes-based deployment proposed in Chapter 5 and considered in this section, the network is divided into virtual nodes, and a given number of sensors is deployed in each virtual node. Figure 6.15b shows that, most of the time (probability 0.9), the sensor position shift is at most of 3 m (except for $N = 10$, i.e. one sensor per virtual node). Thus, even in the case of ranking error, the probability that the sensor remains in the same virtual node (since the virtual node length is equal to 10 m) is greater than 95%.

6.5.4 Comparison to related work

In Section 6.2, we presented related work concerning the nodes ranking in LWSN. These works can be classified into two groups: those which use data provided by the hardware to determine the physical proximity between nodes, and those which exploit the neighborhood correlation. In the later group, we presented a centroid-based solution. Based on this solution, we propose in this thesis (see Section 6.2) a new ranking algorithm for nodes deployed to form a LWSN. In our new algorithm, the only parameter manually configured is the first node. Our goal in this section is to compare our ranking algorithm to another one that can be considered as belonging to this second group.

6.5.4.1 A common neighbors based solution

The number of common neighbors is a metric usually used in the literature to determine the physical proximity between sensors in WSN [124, 125, 137]. Typically, more two sensors have common neighbors, more they are physically close. Let S_i be a ranked sensors running the ranking algorithm and S_j, a neighbor of S_i, not yet ranked. The proximity calculated by S_i related to S_j is defined as follows:

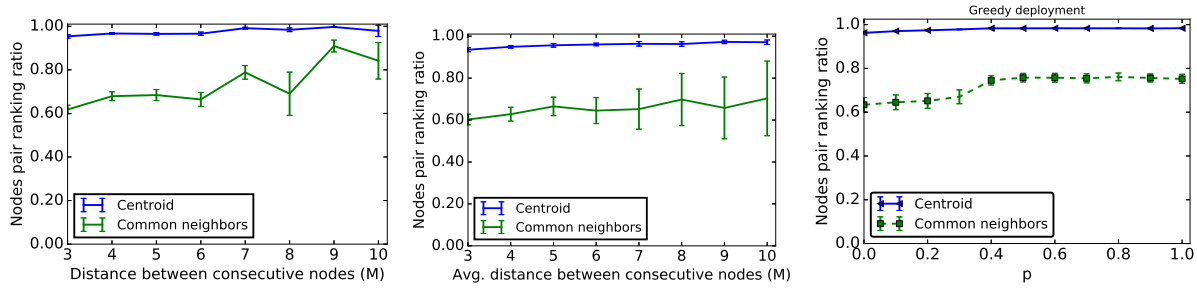
$$common(S_i, S_j) = |\Gamma(S_i) \cap \Gamma(S_j)| \tag{6.4}$$

We implemented a new version of our iterative ranking algorithm. At each iteration, the current sensor, S_i, computes the objective function (see Eq. 6.4) for all its neighbors. The one with the maximum number of common neighbors is considered as the physically closest one and is selected to execute the next iteration of the algorithm.

6.5.4.2 Comparison results

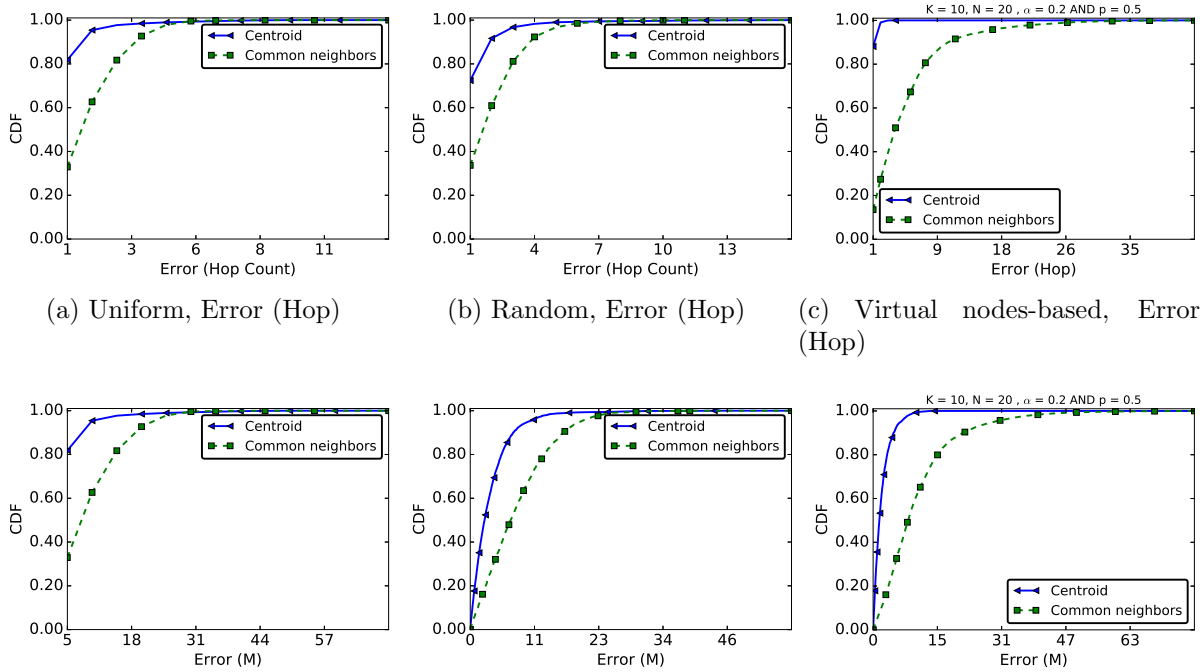
Figures 6.16 and 6.17 compare our centroid-based algorithm to the common neighbors-based one presented in the previous section in terms of sensors pairs ranking and ranking error respectively. For this comparison, we consider a network of 20 sensors and a channel model built from the empirical data. We also consider the three different deployments considered in the previous sections: uniform, random and virtual nodes-based deployment proposed in Chapter 5. For the virtual nodes-based sensors deployment, we consider a LWSN organized to form 10 virtual nodes of length 10 m each and $\alpha = 0.2$.

No matter the deployment considered, the centroid-based ranking algorithm outperforms the common neighbors-based algorithm in terms of sensors pairs ranking ratio (see Figure 6.16a-6.16c) or ranking error (see Figure 6.17a-6.17f). Considering the virtual nodes-based deployment, while we observe a ranking ratio of 60% with the common neighbors-based approach, our centroid-based approach gives a sensors pair ranking ratio more than 95%. And in case of ranking error (still considering the virtual nodes-based deployment), while the probability to have an error of one hop is almost 90% for the centroid-based algorithm, this probability is only 15% for the common neighbors-based approach. These results highlight the robustness of the centroid-based approach



(a) Uniform deployment (b) Random deployment (c) Greedy deployment

Figure 6.16: Sensors pair ranking ratio of the centroid-based and common neighbors-based algorithm for uniform, random and virtual nodes-based deployment for $N = 20$. In the case of the virtual nodes-based deployment, $K = 10$ and $\alpha = 0.2$ and the virtual node length is 10 m.



(a) Uniform, Error (Hop) (b) Random, Error (Hop) (c) Virtual nodes-based, Error (Hop)

(d) Uniform, Error (meter) (e) Random, Error (meter) (f) Virtual nodes-based, Error (meter)

Figure 6.17: Ranking error of the centroid-based and common neighbors-based algorithm for uniform, random and virtual nodes-based deployment for $N = 20$. In the case of the virtual nodes-based deployment, $K = 10, \alpha = 0.2$ and $p = 0.5$ and the virtual node length is 10 m.

in presence of unreliable and asymmetric links, compared to a simple common neighbors-based one.

6.6 Conclusion

In this chapter, we proposed an iterative and distributed centroid-based algorithm for sensor ranking in a WSN with linear topology. Our algorithm takes as input local information regarding sensor connectivity and the single manually configured parameter is the first sensor of the network. The proposed algorithm iteratively selects two sensors as anchors and correctly ranks

more and more sensors in the network. At each iteration, all the processing done by the selected sensor depends on its neighborhood size and not on the number of sensors in the entire network. Thus, our solution is scalable and lightweight considering the deployment and maintenance of a large network.

Using a simulator developed during this thesis, we evaluated through extensive simulations the performance of our ranking algorithm. We consider various sensor degrees, realistic channel models (a model build from the simple UDG model and another one build from empirical PRR values) and different deployments (uniform, random and the virtual nodes-based deployment proposed in Chapter 5). In most of our simulation scenarios, we observed a sensors pair ranking ratio greater than 95%. In case of ranking error, results show a high probability to have an error of one hop only.

Chapter 7

Data collection in a WSN with linear topology

This chapter puts together all our contributions and evaluates the capability of WARIM to collect vehicular traffic at an intersection in as realistic conditions as possible. Thus, this chapter does not gives a new contribution of the thesis, but validates WARIM, the new infrastructure-free WSN architecture proposed in this thesis for vehicular traffic monitoring at intersections. Since in WARIM nodes are deployed on road lanes, similarly to the previous chapters, this one focuses on LWSN.

In WSN, sensors are limited by their radio transmission power and thus their coverage area. In such conditions, a source sensor may not be able to directly communicate with the sink. Thus, we design a routing protocol in our evaluation. In WSN, the poor quality and instability of links are constraints widely studied and highlighted in the literature through theoretical analysis or experiments on physical sensors. Experimental results presented in Chapter 3 considering nodes deployed at ground level, as in WARIM, confirm these observations. Moreover, these results highlight the negative impact of the ground proximity on link quality.

For a realistic evaluation of WARIM, the vehicular traffic at the intersection and the link properties at ground level must be taken into account. These two elements are considered in the evaluation conducted in this chapter. Concerning the topological aspect, we consider in this chapter a WSN deployed using the greedy deployment strategy proposed in Chapter 5. Given the poor and unstable nature of links at ground level, as shown in Chapter 3, we use a gradient-based opportunistic routing protocol to collect the messages generated by car detectors. In our simulations, we use the sensor rank as the gradient.

The content of this chapter is organized as follow. Section 7.1 describes the problem statement addressed in this chapter. We focus on vehicular traffic in Section 7.2. In this section, by analyzing a dataset of the city of Koln (Germany), we study messages generation strategies which reduce the total number of messages generated by sensors in the network. By reducing the number of messages circulating in the network, we also reduce the energy consumption of sensors. In Section 7.3, we present the main idea and some advantages of the opportunistic routing protocols, a category of routing techniques adapted for WSN given their properties and constraints. We also propose a simple gradient-based opportunistic routing protocol used in our simulations. Using a

WSN simulator which takes into account various aspects of wireless communication (propagation model, encoding, collision, interferences, etc), considering the greedy deployment and the ranking mechanism proposed in Chapter 5 and 6 respectively, we evaluate in Section 7.4 the performance of Warim in terms of packet delivery ratio, traffic load per sensor and sink reception cycle. Finally, we conclude this chapter in Section 7.5.

7.1 Problem Statement

In this chapter, our goal is to evaluate the data collection or the monitoring capability of Warim. Using a dedicated routing protocol, this evaluation takes into account the vehicular traffic model and the links properties at ground level. It also considers the greedy deployment proposed in Chapter 5, as well as the sensors ranks (used as gradient for the gradient-based opportunistic routing protocol) provided by the centroid-based ranking algorithm proposed in Chapter 6.

Because of the radio quality/characteristics and the environment constraints, the coverage area of a sensor is time and space dependent. During this thesis, we conducted extensive experiments with physical sensors deployed at ground level. Experimental results presented in Chapter 3 show that the ground proximity has a negative impact on the link quality, compared to a link between nodes deployed at a given height. Even if the available WSN simulators propose more or less realistic models for wireless communication channels, the particular properties of the link at ground level must be taken into account in our simulations.

In a multi-hop WSN, a sensor may build and maintain a routing path to forward messages to the sink [111]. With such a routing protocol, when a sensor has a message to send, it just selects, in its local routing table, the next hop to which the message will be relayed. Sensors may also decide to not maintain any neighbors or routing table since links are unstable by nature in WSN. In such a routing protocol, when a sensor has a message to forward, it just sends it in broadcast mode. Some of its neighbors which receive the message organize themselves to select the next hop: it is the opportunistic routing [141]. These neighbors must be those closer to the sink than the previous sensor. Warim is designed with a multi-hop perspective. Moreover, considering the results presented in Chapter 3, such a routing protocol must be implemented in our simulator for a realistic evaluation of Warim.

It is well known that generally the radio is the component of a sensor that consumes the most energy, for messages transmission and reception [46]. Thus, reducing the radio activity or the number of messages sent or received is a real challenge in WSN. In WSN, a sensor can generate a message each time it detects an event (event-based messages generation). With such a strategy, the number of messages generated by the sensors will depend on the number of events or the application. A sensor can also divide its time into slots and generate a single message in each slot (this is the periodic message generation approach). For this message generation strategy, the number of messages generated by sensors depends on the slot duration.

Considering the case of a vehicular traffic monitoring application, the number of vehicles passing at an intersection may depend not only on the geographical position of the intersection,

but also on the time of day. Thus, it is interesting to design a message generation strategy that will be used by sensors with the goal to reduce the number of messages in the network in order to prolong the sensors functioning time.

An energy efficient greedy deployment and a centroid-based ranking algorithm for LWSN are proposed in Chapter 5 and 6 respectively. In this chapter, we evaluate the capability of WARIM to collect vehicular traffic data at road intersection considering this deployment. The gradient-based opportunistic routing protocol used to collect messages uses nodes ranks as gradient in order to forward the messages generated by sensors toward the sink. The main objective of this chapter is to highlight if, by putting all our contributions together, WARIM is able to collect vehicular traffic data at an intersection. Moreover, even if the results presented in Chapter 6 show good ranking performance, it is interesting to evaluate the impact of using a suboptimal sensor rank as gradient on the packet delivery rate (PDR).

7.2 Messages generation strategies

7.2.1 Event-based and periodic messages generation

Our goal in this section is to present the relationship between the message generation approach and the total number of messages effectively transmitted in the network.

7.2.1.1 System model

We consider sensors deployed linearly on a lane leading to an intersection. Two approaches can be used by sensors to generate messages: the event-based approach and the periodic approach. In the event-based approach, the sensor generates a message once it detects a vehicle. In the periodic approach, each sensor generates one message after a given period. For the periodic message generation approach, we denote by φ the period. Vehicles arriving on a lane can be modeled as a Poisson process of parameter γ [138]. γ is the average number of vehicles passing on the corresponding lane during a unit of time. Then, if the system is observed during a time window T, the number of messages generated in the network by a given sensor using an event-driven approach is defined as: $M_{evt} = \gamma \cdot T$. For a periodic approach, the number of messages generated by a sensor during a time window T is defined as: $M_{pec} = \frac{T}{\varphi}$

7.2.1.2 Number of messages generated

For $T = 100$ minutes, Figure 7.1 and 7.2 show the number of messages generated in the network as a function of the period (Figure 7.1) and the average number of vehicles passing through a given lane per minute (Figure 7.2). As it can be observed, the performance of the two approaches depends on the intensity of the vehicular traffic. Indeed, for an average arrival rate of $\gamma = 3$ vehicles per minute, less messages are generated by a periodic approach when the period φ is greater than 20 s. Given the dynamic nature of vehicular traffic, neither a periodic nor an event-driven approach can be optimal regarding the overall number of messages generated in the network.

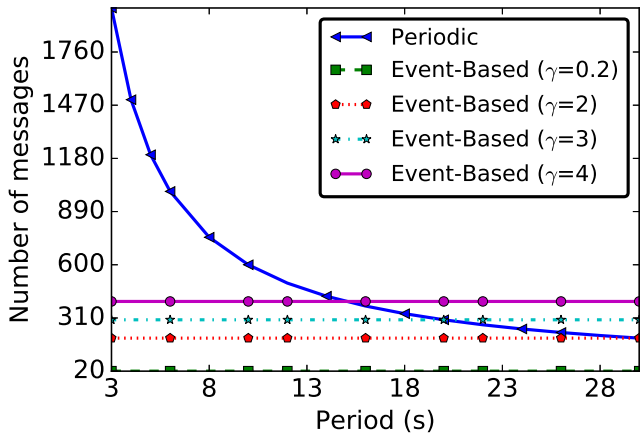


Figure 7.1: Periodic: Number of messages as a function of φ and γ.

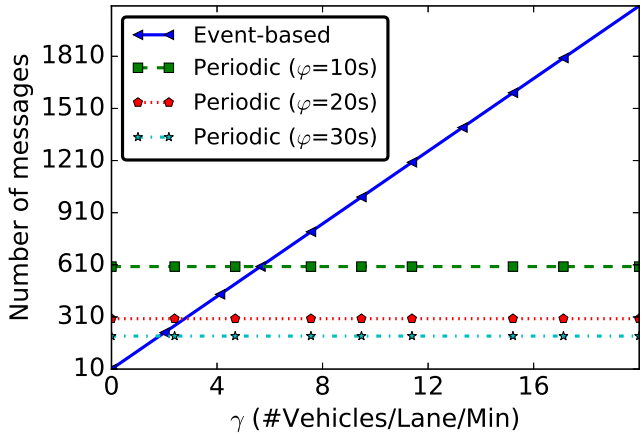
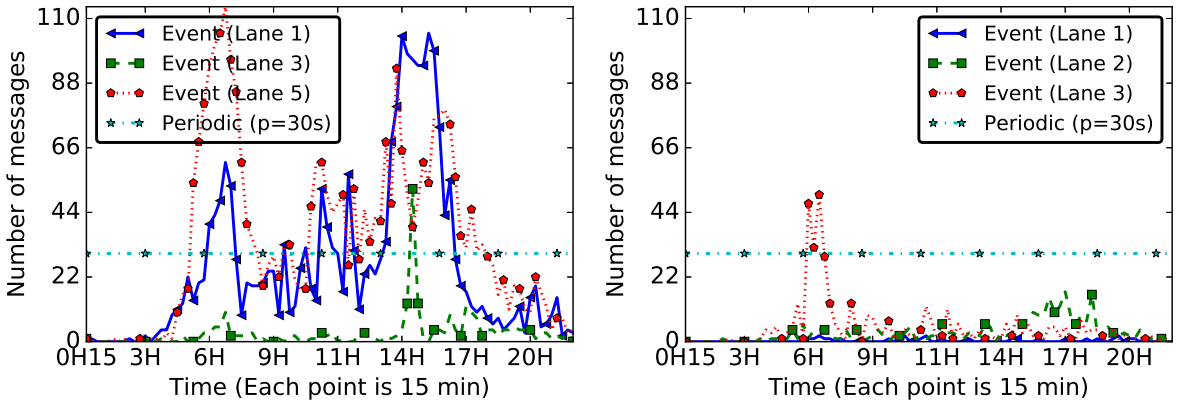


Figure 7.2: Event-based: Number of messages as a function of φ and γ.

7.2.1.3 An application example : Koln vehicular data traffic analysis

Considering a real vehicular traffic data, we analyze in this section the number of messages generated in the network using the two messages generation approach described in the previous section. We have analyzed the vehicular traffic at intersections in the city of Koln (Germany) using a $24h$ dataset [139, 140] indicating, at each second, the coordinates and the speed of each vehicle. We divided this data into blocks of 15 $min.$, and for each block, we calculate for each lane of an intersection the number of vehicles passing through it. The number of messages generated by the two approaches is then compared. For the periodic approach, we consider a period of 30 s, a reasonable delay for traffic light management. Figure 7.3 shows, for two different intersections (a dense one (named Intersection 1 in Figure 7.3) and a sparse one (named Intersection 2 in Figure 7.2)), the number of messages generated as a function of the block number (x-axis), the generation approach and the lane. Intersection 1 has 8 entrance lanes, while intersection 2 has 4. For presentation purpose, Figure 7.3 shows, for each intersection, only results for three lanes. In these figures, each point on the x-axis represents an interval of 15 min.

Results presented in Figure 7.3 show that the vehicular traffic is space and time dependent. Indeed, from an intersection to another, or from one lane to another, vehicular traffic may be different. And on a given lane, the vehicular traffic varies depending on the time of day. For



(a) Intersection 1: 15790 VEHICLES/DAY (b) Intersection 2: 908 VEHICLES/DAY

Figure 7.3: Impact of vehicular traffic on the number of messages.

instance, Figure 7.3a shows that, during the morning or the evening, when there is a peak in the vehicular traffic, it is better to use a periodic approach to reduce message generation.

7.2.2 Adaptive messages generation

Results presented in the previous section highlight the dynamic nature of vehicular traffic, and show that neither periodic nor event-driven approach is better in a realistic context. In this section, we propose a new adaptive message generation approach using only the local vehicular traffic, as detected by the sensors.

7.2.2.1 Adaptive algorithm

We assume that time is divided into w fixed-size windows $W_0, W_1, W_2, ..., W_{w-1}$ of duration W and that, at startup, all nodes use an event-driven approach. We define by V_i the number of vehicles that passed on a sensor deployed on a lane during the window W_i. We also define by θ a threshold allowing a sensor to decide if it should change the message generation strategy or not, based on the vehicular traffic it sensed in the previous window. The Algorithm 7.1 is executed by each sensor at the beginning of window $W_i, i > 0$. This algorithm takes as input the number of vehicles (V_{i1}) detected during the previous windows (W_{i1}) and the threshold (θ).

Algorithm 7.1 Adaptive messages generation

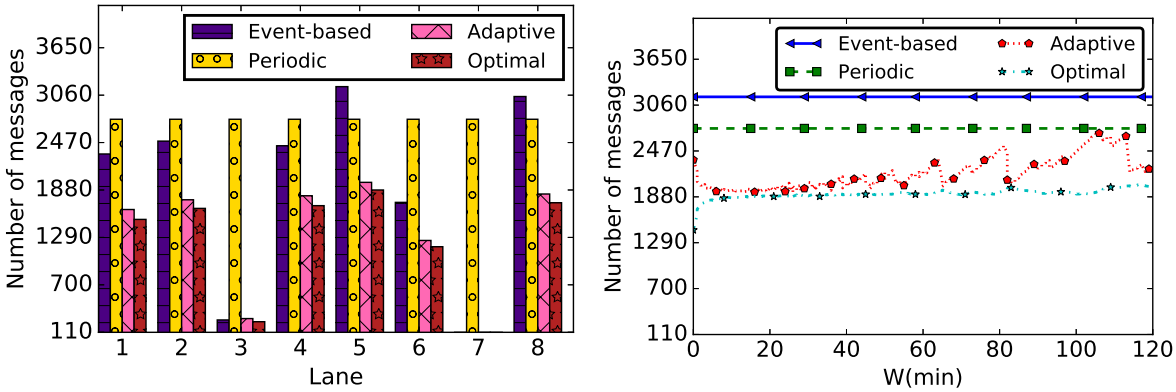
Input: V_{i-1}, θ

```
1: if Current_Approach = Event_Based then
2:     if V_{i-1} > θ then
3:         Current_Approach ← Periodic
4:     end if
5: else
6:     if V_{i-1} < θ then
7:         Current_Approach ← Event_Based
8:     end if
9: end if
```

A sensor using the event-based approach (Line 1) will switch to the periodic one if the number of vehicles detected during the previous windows is greater than the threshold (Line 2-4). And a sensor using the periodic approach (Line 5) will switch to the event-based one if the number of vehicles detected during the previous windows is less than the threshold (Line 6-9).

7.2.2.2 Adaptive approach versus event-based and periodic approaches

In this section, we compare our adaptive approach to periodic and event-driven ones. In this evaluation, we set $\theta = \frac{W}{\varphi}$, which corresponds to the number of messages generated by a sensor in a window W_i using a periodic approach with period φ. As in Section 7.2.1.3, we set the period φ to 30 seconds. Figure 7.4 presents the total number of messages generated in the network per lane only for Intersection 1, which observes more vehicular traffic. The optimal approach in this figure corresponds to the situation where each sensor knows, for each W_i, the value of V_{i+1} and sets the message generation strategy accordingly. In Figure 7.4a, we set $W = 900\ s$ (15 minutes) and present results for each lane of the corresponding intersection, while in Figure 7.4b, we evaluate the impact of W on the performance of our adaptive strategy. Figure 7.4a shows



(a) Intersection 1: All lanes, $W = 15min.$ for adaptive and optimal

(b) Intersection 1, Lane 5: Impact of W

Figure 7.4: Impact of adaptive message generation. For periodic messages generation, $\varphi = 30\ s$.

that our approach significantly reduces the number of messages generated compared to the two static approaches. The difference is largest in case of lanes with dense vehicular traffic. For instance, on lane 5 (Figure 7.4a), compared to an event-driven approach, our adaptive approach reduces by 40% the number of generated messages. Compared to the optimal approach, our adaptive solution needs one W period before switching to the best message generation approach, resulting in a certain number of unnecessary messages. This difference with respect to the optimal can be observed in 7.4b. We can notice that, for very small values of W, our solution shows a degraded performance, as for such a short time scale, the correlation between V_i and V_{i+1} is weak. A degraded performance can also be noticed for large W values, as in these cases, selecting the wrong strategy has a more important cost, as the system is allowed to change strategies less often. However, even in these situations of degraded performance, our solution still performs better than a fixed approach, either periodic or event-driven. By using the adaptive

message generation proposed in this section, sensors will reduce their energy consumption and may thus prolong the WARIM lifetime. Our adaptive message generation is also interesting from a networking perspective. Indeed, less message generation means also less contention and less collision.

Once a message is generated, it must be relayed through a multi-hop communication mechanism from the source to the sink (the light controller). Thus, in the next section, we describe the routing protocol designed to be used with the WARIM architecture and the related contributions.

7.3 Routing of a message from the sensor to the light controller

In the previous section, we proposed an adaptive messages generation strategy which reduces the number of messages generated in the network. Our goal in this section is to design the mechanism used by nodes to transmit a message from the source node to the sink.

7.3.1 An overview of routing

Two layers are particularly concerned by the transmission and forwarding of a message from the source device to the destination device: the MAC and the routing layers. The goal of the MAC layer is to provide mechanisms which not only allow to sensors to share the wireless medium, but also to solve the collision problem. Some communications protocols for LWSN are proposed in [141, 142, 143, 144]. In these papers, authors exploit the linear structure of the network to improve the performance (in terms of network throughput) of the communication protocol. In the remaining of this section, we focus on routing protocols.

In WSN, because of the limited communication range of sensors, a sensor might not be able to establish a direct communication link with the sink. Thus, it uses relay sensors to forward its messages in multi-hop to the sink, as proposed in WARIM. In such a network, the routing is then a critical task executed by nodes. The routing can be defined as the mechanism through which paths are built in a network to carry the data from its source to the sink.

The selection of the link to a neighbor S_j to forward a message might be done based on metrics like the number of hops from S_j to the sink, the distance between S_j and the sink, the remaining energy of S_j or the transmission delay on the link $S_i \rightarrow S_j$ [145]. Because of the environment, the limited energy of nodes or the interference from the surrounding networks, the wireless links in low power WSN are by nature unstable [72, 73]. So, at a given time, some links may be available, and become unavailable a few moments later: they are said to be opportunistic. In traditional routing protocols (i.e. routing protocols in which sensors construct and maintain routing tables), if a link of better quality than the one predefined in the routing table appears, the predefined link will still be used, even if it is unavailable (because of the environment, the neighbor has failed or is not active). If a new message is sent on that (predefined) link, there might be unnecessary energy consumption and moreover, the message might not be delivered to the final destination. The wireless channel used in WSN is by nature a broadcast one. Thus,

when a node sends a message, all the active sensors within its communication range receive it: it is the overhearing phenomenon. A class of routing protocols, named opportunistic routing [141], which takes advantage of the broadcasting nature of the wireless channel and which exploits opportunistic links, has been proposed for WSN.

By using opportunistic routing, each packet is allowed to dynamically build the route toward the destination. This is done according to the condition of the wireless links at the moment when the packet is being transmitted. In opportunistic routing, the nodes do not select a next-hop before the transmission starts. The neighbors that receive the transmitted packet use a predefined mechanism to decide which of them must forward the packet and which must discard it. This process is usually called candidate coordination [141]. In other words, a sender broadcasts the packet first, and one of its neighbors that has received the packet will continue the forwarding process. Therefore, the opportunity of delivering the packet to the destination is increased.

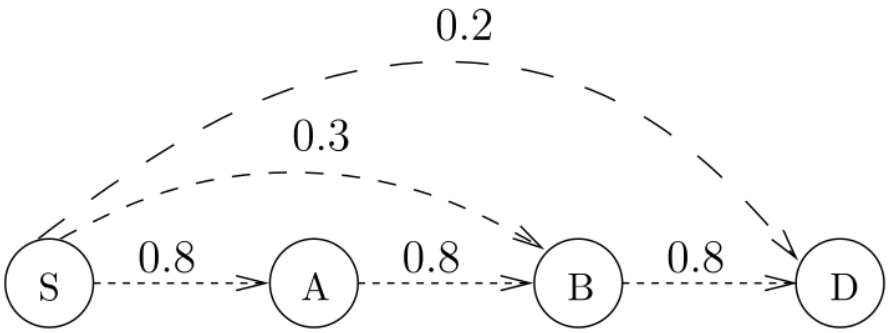


Figure 7.5: An example of opportunistic routing.

Another benefit of the opportunistic routing is the preservation of network resources. Consider the network of Figure 7.5, in which the link weight represents the delivery probability. If a traditional routing protocol is applied on this network, the best path from source S to destination D is $S \rightarrow A \rightarrow B \rightarrow D$. In this case, when the source sends a packet, even though node B correctly receives the packet transmitted from the source, it cannot forward it. This packet has to be resent until it reaches the predefined next-hop A. Therefore, as we can see, in this kind of routing scheme, many retransmissions and wasted network resources may occur. Another situation that might occur is that the packet may be received by both nodes A and B. Because the distance of B to the destination is less than the one of node A, the packet might be delivered to the destination with a smaller number of transmissions if node B forwards it. Because source S has not selected node B as its next-hop forwarder to destination D, node B is not allowed to forward the received packet. By contrast, opportunistic routing can increase the progress of a packet toward the destination by taking advantage of these situations.

In this section, we give a simplified view of traditional and opportunistic routing protocols for WSN. More information can be found in [141] which reviews opportunistic routing protocols proposed for WSN. In the next section, we describe a simple gradient-based opportunistic routing protocol implemented in our simulator to evaluate WARIM.

7.3.2 A gradient-based opportunistic routing protocol for Warim

Gradient-based routing approach [146] is inspired from position-based routing but did not use sensor location, but the concept of gradient. The gradient of a sensor can be defined as the number of hop from the destination (sink). In this section, we use the sensor rank as gradient.

When a sensor has a message to transmit to the sink, it broadcasts a message containing its rank (r_rank) and a sequence number. These two informations contained in the message header allow to uniquely identify the message in the network. The message header also contains the rank of the previous relay sensor (r_p). When a sensor S_i with rank r_i receives a message, if S_i is closer to the sink than the previous relay sensor ($r_i < r_p$), it calculates the value $shift = r_p - r_i$. This value represents the number of sensors between the relay sensor and S_i, S_i included. If $shift$ is less than the constant max_shift, S_i initializes a timer to $(max\_shift - shift) \times slot\_duration$. The sensor S_i listens the channel until the end of its timer or when it hears the forwarding of the message by an other sensor. S_i forwards the message only if it does not hear the forwarding of this message by one of its neighbors. If S_i is the sink (S_0), just after the reception of the message from the previous relay sensor, it sends an acknowledgement in broadcast to inform the previous relay sensor and other candidates that it has received the message. If after $(max\_shift + 1) \times slot\_duration$ the previous relay sensor does not hear the retransmission of the message by one of its neighbors, nor the acknowledgement from the sink, it considers that no neighbors have received the message and retransmits it. We assume that, at each hop, the current relay sensor can do max_rtx retransmissions. The routing process ends when the last relay receives the acknowledgement from the final destination or when, after max_rtx transmissions, none of the candidate forwarders succeeds to forward the message. The parameters of this routing strategy are summarized in Table 7.1.

max_shift	The maximum number of hops between the relay node and a forwarding candidate.
max_rtx	The maximum number of retransmissions at each hop.
$slot\_duration$	The slot duration, its value will be defined taking into account applications and network constraints.

Table 7.1: The parameters of the implemented gradient-based opportunistic routing protocol.

7.4 Performance evaluation

Using a network simulator for WSN [147], we evaluate the capability of Warim to collect vehicular traffic at an intersection. In the following section, we describe our simulation methodology.

7.4.1 Simulation methodology

For our simulations, we use the WSNet simulator [147]. It is an event-driven network simulator for large scale WSN. It already integrates many modules at the physical, MAC, routing and application layers. Its modular architecture allows to easily develop and add new modules.

Thus, we developed for this simulator a module implementing the routing protocol described in Section 7.3.2. However, for application, MAC and physical layers, we used modules already integrated in the simulator.

7.4.1.1 Application layer

In the previous section, we proposed an adaptive message generation strategy which reduces the total number of messages in the network. Initially, an event-based messages generation strategy is used by sensors: i.e., a node generates one message each time it detects a vehicle. But, when the vehicular traffic increases and becomes greater than a threshold, sensors periodically generate and send a message containing the number of vehicles detected. Thus, the periodic approach corresponds to the situation of high vehicular traffic, high contention and probably high collision. Therefore, in the application layer used for this evaluation, a periodic message generation approach is used. We varied the period (φ) of messages generation from 2 s to 60 s.

7.4.1.2 Routing and MAC layers

At the routing layer, we implemented the routing protocol described in Section 7.3.2. In our simulation, we set max_shift, the maximum number of nodes between a forwarding candidate and the message sender, to 5. Results presented in this chapter is for a *slot_duration* of 50 ms. Simulations are conducted for different values of max_rtx (the maximum number of retransmission at each hop). For the MAC layer, we used the $CSMA/CA$ protocol proposed in the IEEE 802.15.4 standard [39]. This protocol is already integrated in the WSNet simulator. We used the default parameters values proposed in the standard. Since all messages are sent in broadcast mode, there is no retransmission at the MAC layer. But, the max_rtx parameter used by our routing protocol improves the transmission reliability at each hop.

7.4.1.3 Physical layer

For the physical layer, we still consider the IEEE 802.15.4 standard. We consider an Offset Quadrature Phase-Shift Keying (OQPSK) modulation scheme, an orthogonal interference model and the communication frequency 2400 Mhz. Many signal propagation models proposed for wireless network are implemented in WSNet. We consider the log-normal shadowing model widely used in the evaluation of WSN protocols [84]. The main parameters of this model is the path-lost exponent (ψ) and the standard deviation (ϕ). For links at ground level, like in WARIM, we estimated the values of these parameters by simulations. The idea is to estimate using simulations (we used a python program wrote during this thesis), the values of ϕ and ϕ which allow to have links with PRR close to the empirical PRR collected during our experiments. Simulation results give $\psi = 5$ and $\phi = 15$. Combining the log-normal shadowing propagation model with the orthogonal interference model and the OQPSK modulation, we configure the physical layer in our simulation.

7.4.1.4 Deployment and nodes coordinates

Recall that our goal in this chapter is to put together all our contributions and evaluate the performance of WARIM in terms of data collection. We take into account the link properties at ground level (See Chapter 3), the deployment proposed in Chapter 5 and the sensors ranks obtained by the centroid-based algorithm proposed in Chapter 6. We consider a linear network of length $100m$. We assume that the network is organized to form 10 virtual nodes. We also assume that $p = 0.5$ (50% of messages generated by virtual node B_i are relayed by B_{i+2} and the others 50% by B_{i+1}) and $\alpha = 0.2$ (each sensor overhears 20% of communications addressed to a different sensor in its neighborhood). Using the greedy algorithm proposed in Chapter 5, we deploy 20 sensors in the network. The routing protocol described in Section 7.3.2, implemented and integrated in the simulator, uses sensors ranks (obtained by the algorithm proposed in Chapter 6) as gradient. Results are shown for real sensors ranks and also for the sensors ranks obtained by our ranking algorithm. Two sets of simulations are then conducted. In the first set, the real ranks of sensors are used as gradient while in the second set, the ranks obtained by our ranking algorithm are used. In the remaining of this chapter, the ranks obtained by our centroid-based algorithm are named estimated ranks. Figure 7.6 shows the sensors deployment.

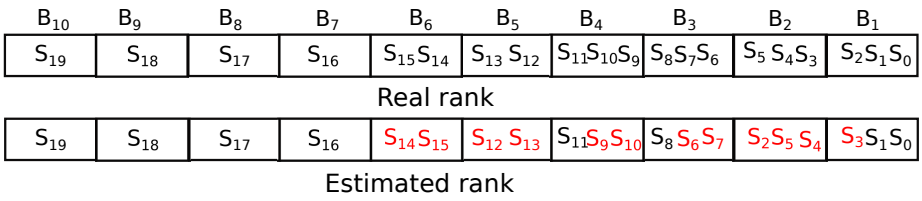


Figure 7.6: Sensors deployment: physical topology (with real rank) and virtual topology (estimated rank). Sensors colored in red are those at bad positions.

7.4.1.5 Evaluation metrics

In our evaluation, we consider three metrics: the PDR (Packet Delivery Ratio), the average traffic (transmissions and receptions) per node and the sink reception cycle. The PDR is the ratio of the number of messages received by the sink to the total number of messages generated by all nodes in the network. It is the main evaluation metric, since it allows to measure the capability of WARIM to effectively collect data at the sink level. The traffic of a node is the number of messages transmitted and received. It is obvious that, when the distance to the sink is reduced, the traffic load of nodes is higher. In the results presented in the next section, we present the average traffic load per node in the network. This gives an indication on the energy consumption of nodes since radio is the component which consumes the most sensor energy for messages transmission and reception. The traffic light controller uses the information provided by WARIM to periodically change the light plans in order to take into account the changes in vehicular traffic at the intersection. Thus, the final metric used in our evaluation, the sink reception cycle, can be considered as an application metric: it is the delay between two receptions at the sink from a given source. This allows to know the period of messages update at the light

controller and the possible impacts on the intelligent traffic light management system. Results presented in Section 7.4.2 are the average over 1000 executions. A summary of the simulation configuration is presented in Table 7.2.

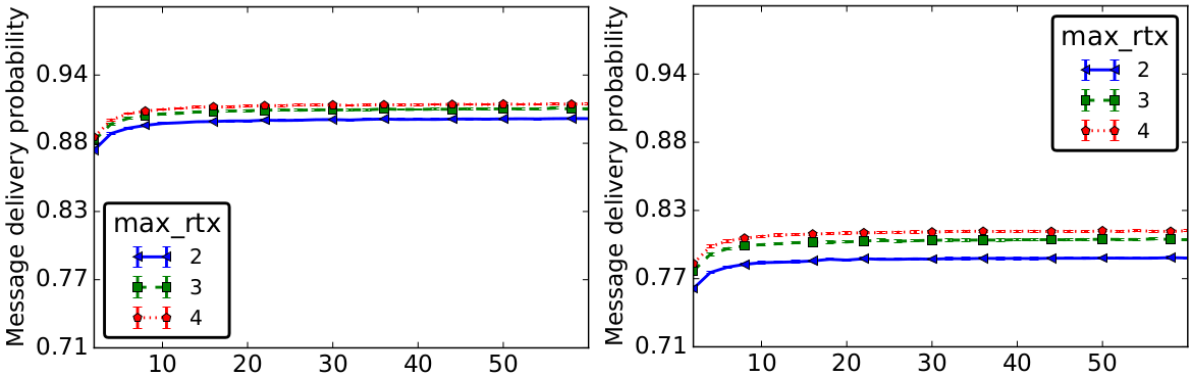
Parameter	**Description**	**Value**
Application layer (periodic messages generation)		
φ	Messages transmission period	$2s - 60s$
M	Number of messages sent by each sensor	100
N	Number of sensors	20
Routing layer (opportunistic routing)		
max_shift	The maximum number of hops between the relay node node a forwarding candidate	5
max_rtx	The maximum number of retransmission at each hop	2, 3 and 4
$slot\_duration$	The slot duration, its value will be defined taking into account applications and network constraints	$50ms$
MAC layer (CSMA/CA)		
IEEE 802.15.4 MAC layer	Parameters values are the default values defined in the standard	–
Physical layer (802.15.4)		
Modulation	The modulation scheme	$OPQSK$
Interference	The interference model	Orthogonal
Propagation	The signal propagation model, Log-Normal Shadowing	$\psi = 5$ and $\phi = 15$

Table 7.2: Simulation parameters.

7.4.2 Simulation results

7.4.2.1 Message delivery ratio

Figure 7.7 presents the PDR as a function of the message transmission period. The PDR is given when the real ranks of sensors are used as gradient (Figure 7.7a) and when the estimated ranks (ranks obtained by the centroid-based algorithm proposed in Chapter 6) are used (Figure 7.7b). This figure shows that, when the messages transmission period (the duration between two packet generations) increases, the PDR increases too. As function of the number of retransmissions at each hop, the PDR is between 88% and 92% when real ranks are used, and between 77% and 80% when estimated ranks are used. When the message transmission period is low, we observe the lowest PDR. Indeed, by reducing the duration between two message generations, the number of messages generated in the network per time unit increases. A direct consequence is an increase of the contention at the MAC layer, which increases the probability of collisions and packets loss. But when the transmission period increases, the PDR increases too, and from a period of 15 s, it remains more or less constant. Indeed, all the mechanisms implemented at the MAC layer and the retransmission mechanism implemented by the opportunistic routing protocol introduce a maximum message transmission delay at each hop. A transmission period greater than this delay does not have an impact on the PDR.



(a) Real nodes rank

(b) Estimated nodes rank

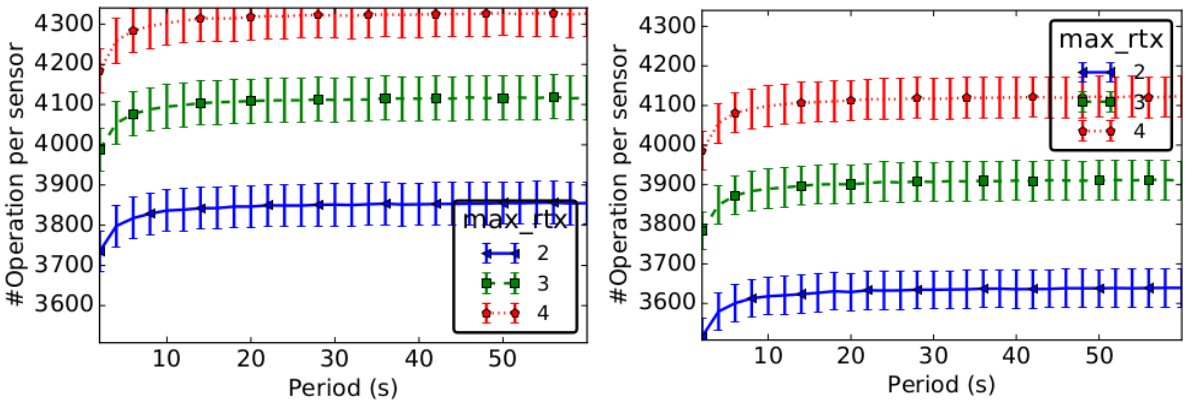
Figure 7.7: Packet delivery ratio.

A particular observation in this figure is the better PDR obtained with real ranks compared to the PDR obtained with estimated ranks. Indeed, while we can obtain a PDR of 92% with real ranks, the maximum PDR obtained with the estimated ranks is 80%. We guess that this is due to the ranking error introduced by our ranking algorithm. Indeed, because of a ranking error, a node physically far from the sink can be selected as a forwarder candidate. Moreover, a node closer to the light controller can receive and ignore a message because its estimated rank places it far than the current relay. These ranking errors may conduct to message loss or may introduce delay in the transmission as we will show in Section 7.4.2.3. But we think that the PDR can be improved by using network mechanisms. A solution is to increase the value of max_rtx. As shown in Figure 7.7, the PDR increases with the value of max_rtx. Human intervention for sensors ranking might give perfect ranks. But, this operation is costly and time consuming.

By adding reliability at the application layer, information contained in the lost messages can be received by the controller, but with a delay. Indeed, the information contained in a message is usually an integer which may be coded on ω bytes (for many applications, may be 2 or 4 bytes). We can assume that each time a node transmits a message, the data inserted in the message corresponds not only to the data measured during the current period (event), but to the data measured for the last periods (events). Thus, the size of the data contained in the message will be κ bytes if we assume that is the size of the data generated in one period (or for an event). In this case, an interesting metric to investigate will be the data latency as the elapsed time between a data generation and its reception by the sink.

7.4.2.2 PDR and traffic load per node

A high value of max_rtx increases the traffic in the network as shown in Figure 7.8. Thus, increasing the maximum number of retransmissions at each hop increases the PDR, while consuming the energy of sensors (since radio consumes the most energy of sensors for messages transmission and reception). How to guarantee a good PDR while minimizing the sensor energy consumption? This is a trade-off widely studied in the field of WSN. For this purpose, energy-efficient physical, MAC, routing and transport layer have been designed for WSN [98]. But in this chapter, our



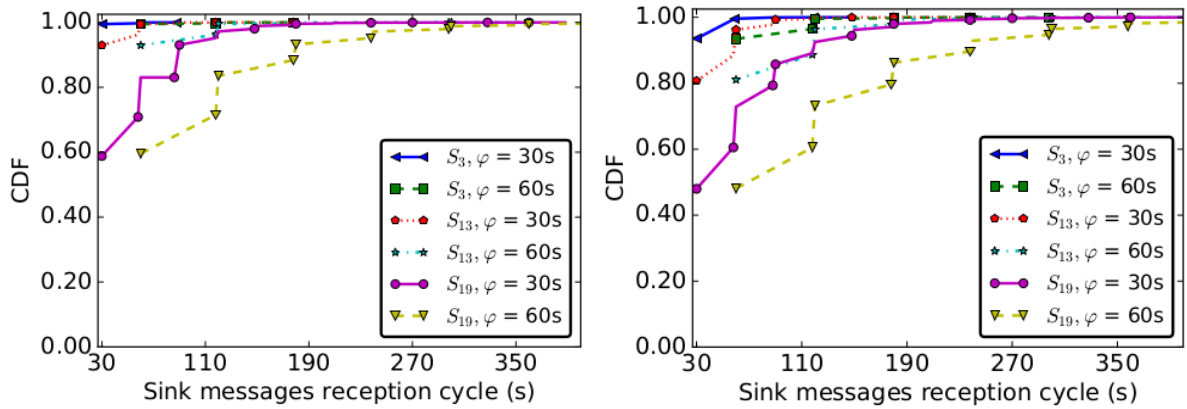
(a) Real nodes rank (b) Estimated nodes rank

Figure 7.8: Average traffic load per node.

goal is not to study all these solutions or to propose new ones. As we previously claimed, the goal of this chapter is to evaluate the capability of WARIM to collect vehicular traffic data at an intersection.

Figure 7.8 also shows that, when estimated ranks are used as gradient, the average traffic load per sensor is relatively smaller compared to the case where real ranks are used. This result is coherent with the one presented by Figure 7.7, which shows that the PDR is smaller when the estimated ranks are used. Indeed, each lost message reduces the total number of messages circulating in the network and then the traffic per nodes.

7.4.2.3 Sink messages reception cycle



(a) Real sensors rank (b) Estimated sensors rank

Figure 7.9: CDF of the duration between two message receptions of the sink from the source S_3 (in the virtual node B_2) S_{13} (in virtual node B_5) and S_{19} (in the virtual node B_{10}).

The goal of WARIM is to collect data concerning the vehicular traffic at the intersection. This data allows the light controller to periodically (using a predefined period) update the light plans. Thus, it is important to study or to show the period at which WARIM will provide up to date information concerning a particular area of the intersection. In this section, we present

in Figure 7.9 the CDF of the sink reception cycle for some sources: S_3 of virtual node B_2, S_{13} of virtual node B_5 and S_{19} of virtual node B_{10} (B_{10} being the last virtual node since, as we explained in Section 7.4.1.4, we consider 10 virtual nodes). As in the previous sections, we assume $slot\_duration = 50$ ms and we consider the maximum number of retransmissions at each hop which gives the best PDR ($max\_rtx = 4$). Results are shown for real ranks (Figure 7.9a) and for estimated ranks (Figure 7.9b). In each case, results are shown for messages transmission period of 30 s and 60 s.

An important observation in these figures is that the sink reception cycle is higher when the estimated ranks are used compared to the case where reals ones are used. Indeed, considering sensor S_{19} as source and $\varphi = 30$ s, we have a probability of 70% to have a sink reception cycle less than 60 s when real sensors rank are used and 60% when estimated ranks are used. This observation is coherent with the results presented in the previous sections (Figure 7.7). Indeed, the lower PDR observed, when considering estimated ranks, is a sign that there are more retransmissions at each hop, compared to the case where real ranks are used. These retransmissions, more numerous in the case of estimated ranks, which may vary from one message to another, is a source of the variations of the sink messages reception cycle from the same source.

For source node S_3 of the virtual node B_2, we observe that the sink reception cycle is almost equal to the message transmission period when real ranks are used (see Figure 7.9a). In the case of the estimated ranks, we observe a probability greater than 90% to obtain a sink messages reception cycle equal to the messages transmission period (see Figure 7.9b). This is expected since sensors in the first virtual nodes have a high probability to communicate directly with the sink (S_1). Thus, messages from these sensors are not subjected to delay introduced by multi-hop routing. For a far source, messages transmitted by this source are relayed through multi-hop until the sink. The delay introduced at all hops may be different from one message to another. Results presented in this figure also show that the farther a node S_i is from the sink, the higher is the maximum reception cycle of sink from S_i.

In summary, results presented in Figure 7.9 show that the reception cycle of the sink for messages coming from the same source depends on the position of the source. Indeed, when the source is far from the sink, its messages are forwarded through many hops until the sink. At each hop, a delay is introduced. This delay depends on the processing capability of each node and the physical layer (data rate, environment constraints, collisions, etc). All these factors may influence the period at which the light controller updates the light plans and must be taken into account when designing communication protocols for WARIM.

7.5 Conclusion

The goal of this chapter was to put together all our contributions and evaluate the performance of WARIM, the WSN-based architecture proposed in this thesis (Chapter 4) for vehicular traffic monitoring at an intersection. In order to know which is the most energy-efficient message

generation strategy, we start by analyzing a dataset of vehicular traffic of the city of Koln in Germany. We proposed an adaptive message generation strategy which reduces by 40% the number of messages generated in the network compared to a static (event-based or periodic) approach. This messages generation strategy switches between event-based (a message is generated each time a vehicle is detected) and periodic messages generation as function of vehicular traffic at the intersection.

Considering the deployment proposed in Chapter 5, the ranking proposed in Chapter 6 as gradient and a simple gradient-based opportunistic routing protocol described in this chapter, we evaluated the capability of WARIM to collect data at an intersection. In the network simulator used, we consider the CSMA/CA MAC protocol of the 802.15.4 standard. We also considered a realistic physical layer with a channel model close to the one at the ground level like in WARIM.

Results show that, by adding network mechanisms like message retransmission at each hop, even if the traffic load of nodes increases, almost 92% of messages generated by nodes are delivered to the sink when real nodes ranks are used and 80% when the rank obtained by the centroid-based algorithm proposed are used. We also look at the sink reception cycle, which is the duration between two receptions by the sink for messages coming from the same source. Results show a high probability to have a reception cycle equal to the messages transmission period. In summary, results presented in this chapter show that, WARIM, combined with the greedy deployment proposed in Chapter 5 and sensors ranks obtained by our centroid-based ranking algorithm proposed in Chapter 6 can be used for a reliable and real time vehicular traffic monitoring at an intersection.

Chapter 8

Conclusion and perspectives

8.1 Remainder of goals and challenges

Intelligent transportation systems [3] represent the introduction of information and communication technologies in the management of transport issues. They appeared in a context marked by a significant growth in the number of vehicles in major cities worldwide causing traffic jams. Some consequences of traffic jams are: excessive fuel consumption, environmental pollution or accidents. Since most of the traffic jams are observed at road intersections, adaptive traffic lights management was one of the the first services offered by intelligent transportation systems. The high cost of the first solutions implementing this service limited their deployment only to some cities of developed countries. Thanks to the low-cost of sensor nodes, many WSN solutions for vehicular traffic monitoring have been proposed.

At the beginning of this thesis, our goal was to propose a low-cost, realistic, lightweight and autonomous WSN architecture for vehicular traffic monitoring at an intersection. We wanted to propose a WSN architecture which may be adopted by many cities, event those limited by the financial budget; which is realistic in terms of network communication constraints; which can be rapidly deployed by an operator without a constraint technical training; which is autonomous in terms of network configuration and organization in order to reduce the deployment time and maintenance effort; and which can report data at lane level (the number of vehicles per lane) in order to allow an optimal and real-time traffic lights management.

8.2 Contributions

Two types of sensors for vehicles detection are proposed by industrials: underground sensor and surface-mounted sensor. Most WSN-based architectures proposed by researchers for vehicular traffic monitoring at intersections are unrealistic in terms of communication requirements [8, 10, 11]. Thus, our **first contribution** in this thesis is the **links characterization in a WSN with nodes deployed at ground level**. We evaluated the performance of a wireless link as function of many parameters like the proximity of the ground surface, the road topography, the message size, the radio transmission power and communication frequency. We expressed

the link quality in terms of PRR and RSSI. Results showed a poor link quality at ground surface compared to a deployment at a given height. At ground level, we observed a maximum communication range of 10-15 m with acceptable link quality. This communication range is very far from the 100m proposed by the IEEE 802.15.4 standard and widely adopted by all existing WSN-based architectures. **The second contribution** of this thesis is **Warim**, **a realistic WSN architecture for vehicular traffic monitoring at intersections**. This architecture is proposed based on the experimental results obtained during the first part of this thesis. In this architecture, many nodes are deployed on the road surface per lane and a node deployed on the light support is considered as the sink. The number of nodes per lane and the distance between two nodes depend on the application and the environment constraints, as well as on the radio module properties. Since experimental results showed a limited communication range with nodes deployed at ground level, in Warim nodes use a multi-hop communication model to forward a message from any node to the sink.

The **third contribution** of this thesis is an **energy efficient and simple nodes deployment strategy for LWSN using the virtual nodes concept**. A virtual node is formed by sensors deployed in a given area. We proposed a model of the average communication traffic per virtual node. This traffic includes messages transmission and reception. It is important to note that receptions mentioned here include those due to the overhearing phenomenon. We then formalized the sensors deployment problem as a mixed-integer non linear programming problem and proposed a greedy algorithm to compute the number of sensors that form each virtual node. Performance evaluation showed that, compared to a uniform deployment (i.e. when the same number of sensors is deployed in each virtual node), our virtual nodes-based approach improves the network lifetime by up to 40%. Moreover, our virtual nodes-based approach outperforms the distance-based one when a good scheduling algorithm, which reduces the overhearing phenomenon, is used. We also studied an alternative solution which considers a network in which the batteries of nodes are nonuniform; i.e the battery of a node is proportional to its traffic. Results showed that, by properly selecting the battery to assign to each node, such an approach can significantly improve the network lifetime.

The **fourth contribution** of this thesis is **an iterative, autonomous and centroid-based nodes ranking algorithm for LWSN**. The centroid of a node is the average of its neighbors coordinates. At each iteration, all the processing done by a selected node depends on its neighborhood size and not on the number of nodes in the entire network. Thus, our ranking algorithm is scalable and lightweight considering the network deployment and maintenance. In this algorithm, the single parameter manually configured is the first node of the network. Using a simulator developed during this thesis, we evaluated the performance of our ranking algorithm. We evaluated its performance considering various nodes degree, realistic channel models (built from empirical PRR values collected from experiments with nodes deployed at ground level) and different deployments (uniform, random and the virtual nodes-based deployment proposed in this thesis). Results showed a node ranking ratio greater than 95%. In case of ranking error, results showed a high probability to have an error of one hop. Results also showed that our ranking

algorithm outperforms the common-neighbors approach widely used in the literature to estimate the proximity between nodes. Our ranking algorithm showed not only its efficiency in terms of nodes ranking in a LWSN, but also its robustness to deployment errors.

Finally, we **validated Warim using the WSNet simulator**. The analysis of the vehicular traffic at intersection using a dataset of the city of Koln in Germany showed a temporal and spatial variation in vehicular traffic. In order the reduce the number of messages in the network and then the node energy consumption, we proposed an adaptive message generation strategy. The proposed strategy reduces by 40% the number of messages generated in the network compared to a static (event-based or periodic) one. To evaluate Warim, we implemented in the simulator a simple gradient-based opportunistic routing protocol using sensors rank as gradient, a physical layer considering the link properties at ground level and we used the MAC layer of the 802.15.4 standard. We also considered the virtual nodes-based deployment proposed in this thesis. Results showed a PDR of 85%. We also investigated the controller reception cycle which is defined as the time between two receptions from the same sensor. Results showed a high probability (from 50% to 90% depending on the sensor position) to have a reception cycle equal to the message generation period. In summary, simulation results showed that putting together all our contributions, Warim, the WSN-based architecture proposed in this thesis can be used for a reliable and real-time vehicular traffic monitoring at intersection.

8.3 Perspectives

8.3.1 The case of intersections with complex topologies

In this thesis, we considered sensors deployed on a single lane to form a LWSN. A typical intersection may have several entrances and exits depending on its topology, and each entrance (or exit) may have many lanes. Such an intersection presents new challenges. Indeed, in such a configuration, communication between sensors deployed at different entrances (or exits) may interfere. Even on a road segment with multiple lanes, a sensor might receive or forward messages from other lanes. This should be take into account when modeling the traffic per sensor and then the deployment strategy. If we consider the sensors ranking problem, the case of multi-lane road requires first the identification of the lane on which each sensor is deployed, and second the rank of each sensor on that lane. Thus, proposing a solution to this problem, which does not require the human intervention is an interesting challenge.

8.3.2 LWSN

In this thesis, we proposed a simple and energy efficient deployment strategy for WSN with linear topology. In the deployment proposed, the number of sensors increases as the distance to the sink is reduced. This deployment allows to address the problem of the high traffic observed in the area close to the sink. The high number of sensors in the neighborhood of the sink, combined with the high data traffic in this area, increases the contention at the MAC layer and the collision probability. Thus, it would be interesting to investigate communication protocols considering a

network with asymmetric data traffic and sensor density. These communication protocols could exploit the linear property of the network to improve the access to the channel and then to reduce the collision probability and the overhearing. LWSN can also be deployed for applications like border control and infrastructures monitoring. Considering such large networks, the number of messages circulating in the network could be very huge. Taking this into account, communication protocols, sensors organization (hierarchical network) and data aggregation mechanisms could be investigate in order to improve the network efficiency.

8.3.3 LPWAN

LPWAN is a more recent communication technology used in WSN, which allows long range and low power communications. Using such a communication technology in WARIM, each sensor will directly communicate with the light controller. Thus, considering the poor wireless link quality at ground level and the limited communication range which is addressed by a routing protocol, LPWAN technology might be a good communication technology choice. One of the main goal of WARIM is the real time vehicular traffic monitoring at intersection for an intelligent traffic lights management. During some periods of the day characterized by a high vehicular traffic, the number of message can be very high. One drawback of LPWAN technology is the low throughput and the high transmission delay. Thus, an important challenge would be to investigate how such a communication technology could be used in a network with a high data traffic considering an application requiring real time data transmission.

8.3.4 Energy Harvesting

Generally speaking, wireless sensor nodes are battery-powered. When the battery is exhausted, it can be replaced if the deployment area is accessible. Recently, with the evolution of energy harvesting technologies, and particularly solar energy, small size solar cells have been proposed for WSN [61, 62].

In WARIM, to reduce the maintenance cost, sensor nodes, which take benefit from renewable energy sources, may be used. Energy-aware networking protocols are generally used in battery-powered WSN to manage efficiently the energy consumption of nodes. In energy harvesting WSN, the network lifetime is theoretically infinite, and it is only limited by hardware lifetime. Then, the main considerations regarding energy should be *i)* to schedule the data transmission according to the energy available and to anticipate the energy depletion using a data traffic schedule, *ii)* to maximize throughput and *iii)* to guarantee perpetual network functioning, i.e. allow the network to function even in the lack of harvesting sources for a long time.

Considering sensors equipped with solar panels, the energy produced by a solar panel depends on its size, the weather, the period of the day, the area where the panel is deployed and the availability of sunshine radiation. If nodes in WARIM have energy harvesting capabilities, since vehicles stop at the intersection when the light is red, sensors close to the traffic light are shadowed much of the time by vehicles and harvest less energy compared to those which are far. Moreover,

such sensors have to relay control packets sent by other nodes. Clearly, this means that the power consumption of such nodes has to be studied finely, by taking into account data traffic intensity and the available energy sources. Transmission scheduling, topology management, and dynamic routing protocols should be designed by considering renewable energy cycles.

8.3.5 Other perspectives

An important observation of the experiment results presented in this thesis is the poor link quality when the message size is too small (2 bytes). So, a perspective to this thesis is to continue investigation in order to understand how can this be explained. Using statistical tests, we also compared the empirical PRR collected during our experiments campaign to the PRR values generated by the log-normal shadowing model widely used by simulator when evaluating communication protocols for WSN. The Mann-Whitney-Wilcoxon test validated H_0 in the case of the flat area deployment, but rejected it for the hill deployment. This indicates that the poor link quality on the hill deployment is the consequence of a phenomenon that is not accounted for in the theoretical log-normal shadowing model. Thus, understanding what happens when nodes are deployed on a hill could be an interesting challenge to investigate in order to improve the log-normal shadowing propagation model.

On modern sensor platforms, nodes have a set of discrete power levels, which can allow them to transmit data for different distances. It might be interesting to investigate deployment strategies which not only give the number of sensors to deploy in each virtual nodes, but also the transmission power used by each sensor in order to reduce the energy consumption and guarantee the network connectivity.

In this thesis, we proposed a new centroid-based algorithm for nodes ranking in a LWSN. Concerning the ranking ratio, the results obtained are interesting. Nevertheless, in some cases, we observe ranking errors, i.e. the rank calculated by our ranking algorithm for a node does not corresponds to its real rank. In this thesis, we argued that the nodes ranks might be useful not only for some network protocols and mechanisms like routing and data aggregation, but also for the target application (the application for which the LWSN is designed). Concerning the routing, we observed a PDR of 80% when estimated ranks are used, compared to 92% when real rank are used. To recover information contained in the lost packet, a simple solution is to integrate application reliability. Indeed, the information contained in a packet is usually an integer which may be coded on ω bytes. With such a mechanism, each time a node transmits a packet, the data inserted in the packet corresponds not only to the data measured during the current period (event), but to the data measured for the last periods (events). Thus, the size of the data contained in the packet will be κ bytes if we assume that ω is the size of the data generated in one period (or for an event). An interesting perspective would be to evaluate the impact of such a mechanism on the message latency. Indeed, if this latency can have a significant impact on the adaptive traffic lights control system, the impact might be less pronounced on other services or applications (like routes guidance) that can be offered using the collected data. At the application level, ranking error might also have a negative impact. Indeed, concerning the

vehicular traffic monitoring at an intersection, the node rank is used by the light controller to identified exactly at which location of the intersection a vehicle or an incident have been detected. Thus, evaluating and analyzing the impact of the ranking error on the target application remain interesting and opened questions.

Finally, our future work will be the implementation of WARIM in the physical world. The objective will be to implement our contributions on physical sensors deployed on road. We will consider the deployment strategy proposed, the centroid-based ranking algorithm proposed and a simple opportunistic routing algorithm to evaluate the capability of WARIM to collect vehicular traffic data at an intersection.

Bibliography

[1] United Nation., "https://esa.un.org/unpd/wpp/publications/files/wpp2017_keyfindings.pdf", Visited on January 30th, 2019

[2] Texas Transport Institute."Urban Mobility Scorecard 2015: https://mobility.tamu.edu/ums/report/", Visited on January 30th, 2019

[3] Wikipedia, "https://en.wikipedia.org/wiki/Intelligent transportation system", Visited on June 7th, 2019

[4] Mobilité intelligente, "https://www.mobilite-intelligente.com/ressources/technologies/capteurs", Visited on August 23rd, 2019

[5] F. Losilla, A.-J. Garcia-Sanchez, F. Garcia-Sanchez, J. Garcia-Haro, and Z. Haas, "A Comprehensive Approach to WSN-based ITS Applications: A Survey", In *Sensors*, 11(11): 10220–65, October 2011.

[6] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "Wireless sensor networks: a survey", Computer networks, 2002, 38(4), pp.393-422.

[7] P. Rawat, K.D. Singh, H. Chaouchi and J.M. Bonnin, "Wireless sensor networks: a survey on recent developments and potential synergies", The Journal of supercomputing, Springer, 2014, 68(1), pp.1-48.

[8] S. Faye, C. Chaudet, and I. Demeure, "A Distributed Algorithm for Single Intersections Adaptive Traffic Lights Control using a Wireless Sensor Networks", In *IEEE ITSC 2012*, Anchorage, AK, USA, September 2012.

[9] S. Faye, C. Chaudet, and I. Demeure, "A distributed algorithm for multiple intersections adaptive traffic lights control using a wireless sensor networks", In Proceedings of the first ACM workshop on Urban networking, pp. 13-18, Nice, France, December 2012.

[10] B. Zhou, J. Cao, X. Zeng, and H. Wu, "Adaptive Traffic Light Control in Wireless Sensor Network-Based Intelligent Transportation System", In *IEEE VTC Fall 2010*, Ottawa, ON, Canada, September 2010.

[11] M. Tubaishat, Q. Qi, Y. Shang, and H. Shi, "Wireless Sensor-based Traffic Light Control", In *IEEE CCNC 2008*, Las Vegas, NV, USA, January 2008.

[12] Sensys Networks, "www.sensysnetworks.com", Visited on February 12th, 2019

[13] L. Mimbela, L. Klein, P. Kent, J. Hamrick, K. Luces, and S. Herrera, "Summary of Vehicle Detection and Surveillance Technologies Used in Intelligent Transportation Systems", In *Federal Highway Administration's Intelligent Transportation Systems Program Office*, August 2007.

[14] L. A. Klein, M. K. Mills, and D. R.P. Gibson, "Traffic Detector Handbook", US Dept, Transp., Fed. Highway Admin., Washington, DC (2006).

[15] Wikipedia, "https://en.wikipedia.org/wiki/Magnetometer", Visited on April 9th, 2019

[16] G. De Angelis, A. De Angelis, V. Pasku., A. Moschitta and P. Carbone, "A simple magnetic signature vehicles detection and classification system for smart cities", In IEEE International Symposium on Systems Engineering (ISSE), (pp. 1-6), October 2016.

[17] P.T. Martin, Y. Feng, and X. Wang, "Detector Technology Evaluation, Department of Civil and Environmental Engineering", University of Utah Traffic Laboratory, Final Report, 2003.

[18] W. Birk, E. Osipov and J. Eliasson, "iRoadCooperative Road Infrastructure Systems for Driver Support", In Proceedings of the 16th ITS World Congress, Stockholm, Sweden, September 2009.

[19] H. Qin, Z. Li, Y. Wang, X. Lu, G. Wang and W. Zhang "An Integrated Network of Roadside Sensors and Vehicles for Driving Safety: Concept, Design and Experiments", In Proceedings of the 2010 IEEE PerCom, Manheim, Germany, 29 March - 2 April 2010.

[20] S. Yoo, P.K. Chong, T. Park, Y. Kim, D. Kim, C. Shin, K. Sung and H. Kim, "DGS: Driving Guidance System Based on Wireless Sensor Network", In Proceeding of the 22nd International Conference on Advanced Information Networking and Applications, Okinawa, Japan, 2528 March 2008; pp. 628-633

[21] S. Yoo, P. Chong and D. Kim, "S3: School zone safety system based on wireless sensor network", Sensors 2009, 9, pp. 5968-5988.

[22] T. Lin, H. Rivano and F. Le Moul, "A survey of smart parking solutions", IEEE Transactions on Intelligent Transportation Systems, 18(12), 2017, pp.3229-3253.

[23] T. Lin, Smart Parking: Network, Infrastructure and Urban Service, "http://theses.insa-lyon.fr/publication/2015ISAL0138/these.pdf", INSA Lyon, 2015, Visited on June 7th, 2019

[24] S. Faye, C. Chaudet, and I. Demeure, "Contrle du trafic routier urbain par un rseau fixe de capteurs sans fil", Rapport Technique, Février 2011.

[25] D.I. Robertson and R.D. Bretherton, "Optimizing Networks of Traffic Signals in real-time: the SCOOT method", IEEE Transactions on Vehicular Technology, 40(1), pages. 11-15, Febuary 1991.

[26] Traffic Research Laboratory, “https://trl.co.uk/”, Visited on January 30th, 2019

[27] A.G. Sims and K.W. Dobinson. “The Sydney coordinated adaptive traffic (SCAT) system philosophy and benefits”, IEEE Transactions on Vehicular Technology, 29(2), pages. 130-137, May 1980.

[28] SCAT, “http//:www.scats.com.au/product_base_packg_compnts.html”, Visited on January 30th, 2019

[29] L.C. Liao, “A Review of the Optimized Policies for Adaptive Control Strategy (OPAC)”, California PATH Program, Institute of Transportation Studies, University of California at Berkeley, 1998

[30] N.H. Gartner, F.J. Pooran, and C.M. Andrews. “Implementation of the OPAC adaptive control strategy in a traffic signal network”, In Proceedings of the IEEE Intelligent Transportation Systems, pages 195 200, August 2001

[31] C. Reggie and G. Chris, “Insync adaptive traffic signal technology : Real-time artificial intelligence delivering real-world results”, March 2012

[32] S.G. Shelby, D.M. Bullock, D. Gettman, R.S. Ghaman, Z.A. Sabra, and N. Soyke. “An overview and performance evaluation of ACS lite - a low cost adaptive signal control system”. In Transportation Research Board Annual Meeting, 2008.

[33] M. Selinger and L. Schmidt, “A review of the cost, maintenance and reliability of popular adaptive traffic control technologies”, In Adaptive Traffic Control Systems in the United States, September 2010.

[34] P. Mirchandani, and L. Head, “A Real-Time Traffic Signal Control System: Architecture, Algorithms, and Analysis”, In *Transportation Research Part C: Emerging Technologies*, 9(6): 415–432, December 2001.

[35] G. Guido, A. Vitale, F.F. Saccomanno, D.C. Festa, V. Astarita, D. Rogano and V. Gallelli, “Using smartphones as a tool to capture road traffic attributes.“ In Applied Mechanics and Materials, Vol. 432, pp. 513-519, 2013

[36] S. Axer, F. Pascucci and B. Friedrich, “Estimation of traffic signal timing data and total delay for urban intersections based on low-frequency floating car data“. Proceedings of the 6th Mobil.TUM, Munich, Germany, June 2015

[37] V. Astarita, V.P. Giofr and A. Vitale, “A cooperative intelligent transportation system for traffic light regulation based on mobile devices as floating car data (FCD)“. American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS), 19(1) , pp.166-177, 2016

[38] IEEE 802.11, ”http://www.ieee802.org/11/”, Visited on June 7th, 2019

[39] Institute of Electrical and Electronics Engineers, "IEEE Standard for Information Technology - Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks Specific Requirements Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs)", 1–670, October 2003.

[40] K. Mikhaylov, J. Petaejaejaervi and T. Haenninen, "Analysis of Capacity and Scalability of the LoRa Low Power Wide Area Network Technology", 22th European Wireless Conference, Oulu, Finland, 2016, pp. 1-6.

[41] SigFox, "https://www.sigfox.com/", Visited on February 15th, 2019

[42] A. Sharma, R. Chaki, and U. Bhattacharya, "Applications of Wireless Sensor Network in Intelligent Traffic System: A Review", In *ICECT 2011*, Kanyakumari, India, April 2011.

[43] K. Nellore and G.P. Hancke "A Survey on Urban Traffic Management System Using Wireless Sensor Networks", Sensors, vol. 16 (2), 2016

[44] M. Bernas, B. Paczek, W. Korski, P. Loska, J. Smya, P. Szymaa "A Survey and Comparison of Low-Cost Sensing Technologies for Road Traffic Monitoring", Sensors, 18(10):3243, 2018.

[45] X. Bao and H. Li and D. Xu and L. Jia and B. Ran and J. Rong "Traffic Vehicle Counting in Jam Flow Conditions Using Low-Cost and Energy-Efficient Wireless Magnetic Sensors", Sensors, vol. 16 (11), 2016

[46] T. S. Rappaport. "Wireless Communications: Principles and Practise." Prentice Hall, 1996.

[47] HiKoB, "www.hikob.com", Visited on February 12th, 2019

[48] S. Roy, R. Sen, S. Kulkarni, P. Kulkarni, B. Raman, and L.K. Singh, "Wireless Across Road: RF based Road Traffic Congestion Detection", In *COMSNETS 2011*, Bangalore, India, January 2011.

[49] R. Sen, A. Maurya, B. Raman, R. Mehta. R. Kalyanaraman, N. Vankadhara, S. Roy, and P. Sharma, "Kyun Queue: A Sensor Network System to Monitor Road Traffic Queues", In *ACM SenSys 2012*, Toronto, ON, Canada, November 2012.

[50] C. Day, H. Premachandra, T. Brennan, J. Sturdevant, and D. Bullock, "Operational Evaluation of Wireless Magnetometer Vehicle Detectors at Signalized Intersection", In *Transportation Research Record: Journal of the Transportation Research Board*, 2192(1): 11–23, February 2010.

[51] Libellium, "http://www.libelium.com/", Visited on February 12th, 2019

[52] H. Karl and A. Willig, "Protocols and Architectures for Wireless Sensor Networks", Wiley, 2006

[53] Crossbow, "http://www.willow.co.uk/TelosB_Datasheet.pdf", Visited on February 12th, 2019

[54] Tmote Sky Sensors. http://www.moteiv.com/products/tmotesky.php,2007, Visited on February 12th, 2019

[55] Contiki OS, "http://www.contiki-os.org/", Visited on February 12th, 2019

[56] Sing-Yiu Cheung and Pravin Varaiya, "Traffic Surveillance by Wireless Sensor Networks: Final Report for PATH TO 5301", University of California, Berkeley

[57] F. Stajano, N. Hoult, I. Wassell, P. Bennett, C. Middleton, K. Soga, "Smart Bridges, Smart Tunnels: Transforming Wireless Sensor Networks from Research Prototypes into Robust Engineering Infrastructure", *Ad Hoc Networks*, 8(8):872–888, Nov. 2010.

[58] W. Fisher, T. Camp, V. Krzhizhanovskaya, "Crack Detection in Earth Dam and Levee Passive Seismic Data Using Support Vector Machines", *Proc. ICCS 2016*, San Diego, CA, USA, Jun. 2016.

[59] Z. Sun, P. Wang, M. Vuran, M. Al-Rodhaan, A. Al-Dhelaan, I. Akyildiz, "BorderSense: Border Patrol through Advances Wireless Sensor Networks", *Ad Hoc Networks*, 9(3):468–477, May 2011.

[60] K. L. Mills, "A Brief Survey of Self-organization in Wireless Sensor Networks: Research Articles", In *IEEE Wireless Communications & Mobile Computing*, 7(7): 823–834, September 2007.

[61] A. Kansal, J. Hsu, M.B. Srivastava, and V. Raghunathan, "Harvesting Aware Power Management for Sensor Networks", In *ACM/IEEE DAC 2006*, San Francisco, CA, USA, July 2006.

[62] S. Sudevalayam, and P. Kulkarni, "Energy Harvesting Sensor Nodes: Survey and Implications", In *IEEE Communications Surveys & Tutorials*, 13(3): 443–461, September 2011.

[63] J.N.Al-Karaki, and A.E. Kamal, "Routing techniques in wireless sensor networks: a survey", In *IEEE Wireless Communications*, 6(1): 6-28, December 2004

[64] L. E. Y. Mimbela and L. A. Klein. "Summary of vehicle detection and surveillance technologies used in intelligent transportation systems".*Federal Highway Administration, Intelligent Transportation Systems Joint Program Office*, 2007

[65] S. Rost, and H. Balakrishnan, "Memento: A Health Monitoring System for Wireless Sensor Networks", *Proc. IEEE SECON 2006*, Reston, VA, USA, Sep. 2006.

[66] D. Benferhat, F. Guidec, and P. Quinton, "Disruption-Tolerant Wireless Sensor Networking for Biomedical Monitoring in Outdoor Conditions', *Proc. BodyNets 2012*, Oslo, Norway, Sep. 2012.

[67] A.B. Mohamed, T. Val, L. Andrieux, and A. Kachouri, "Using a Kinect WSN for Home Monitoring: Principle, Network and Application Evaluation", *Proc. ICWCUCA 2012*, Clermont Ferrand, France, Aug. 2012.

[68] T. He, S. Krishnamurthy, J.A. Stankovic, T. Abdelzaher, L. Luo, R. Stoleru, T. Yan, L. Gu, J. Hui, B. Krogh, "Energy-Efficient Surveillance System using Wireless Sensor Networks", *Proc. ACM MobiSys 2004*, Boston, MA, USA, Jun. 2004.

[69] M. Zuniga and B. Krishnamachari, "Analyzing the Transitional Region in Low Power Wireless Links", *Proc. IEEE SECON 2004*, Santa Clara, CA, USA, Oct. 2004.

[70] M. Z. Zamalloa and B. Krishnamachari, "An Analysis of Unreliability and Asymmetry in Low-power Wireless Links", *ACM Transactions on Sensor Networks*, vol. 3, no. 2, Jun. 2007.

[71] W. Yuan, X. Wang, and J.-P. Linnartz, "A Coexistence Model of IEEE 802.15.4 and IEEE 802.11b/g", *Proc. IEEE SCVT 2007*, Delft, Netherlands, Nov. 2007.

[72] K. Srinivasan, P. Dutta, A. Tavakoli, and P. Levis. "An Empirical Study of Low-Power Wireless", *ACM Transactions on Sensor Networks*, vol. 6, no. 2, Mar. 2010.

[73] L. Mottola, G. Pietro Picco, M. Ceriotti, . Gună, and A. L. Murphy. "Not All Wireless Sensor Networks Are Created Equal: A Comparative Study on Tunnels", *ACM Transactions on Sensor Networks*, vol. 7, no. 2, Sep. 2010.

[74] H. Bizagwira, J. Toussaint, and M. Misson, "Experimental Protocols and Testbed for Radio Link Quality Evaluation over the Freshwater", *Proc. IFIP Wireless Days 2014*, Rio de Janeiro, Brazil, Nov. 2014.

[75] C. A. Boano, H. Wennerstrom, M. A. Zúñiga, J. Brown, C. Keppitiyagama, F. Oppermann, U. Roedig, L. Norden, T. Voigt, and K. Römer, "Hot Packets: A Systematic Evaluation of the Effect of Temperature on Low Power Wireless Transceivers", *Proc. ExtremeCom 2013*, Eyjafjallajokull, Iceland, Aug. 2013.

[76] H. Wennerstrom, F. Hermans, O. Rensfelt, C. Rohner and L.A. Norden, "A Long-Term Study of Correlations between Meteorological Conditions and 802.15.4 Link Performance", *IEEE SECON 2013*, New Orleans, USA, Jun. 2013.

[77] N. Baccour, A. Koubaa, L. Mottola, M. A. Zúñiga, H. Youssef, C. A. Boano, and M. Alves, "Radio Link Quality Estimation in Wireless Sensor Networks: A Survey", *ACM Transactions on Sensor Networks*, vol. 8, no. 4, Sep. 2012.

[78] K. Heurtefeux, F. Valois, "Is RSSI A Good Choice for Localization in Wireless Sensor Network?", *Proc. IEEE AINA 2012*, Fukuoka, Japan, Mar. 2012.

[79] A.-S. Tonneau, N. Mitton, and J. Vandaele, "How to Choosean Experimentation Platform for Wireless Sensor Networks? A Survey on Static and Mobile Wireless Sensor Network Experimentation Facilities", *Ad Hoc Networks (Elsevier)*, 2015.

[80] C. Adjih, E. Baccelli, E. Fleury, G. Harter, N. Mitton, T. Noel, R. Pissard-Gibollet,F. Saint-Marcel, G. Schreiner, J. Vandaele, and T. Watteyne, "FIT IoT-LAB: A large scale open experimental IoT testbed", *In 2015 IEEE 2nd World Forum on Internet of Things (WF-IoT)*, 459464

[81] R. Lim, F. Ferrari, M. Zimmerling, C. Walser, P. Sommer, and J. Beutel, "FlockLab: A testbed for distributed, synchronized tracing and profiling of wireless embedded systems", *In 2013 ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, 153165.

[82] L. Sanchez, J.A. Galache, V. Gutierrez, J.M. Hernandez, J. Bernat, A. Gluhak, T. Garcia, "SmartSantander: The meeting point between Future Internet research and experimentation and the smart cities", *In IEEE Future Network & Mobile Summit*, Poland, Warsaw, 2011

[83] M. Bradbury, A. Jhumka and C. Maple, "The impact of decreasing transmit power levels on FlockLab to achieve a sparse network", *In Proceedings of the 2nd Workshop on Benchmarking Cyber-Physical Systems and Internet of Things*, ACM, New York, NY, USA, 7-12, April, 2019

[84] A. Goldsmith, "Wireless Communications", Cambridge University Press, New York, NY, USA, 2005

[85] V. Erceg, L.J. Greenstein, S.Y. Tjandra, S.R. Parkoff, A. Gupta, B. Kulic, A.A Julius, and R. Bianchi, "An Empirically based Path Loss Model for Wireless Channels in Suburban Environments", *IEEE Journal on Selected Areas in Communications*, vol.17, no.7, pp.1205-1211, Jul 1999.

[86] A. Sikora and V. F. Groza, "Coexistence of IEEE 802.15.4 with other Systems in the 2.4 GHz-ISM-Band," IEEE Instrumentation and Measurement Technology Conference Proceedings, Ottawa, 2005, pp. 1786-1791.

[87] I. Khoufi, P. Minet and B. Rmili, "Beacon advertising in an IEEE 802.15. 4e TSCH network for space launch vehicles", Acta Astronautica, 158, 76-88, 2019.

[88] R.A. Koutsiamanis, G.Z. Papadopoulos, X. Fafoutis, J.M. Del Fiore, P. Thubert and N. Montavont. "From best effort to deterministic packet delivery for wireless industrial IoT networks", IEEE Transactions on Industrial Informatics, 14(10), pp.4468-4480, 2018.

[89] M. Petrova, and al, "IEEE 802.15.4 Low Rate - Wireless Personal Area Network Coexistence Issues," Proc. IEEE WCNC'06, Las Vegas, USA

[90] M. O. Ergin, V. Handziski, A. Behboodi and A. Wolisz, "Determining Node Sequence in a Linear Configuration", *Proc. IPIN 2014*, Busan, Korea, Oct. 2014.

[91] A. Boukerche, H. Oliveira, E. Nakamura, and A. Loureiro, "Localization Systems for Wireless Sensor Networks", *IEEE Wireless Communications*, vol. 14, no. 6, pp. 612, Dec. 2007.

[92] A. Zanella and A. Bardella, "RSS-Based Ranging by Multichannel RSS Averaging", IEEE Wireless Communications Letters, vol. 3, no. 1, pp. 1013, Dec. 2014.

[93] B. Pavkovic, F. Theoleyre, D. Barthel and A. Duda, "Experimental Analysis and Characterization of a Wireless Sensor Network Environment", *Proc. ACM PE-WASUN 2010*, Bodrum, Turkey, Oct. 2010.

[94] S.Y. Cheung, S. Coleri, B. Dundar, S. Ganesh, C.W. Tan and P. Varaiya, "Traffic measurement and vehicle classification with single magnetic sensor", Transportation Research Record, 2005, pp.173-181.

[95] I. Jawhar, N. Mohamed, D.P. Agrawal, "Linear Wireless Sensor Networks: Classification and Applications", *Journal of Network and Computer Applications*, 34(5):1671–1682, Sep. 2011.

[96] M. Perillo, Z. Cheng and W. Heinzelman, "On the problem of unbalanced load distribution in wireless sensor networks," IEEE Global Telecommunications Conference Workshops, November 2004, Dallas, USA, pp. 74-79.

[97] M. Noori and M. Ardakani, "Characterizing the traffic distribution in linear wireless sensor networks", IEEE Communications Letters, vol. 12, no. 8, pp. 554-556, Aug. 2008.

[98] H. Yetgin, and K. T. K. Cheung, and M. El-Hajjar and L. H. Hanzo, "A Survey of Network Lifetime Maximization Techniques in Wireless Sensor Networks," IEEE Communications Surveys & Tutorials, vol. 19, no. 2, pp. 828-854, Second quarter 2017.

[99] L. B. Saad, and Bernard Tourancheau. "Multiple Mobile Sinks Positioning in Wireless Sensor Networksfor Buildings." International Conference on Sensor Technologies and Applications (SensorComm), 2009, Athenes, Greece. pp.264270.

[100] L. B. Saad, and Bernard Tourancheau. "Towards an Efficient Positioning of Mobile Sinks in WirelessSensor Networks inside Buildings." NTMS-WSN workshop, 2009, Le Caire, Egypt, pp.1 -5.

[101] L. B. Saad, and Bernard Tourancheau. "Sinks Mobility Strategy in IPv6-based WSNs for NetworkLifetime Improvement." International Conference on New Technologies, Mobility and Security (NTMS), IFIP, 2011, Paris, France.

[102] Y. Guo, F. Kong, D. Zhu, A. . Tosun, and Q. Deng (2010, April). “Sensor placement for lifetime maximization in monitoring oil pipelines.” In Proceedings of the 1st ACM/IEEE International Conference on Cyber-Physical Systems (pp. 61-68).

[103] S. Olariu and I. Stojmenovic. ”Design guidelines for maximizing lifetime and avoiding energy holes in sensor networks with uniform distribution and uniform reporting“. In INFOCOM 2006, pages 112, April 2006

[104] C. Ok, H. Thadakamalla, U. Raghavan, S. Kumara, S. G. Kim, X. Zhang, and S. Bukkapatnam (2007, September). “Optimal transmission power in self-sustainable sensor networks for pipeline monitoring.” In IEEE International Conference on Automation Science and Engineering (CASE), 2007, (pp. 591-596).

[105] K. Aberer and M. Hauswirth and A. Salehi. “A middleware for fast and flexible sensor network deployment.” In Proceedings of the 32nd international conference on Very large data bases, September 2006.

[106] N. Raveendranathan and S. Galzarano and V. Loseu and R. Gravina and R. Giannantonio and M. Sgroi and R. Jafari and G. Fortino. “From modeling to implementation of virtual sensors in body sensor networks.“ IEEE Sensors Journal, 12(3), pp.583-593, 2012

[107] S. Madria and V. Kumar and R. Dalvi, ”Sensor Cloud: A Cloud of Virtual Sensors.” In IEEE Software, vol. 31, no. 2, pp. 70-77, Mar.-Apr. 2014.

[108] P. Vicaire, T. He, Q. Cao, T. Yan, G. Zhou, L. Gu, L. Luo, R. Stoleru, J. A. Stankovic, and T. F. Abdelzaher. Achieving long-term surveillance in vigilnet. ACM Transaction Sensor Networks, 5(1):139, 2009

[109] Liu, Xin, and Prasant Mohapatra. “On the deployment of wireless data back-haul networks.“ IEEE Transactions on Wireless Communications 6.4 (2007).

[110] S. Plancoulaine, and A. Bachir, and D. Barthel, “WSN Node energy dissipation“, Technical repport, France Telecom R&D, Internal Repport, July 2006

[111] A. M. Zahid and A.B.Kamalrulnizam and A. Muhammad and M.. Hafiz, ’An overview of routing techniques for road and pipeline monitoring in linear sensor networks’, Wireless Networks, 2018

[112] W. Ye, and J. Heidemann, and D. Estrin, ”Medium access control with coordinated adaptive sleeping for wireless sensor networks“, IEEE/ACM Trans. Netw., vol 12, no. 3, pp. 493-506, June 2004.

[113] Y. Zhao, and J. Wu, and F. Li and S. Lu, ”On maximizing the lifetime of wireless sensor networks using virtual backbone scheduling“, IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 8, pp. 1528-1535, Aug. 2012.

[114] Y. Chen, and Q. Zhao, and V. Krishnamurthy and D. Djonin, "Transmission scheduling for optimizing sensor network lifetime: A stochastic shortest path approach", IEEE Trans. Signal Process., vol. 55, no. 5, pp. 2294-2309, May 2007

[115] C. Sevgi and A. Koyiit. "Optimal deployment in randomly deployed heterogeneous WSNs: A connected coverage approach", Journal of Network and Computer Applications, 46,pp.182-197, 2014.

[116] M.Z.A. Bhuiyan and G. Wang and J. Cao and J. Wu, "Deploying wireless sensor networks with fault-tolerance for structural health monitoring", IEEE Transactions on Computers, 64(2), pp.382-395, 2015

[117] F.J. Parrado-Garcia and J. Vales-Alonso and J.J. Alcaraz, "Optimal Planning of WSN Deployments for In SituLunar Surveys", IEEE Transactions on Aerospace and Electronic Systems, 53(4), pp.1866-1879, 2017.

[118] A. Boubrima and W. Bechkit and H. Rivano, "Optimal WSN deployment models for air pollution monitoring", IEEE Transactions on Wireless Communications, 16(5), pp.2723-2735, 2017.

[119] N. Kulkarni and N.R. Prasad and R. Prasa, "A novel sensor node deployment using low discrepancy sequences for WSN", Wireless Personal Communications, 100 (2), pp.241-254, 2018.

[120] S. Potthuri and T. Shankar and A. Rajesh, "Lifetime improvement in wireless sensor networks using hybrid differential evolution and simulated annealing (DESA)". Ain Shams Engineering Journal, Elsevier, Volume 9, Issue 4,Pages 655-663, 2018

[121] A. Boukerche, H. Oliveira, E. Nakamura, A. Loureiro, "Localization Systems for Wireless Sensor Networks", *IEEE Wireless Communications*, 14(6):6–12, Dec. 2007.

[122] G. von Zengen, Y. Schroder, S. Rottmann, F. Busching and L.C. Wolf, "No-Cost Distance Estimation using Standard WSN Radios", *Proc. IEEE Infocom 2016*, San Francisco, CA, Apr. 2016.

[123] D. Niculescu and B. Nath, *Ad hoc positioning system (APS)*, Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE, San Antonio, TX, 2001, pp. 2926-2931 vol.5

[124] S. Zhang, X. Liu, J. Wang, J. Cao, G. Min, "Accurate Range-Free Localization for Anisotropic Wireless Sensor Networks", *ACM Transaction on Sensor Networks*, 11(3):51, May 2015.

[125] S. Merkel, S. Mostaghim, H. Schmeck, "Distributed Geometric Distance Estimation in Ad Hoc Networks", *Proc. Adhoc-Now 2012*, Belgrade, Serbia, Jul. 2012.

[126] X. Zhu, G. Chen, "Spatial Ordering Derivation for One-Dimensional Wireless Sensor Networks", *Proc. IEEE ISPA 2011*, Busan, Korea, May 2011.

[127] X. Zhu, X. Wu, G. Chen, "Relative Localization for Wireless Sensor Networks with Linear Topology", *Computer Communications*, 36(1516):1581–1591, Sep. 2013.

[128] Y. Cui, Q. Wang, H. Yuan, X. Song, X. Hu, L. Zhao, "Relative Localization in Wireless Sensor Networks for Measurement of Electric Fields under HVDC Transmission Lines", *Sensors*, 15(2):3540–3564, Feb. 2015.

[129] I. Jawhar, N. Mohamed and L. Zhang, "A distributed topology discovery algorithm for Linear Sensor Networks," 2012 1st *IEEE International Conference on Communications in China (ICCC)*, Beijing, 2012, pp. 775-780

[130] M. O. Ergin and A. Wolisz, *Node position discovery in wireless sensor networks*, Positioning Navigation and Communication (WPNC), 2012 9th Workshop on, Dresden, 2012, pp. 157-162

[131] M. Onur Ergin, Vlado Handziski, and Adam Wolisz. 2013. *Node Sequence Discovery in Wireless Sensor Networks*. In IEEE DCOSS 2013, Washington, DC, USA, 394-401

[132] M.O. Ergin, V. Handziski, A. Behboodi, A. Wolisz, "Determining Node Sequence in a Linear Configuration", *Proc. IPIN 2014*, Busan, Korea, Oct. 2014.

[133] Malekpour A, Ling TC, Lim WC, *Location determination using radio frequency RSSI and deterministic algorithm*, In IEEE Communication Networks and Services Research Conference, 2008, May 5, pp. 488-495

[134] Z. Lotker, M.M. de Albeniz, and S. Pérénnes, "Range-Free Ranking in Sensors Networks and Its Applications to Localization", *Proc. Adhoc-Now 2004*, Vancouver, BC, Canada, Jul. 2004.

[135] T.P. Hettmansperger and J.W. McKean, "Robust nonparametric statistical methods", Kendall's Library of Statistics. 5th edition. London; New York: Edward Arnold; John Wiley and Sons, Inc. pp. xiv+467. ISBN 978-0-340-54937-7. MR 1604954.

[136] T. Watteyne, I. Augé-Blum, M. Dohler, S. Ubéda, D. Barthel, "Centroid Virtual Coordinates: A Novel Near-Shortest Path Routing Paradigm", *Computer Networks*, 53(10):1697–1711, Jul. 2009.

[137] M. D. Sarr and F. Delobel and M. Misson, I. Niang. "Automatic discovery of topologies and addressing for linear wireless sensors networks", In Wireless Days (WD), 2012 IEEE IFIP, pp. 1-7, November 2012.

[138] N. Wisitpongphan, and F. Bai, and P. Mudalige, and O.K. Tonguz, "On the Routing Problem in Disconnected Vehicular Ad-hoc Networks", In *IEEE INFOCOM 2007*, Anchorage, Alaska-USA, May 2007, Pages 2291-2295

[139] S. Uppoor, O. Trullols-Cruces, M. Fiore, J.M. Barcelo-Ordinas, "Generation and Analysis of a Large-scale Urban Vehicular Mobility Dataset", IEEE Transactions on Mobile Computing, Vol.13, No.5, May 2014

[140] CITI Lab, http://kolntrace.project.citi-lab.fr/, Visited on June 13th, 2019

[141] A. Boukerche and A. Darehshoorzadeh. "Opportunistic Routing in Wireless Networks: Models, Algorithms, and Classifications", ACM Comput. Surv. 47, 2, Article 22, November 2014

[142] M. Zimmerling and D. Waltenegus and M.R. Johnathan, "Energy-efficient routing in linear wireless sensor networks", In MASS 2007, October, pp 1-3.

[143] M. Zimmerling and D. Waltenegus and M.R. Johnathan. "Localized power-aware routing in linear wireless sensor networks", In Proceedings of the 2nd ACM international conference on Context-awareness for self-managing systems, pp. 24-33., ACM 2008.

[144] F. Jacquet and M. Misson and I. Niang, "Using a token approach for the MAC layer of linear sensor networks: Impact of the node position on the packet delivery", In Wireless Days (WD), 2014 IEEE IFIP, pp. 1-4., November 2014

[145] R. Draves and J. Padhye and B. Zill, "Comparison of routing metrics for static multi-hop wireless networks", In ACM SIGCOMM Computer Communication Review, Vol 34, Num. 4, pp. 133-144.

[146] F. Ye, G. Zhong, S. Lu and L. Zhang, "Gradient broadcast: A robust data delivery protocol for large scale sensor networks", Wireless Networks, 11(3), pp.285-298, 2005.

[147] WSNet, "http://wsnet.gforge.inria.fr/", Visited on January 30th, 2019

Appendix A

Linear Wireless Sensors Network simulator

During this thesis, we developed a simulator for quick LWSN benchmarking. To evaluate the performance of the ranking algorithm proposed in this thesis and presented in Section 6.4, we used this simulator, described in the remainder of this section.

A.1 Motivations

Numerous simulators are proposed in the literature for WSN. Many of them are for generalpurpose WSN and are used to simulate networks with various characteristics. They also simulate various aspects of wireless communications, ranging from signal modulation and propagation to medium access control. Generic channel models like the one presented in Chapter 3 are usually used. Concerning this model, the results presented in this thesis, particularly with respect to the RSSI (used by the signal propagation module in WSN simulators), confirm the findings of previous experimental campaigns, which demonstrate that the received signal is not a strictly decreasing function of the distance [132, 133]. Moreover, the deviation from this predicted behavior seems to be exacerbated by a deployment at ground level. That is what motivates the development of a new simulator to evaluate the ranking algorithm proposed in this thesis. Our goal is to develop a simulator which not only takes into account the link properties at ground level, but can also be adapted by other researchers with little effort. A typical use-case of this simulator is a researcher working on a particular network service like localization or topology control, who wants to evaluate a solution under an unusual link model, without implementing all the network communication stack.

A.2 Simulation parameters

In this section, we describe the inputs of the simulator.

A.2.1 Topology parameters

Concerning the topology parameters, the simulator takes as input the number of nodes, the maximum communication range and the node deployment. The deployment can be uniform, random or customized. A uniform deployment is defined by a constant distance between two consecutive nodes. For the random deployment, we assume that the number of nodes per unit length follows a Poisson distribution. Thus, the distance between two consecutive nodes in a random deployment follows an exponential distribution. That is why, in the case of a random deployment, the parameter is a distance which represents the parameter of the exponential law, i.e. the average distance between two consecutive nodes. In the case of a personalized deployment, the user provides to the simulator a file containing for each node, its ID and its coordinate relative to the sink. If N is the number of nodes, the first node, the sink, has ID 0 and the last one has ID $N-1$. The sink is considered as the origin with coordinate 0, and the coordinate of each other node is its distance to the sink.

A.2.2 Network parameters

We modeled the network as an oriented weighted graph. The vertices of the graph correspond to the nodes of the network. An edge (S_i,S_j) corresponds to the link from node S_i to node S_j. This materializes the fact that, in the network, the node S_j can receive messages from S_i. The weight of the edge (S_i,S_j) models the quality of the wireless link $S_i \rightarrow S_j$ or, from the node S_j point of view, the quality of the link with the node S_i. Thus, for link quality, the user can decide to use the simple UDG model widely used in the literature, or proposes a new model adapted to the studied. In this second case, the user must provide to the simulator a datafile containing empirical data. The expected format is a two-columns file, where the first column is the distance between two nodes and the second one is the link metric. This metric could be derived from the RSSI, the packet reception rate or any other metric designed by the user. The edge weight can be randomly selected from the empirical data, or can be generated from a normal distribution of parameter μ and σ calculated from the empirical data.

Appendix B

List of publications

1. Journals

   (a) Rodrigue DOMGA KOMGUEM, Razvan STANICA, Maurice TCHUENTE, Fabrice VALOIS, *Ground Level Deployment of Wireless Sensor Networks: Experiments, Evaluation and Engineering Insight*, In Sensors, 19(15), 2019, p.3358.

   (b) Rodrigue DOMGA KOMGUEM, Razvan STANICA, Maurice TCHUENTE, Fabrice VALOIS, *Sensor deployment in wireless sensor networks with linear topology using virtual node concept*, In Wireless Networks, 2019, pp.1-16.

2. International Conferences

   (a) DOMGA K.R.; STANICA R.; TCHUENTE M.; VALOIS F., *Adaptive message generation in intersection monitoring WSN*, CRI 2019, Yaounde, Cameroon, December, 2019.

   (b) Rodrigue DOMGA KOMGUEM, Razvan STANICA, Maurice TCHUENTE, Fabrice VALOIS, *Nodes Ranking in Wireless Sensor Network with Linear Topology*, IEEE WD2017, Porto, Portugal, March, 2017

   (c) Rodrigue DOMGA KOMGUEM, Razvan STANICA, Maurice TCHUENTE, Fabrice VALOIS, *WARIM : Wireless Sensor Network Architecture for a Reliable Intersection Monitoring*, ITSC - IEEE 17th International Conference on Intelligent Transportation Systems, Qingdao, China, October, 2014

Wireless Networks (2019) 25:4947–4962
https://doi.org/10.1007/s11276-019-02071-x

Sensor deployment in wireless sensor networks with linear topology using virtual node concept

Rodrigue K. Domga[1,2] · Razvan Stanica[2] · Maurice Tchuente[1] · Fabrice Valois[2]

Published online: 3 July 2019

Abstract
In a multi-hop wireless sensor network with a convergecast communication model, there is a high traffic accumulation in the neighborhood of the sink. This area constitutes the bottleneck of the network since the sensors deployed withing it rapidly exhaust their batteries. In this paper, we consider the problem of sensors deployment for lifetime maximization in a linear wireless sensor network. Existing approaches express the deployment recommendations in terms of distance between consecutive sensors. Solutions imposing such constraints on the deployment may be costly and difficult to manage. In this paper, we propose a new approach where the network is formed of virtual nodes, each associated to a certain geographical area. An analytical model of the network traffic per virtual node is proposed and a greedy algorithm to calculate the number of sensors that should form each virtual node is presented. Performance evaluation shows that the greedy deployment can improve the network lifetime by up to 40%, when compared to the uniform deployment. Moreover, the proposed approach outperforms the related work when complemented by a scheduling algorithm which reduces the messages overhearing. It is also shown that the lifetime of the network can be significantly improved if the battery capacity of each sensor is dimensioned taking into account the traffic it generates or relays.

Keywords Linear wireless sensor network · Deployment · Energy efficiency · Virtual node · Greedy algorithm · Lifetime

1 Introduction

The topology and architecture of wireless sensor networks (WSN) generally depend on the target application and the geographical area where sensors are deployed. A linear WSN (LWSN) is a special case, where the physical topology of the network is a line [1]. The applications of LWSN are diverse, e.g. monitoring of large infrastructure such as bridges [2] and dams [3], road traffic observation [4], and border control [5]. Sensors typically operate on a small capacity battery and are thus limited in their active lifetime. It may be infeasible or expensive to change batteries in sensors once a wireless sensor network is deployed. It is therefore important to design energy-efficient communication protocols and to optimize deployment strategies.

In this paper, we are interested in sensors deployments. The goal of this paper is to present an energy-efficient and simple deployment strategy which can allow a network to operate for a long time. In a multi-hop LWSN with convergecast traffic (see Fig. 1), the amount of messages that a sensor has to forward significantly increases as the distance to the sink becomes smaller [6, 7]. Moreover, sensors are more exposed to the overhearing phenomenon: a sensor can receive a message addressed to another one. Thus, sensors closest to the sink, because of their high traffic load and their high overhearing probability, tend to exhaust their batteries early, causing network failure. This problem can

✉ Rodrigue K. Domga
rodrigue.domga-komguem@insa-lyon.fr

Razvan Stanica
razvan.stanica@insa-lyon.fr

Maurice Tchuente
maurice.tchuente@gmail.com

Fabrice Valois
fabrice.valois@insa-lyon.fr

[1] CETIC, LIRIMA, Faculté des Sciences, Université de Yaoundé I, BP 812, Yaoundé, Cameroon

[2] INSA Lyon, CITI-Inria, 69621 Villeurbanne, France

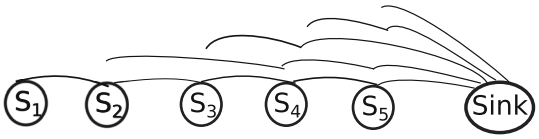


Fig. 1 A LWSN of five sensors and one sink with convergecast traffic

be addressed through suitable deployment, by designing energy-efficient protocols or by considering non homogeneous energy distribution.

Many sensors deployment solutions are proposed in the literature [8–11]. In these solutions, message receptions are not always taken into account, although transmission and reception consume more or less the same energy. Moreover, deployment recommendations are usually expressed in terms of distances between consecutive sensors. For a LWSN with a large number of sensors, a solution with such an accuracy on the inter-node distances may be costly and difficult to implement.

In a LWSN, an event that occurs in an area can be detected by many sensors deployed in that area. A new deployment approach in which the LWSN is divided into virtual nodes and where consecutive virtual nodes should be wirelessly connected is proposed in this paper. An event happening within the area covered by a virtual node can be detected by any sensor, no matter its position in the virtual node. This paper also proposes a greedy algorithm for calculating the number of physical sensors to deploy for a virtual node, taking into account the traffic per virtual node. It is assumed that an average transmission power which guarantees connectivity between consecutive virtual nodes is used by sensors.

The concept of virtual node or virtual sensor is not new in the field of WSN. It is particularly used in the context of WSN programming [12–14]. In [13], every component represents a processing task applied to a stream of data originated from physical sensors and can be modeled as a virtual node. In [12], a virtual node corresponds either to a data stream received directly from sensors or to a data stream derived from other virtual sensors. In this paper, the concept of virtual node is used in the context of sensors deployment to reduce the complexity of the deployment.

The solution proposed here is different from the existing ones on at least two points. First, there is no need to keep a precise distance between sensors. Indeed, after dividing the network into virtual nodes (the number of virtual nodes is smaller than the number of physical nodes), the position of sensors within a virtual node is not predefined. Second, contrarily to our solution, which considers also message receptions, the related work only account for message transmissions. This is not realistic because it is well known that most radio modules consume almost the same energy for receptions and transmissions [15]. It is important to note that message receptions mentioned here include receptions due to the overhearing phenomenon.

Results presented in this paper show that, by properly selecting the number of sensors, a simple greedy deployment can improve the network lifetime by up to 40%, when compared to the uniform deployment. A framework which allows to compare the virtual node-based approach proposed in this paper to the distance-based one defined in [8] is also designed. Results show that our approach outperforms that of [8] when a scheduling algorithm which reduces the messages overhearing is used.

A solution for heterogeneous networks where the battery capacities of sensors are not uniform is also proposed. More precisely, it is assumed that the battery capacity of a sensor is proportional to its traffic load. Results show that this solution significantly improves the network lifetime compared to the homogeneous case.

The remainder of this paper is organized as follows. Section 2 presents related work on sensor deployment. The virtual node approach is formalized in Sects. 3 and 4 describes a greedy algorithm for sensor deployment. Section 5 is devoted to performance analysis. In Sect. 6, the comparison to the related work is done, and Sect. 7 considers the case of heterogeneous networks. Finally, Sect. 8 concludes this paper.

2 Related work

In a multi-hop WSN with a convergecast communication model, the sensors close to the sink have a high transmission activity. Thus, they constitute the bottleneck of the network, since sensors are constrained in terms of battery (limited capacity), memory (limited buffer size) and communication capabilities (high contention). Many energy-efficient solutions have been proposed at all layers of the communication stack, ranging from hardware design, MAC and routing protocols, and data aggregation [16]. In this paper, we are interested in a sensor deployment strategy which could prolong the network lifetime.

Sensor deployment strategies can be designed with different objectives in mind. For example, the objective of the deployment can be to guarantee a balanced energy consumption of sensors in the network. To address this problem, one can deploy more sensors close to the sink and make them transmit at lower power levels [8–10]. In this way, sensors close to the sink transmit more messages, but consume less energy per message. When all sensors use the same transmission power, in network areas with high sensors density, a message can be forwarded by a significant number of sensors. Load-balancing can then be achieved by adopting duty-cycle scheduling [17].

The problem of sensors deployment is widely addressed in the literature [18–23]. In [23], authors propose a hybrid differential evolution and simulated annealing (DESA) algorithm for clustering and choice of cluster heads. They prolong the network lifetime by appropriately selecting the cluster head and affecting sensors to clusters. Sevgi and Koçyiğit [18] addresses the optimal deployment problem defined as the determination of the minimum possible number of sensors aiming to achieve the targeted partial connected coverage, the lowest financial cost and the the highest lifetime. In [22], authors propose a novel node deployment strategy based on Quasi-random method of low-discrepancy sequences to increase the lifetime and the coverage of the network. While most of the previous work consider 2-D or 3-D topologies, this paper focues on LWSN, i.e. a WSN with 1-D topology. Rather than consider a hierarchical network (like in clustered networks), we are interested by flat WSN in which all the messages generated by sensors are forwarded in a multi-hop manner toward the sink deployed at an network edge.

In [7], authors propose an analytic description of the traffic over a linear, randomly deployed WSN. They evaluate the effect of the number of sensors and their distribution on network traffic. They propose a non uniform deployment obtained through an increased network density closer to the sink. Their results show that, given a number N of sensors, such a deployment can significantly reduce the maximum traffic load per node compared to a uniform deployment, and hence improve the network lifetime. However, the authors do not propose a deployment strategy or some concrete recommendations for sensors deployment.

In [11], the problem of sensor deployment is addressed when all sensors have the same transmission power level. Given the required lifetime of a sensor network, the energy constraint of sensors, and the area to be covered, authors study the problem of finding the minimum number of sensors needed to build such a network and the corresponding deployment scheme. In [8], the authors assume that a sensor can select the transmission power among a set of power levels. Given the energy constraint of sensors, the problem of sensors deployment is addressed with the goal of maximizing the covered area and the network lifetime. They formalize the deployment problem as a Mixed-Integer Linear Programming (MILP) problem. Their results show that, by properly selecting the number of sensors, the distance between them, and the corresponding transmission power, the WSN lifetime can be improved by up to 30%.

Despite their performance in terms of network lifetime, a drawback of the solutions proposed by [8, 11] is the complexity of their implementation, particularly in large LWSN. Indeed, the deployment recommendations are expressed in terms of distances between consecutive sensors in the network. This is not practical if the deployment must be done by a human for a large network. In this paper, we focus on a solution that is simple to implement, i.e. it requires reduced deployment effort, and can at the same time improve significantly the network lifetime compared to a uniform deployment. Another important contribution with respect to the state of the art is that, because of the broadcast nature of the wireless communication channel, the proposed solution takes into account messages received due to the overhearing phenomenon.

3 Problem description

Given a set of homogeneous sensors in terms of battery capacity, our goal is to design a simple, energy-efficient and realistic sensor deployment that maximizes the network lifetime. In Sect. 7, we investigate the case where sensors are heterogeneous.

In many WSNs applications, the same event can be detected by many sensors if the network is dense. For instance, if a WSN is deployed to monitor the vehicular traffic on a highway or on lanes of an intersection, a vehicle might be detected at the same time by different sensors. This depends on the average vehicle length and the distance between two consecutive nodes. Another example is the monitoring of the status of an infrastructure like a bridge. In such an application, a damage on an area of the bridge could be detected by many sensors. Thus, rather than considering a flat network like in related work [8, 11], we consider a large network divided into virtual nodes wirelessly connected. A virtual node represents the set of sensors deployed in a given area which can measure the same physical phenomenon. In the related work, the deployment recommendations is expressed in terms of distance between consecutive sensors. In our approach, once the virtual nodes are defined, the position of sensors in a virtual nodes are not predefined. In the next section, our hypothesis are further detailed. Table 1 presents the list of important notations used in our model.

3.1 Assumptions

We consider a set of N sensors $\{S_1, \ldots, S_N\}$ deployed in a $1D$ space to monitor a phenomenon. We assume that the area to be monitored is covered by K virtual nodes $\{B_1, \ldots, B_K\}$, as presented in Fig. 2. We also assume that the sink is a single node which corresponds to the virtual node B_{K+1}. In our work, we assume that consecutive virtual nodes are connected, in order to guarantee that a message generated in a virtual node can be received by the sink. However, we note that this assumption does not imply that

Table 1 Model notations

Parameter	Description
N	Number of sensors
K	Number of virtual nodes
λ	Avg. number of messages generated per virtual node per time slot
Δ	Time slot duration
p	Probability for the virtual node B_i to receive messages from virtual node B_{i-2} or B_{i+2}
α	Probability for a sensor to receive messages destined to another sensor
E_{sensor}	The sensor battery capacity
P	The average transmission power
E_i	Energy consumed per sensor per time slot in the virtual node B_i
τ	The time required to transmit one packet
T_i	The total number of transmissions from a virtual node B_i: generated and relayed messages
n_i	Number of sensors in the virtual node B_i
$\overline{R}_i^{ns}$	Avg. number of receptions per sensor in virtual node B_i when the radio of nodes is always on
$\overline{R}_i^{os}$	Avg. number of receptions per sensor in virtual node B_i when there is no overhearing
$\overline{R}_i^{ps}$	Avg. number of receptions per sensor in virtual node B_i when the overhearing is partial
$\overline{O}_i$	Avg. number of operations per sensor per time slot in the virtual node B_i

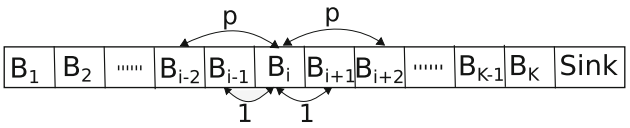


Fig. 2 Network model

physical nodes use homogeneous power levels, since the physical distance between nodes is not necessarily uniform, and several physical nodes can form a virtual node.

An event happening in an area can be detected by many sensors. We consider that for each event detected by a virtual node, only one message will be generated, as a result of a cooperation algorithm executed by sensors of the same virtual node. This cooperation algorithm also equally distributes to the nodes in B_i the traffic coming from B_j, $j<i$. We also assume that sensors can run a ranking algorithm [24] to determine the relative position of each sensor to the sink. Such an algorithm allows sensors to compute the virtual node index in which they are deployed.

3.2 Connectivity model and traffic pattern

It is assumed a multi-hop LWSN with a convergecast communication model, i.e data from all sensors are forwarded toward the sink located at position $K+1$. Unlike a distributed peer-to-peer wireless network, the traffic load is highly asymmetric, i.e. sensors closer to the sink have a heavier relay load. A wireless connectivity is assumed between consecutive virtual nodes B_i and B_{i+1}. A virtual node B_i can communicate with B_{i+2} or B_{i-2} with probability p (see Fig. 2). This models the an-isotropic property of the signal. We consider a shortest path routing protocol [25], i.e. a packet transmitted by B_i and received by B_{i+1} and B_{i+2} is forwarded by B_{i+2}. The average number of messages generated from a virtual node per time unit of duration Δ is denoted by λ. Typically, λ is application dependent and represents the average number of events detected by a virtual node during a time slot. T_i denotes the total traffic load (messages generated or relayed) of virtual node B_i per time unit. This traffic (T_i) comes from three different sources:

- From B_i: λ messages generated
- From B_{i-1}: $(1-p)\cdot T_{i-1}$ messages relayed
- From B_{i-2}: $p\cdot T_{i-2}$ messages relayed.

Thus, the traffic load of virtual node B_i is given by Eqs. 1.

$$T_i = \begin{cases} \lambda, & \text{if } i = 1 \\ (2-p)\cdot\lambda, & \text{if } i = 2 \\ p\cdot T_{i-2} + (1-p)\cdot T_{i-1} + \lambda, & \text{if } 2 < i \le K \end{cases} \tag{1}$$

Equation 1 is a recurrence equation whose solution is given by Eq. 2.

$$T_i = \beta + \gamma\cdot(-p)^i + \frac{\lambda\cdot i}{p+1}, \forall i, 1 \le i \le K \tag{2}$$

In Eq. 2, β and γ are constants calculated from T_1 and T_2

(see Eq. 1). Thus, $\beta = \frac{p \cdot \lambda}{(p+1)^2}$ and $\gamma = -\frac{p \cdot \lambda}{(p+1)^2}$. The traffic load in the virtual node $B_{i,1 \leq i \leq K}$ is thus given by Eq. 3.

$$T_i = \frac{\lambda}{p+1} \cdot \left[\frac{p}{p+1} + \frac{(-p)^{i+1}}{p+1} + i \right], \forall i, 1 \leq i \leq K \tag{3}$$

Thanks to the scheduling algorithm used, it is assumed a uniform distribution of traffic between sensors in a given virtual node. The average number of transmissions per sensor in the virtual node B_i, which contains n_i sensors, is:

$$\overline{T}_i = \frac{T_i}{n_i}, \forall i, 1 \leq i \leq K \tag{4}$$

3.3 Number of operations per sensor

In a WSN, a sensor consumes energy during data sensing and processing, and through its radio for data transmission and reception. It is well known that the radio generally consumes most of the energy of the sensor [26]. In this paper, we denote as energy-consuming operations only the transmission and the reception of a packet. Thus, the number of operations executed by a sensor is the total number of messages sent and received by that sensor. The total number of operations per virtual node is equal to the number of operations of all the sensors which form the virtual node. But, since our goal is to maximize the sensor lifetime, in this section, we focus on the number of operations per sensor. We have already seen that the average number of transmissions per sensor deployed in a virtual node is given by Eq. 4.

A sensor may receive a message destined to another sensor: it is the overhearing phenomenon [27]. Therefore, the number of receptions observed by a sensor depends on the scheduling algorithm used. We assume in this paper that the goal of the scheduling algorithm is to determine the time at which a sensor must switch its radio on or off. Thus, a perfect scheduling is when a sensor switches its radio on only when it has a message to send or to receive. In order to model the number of receptions per sensor per time period, different cases are considered in the following:

- No scheduling, where sensors keep their radio always on and ready to transmit/receive
- Optimal Scheduling, where the overhearing is eliminated by a perfect scheduling.
- Practical Scheduling, where the overhearing is partially, but not totally, eliminated.

3.3.1 Case 1: When the radio is always on

It is assumed here that each sensor maintains its radio on all the time. In such a situation, a sensor will receive all transmissions in its communication range. This includes transmissions from neighboring virtual nodes and from the virtual node to which the sensor belongs. For a virtual node B_i, $1 \leq i \leq K$, there are two sources of messages:

- Receptions from neighboring virtual nodes: this concerns transmissions from virtual nodes $i-2, i-1, i+1$ and $i+2$. This quantity is equal to $p \cdot T_{i-2} + T_{i-1} + T_{i+1} + p \cdot T_{i+2}$, since we assume that a sensor in a virtual node i can always receive transmissions from virtual nodes $i-1$ and $i+1$, and it can communicate with virtual nodes $i-2$ and $i+2$ with a probability p. We define $T_i = 0$ for $i < 1$ and $i > K$.
- Messages transmitted by other sensors in the same virtual node: $T_i - \frac{T_i}{n_i}$. Indeed, for a traffic load T_i in virtual node B_i, the average number of transmissions for a given sensor s is given by Eq. 4. Then, the remaining messages $\left(T_i - \frac{T_i}{n_i}\right)$ will be transmitted by other sensors. Sensor s will also receive these messages, since we assume each sensor continuously maintains its radio on.

The total number of receptions is then given by Eq. 5.

$$\overline{R}_i^{ns} = p \cdot T_{i-2} + T_{i-1} + T_{i+1} + p \cdot T_{i+2} + T_i - \frac{T_i}{n_i} \tag{5}$$

Under the assumption of the radio always on, the average number of operations per sensor per time unit in a virtual node B_i is the sum of transmissions (Eq. 4) and receptions (Eq. 5) per sensor and is expressed by Eq. 6.

$$\begin{aligned} \overline{O}_i &= \overline{T}_i + \overline{R}_i^{ns} \\ &= \frac{T_i}{n_i} + p \cdot T_{i-2} + T_{i-1} + T_{i+1} + p \cdot T_{i+2} + T_i - \frac{T_i}{n_i} \\ &= p \cdot T_{i-2} + T_{i-1} + T_{i+1} + p \cdot T_{i+2} + T_i \end{aligned} \tag{6}$$

From Eq. 6, an interesting property appears: the average number of operations per sensor per time unit is independent of the number of sensors deployed in the virtual node. Thus, a simple solution is to consider the same number of sensors per virtual node; this corresponds to a uniform deployment. Under the assumption of the radio always on, the optimal solution is to deploy one sensor in each virtual node, since it is useless to deploy more sensors in a virtual node.

3.3.2 Case 2: Optimal scheduling

This section assumes an optimal scheduling algorithm where a sensor is able to switch its radio on only when there is a message to forward or only when it has a new message to transmit. If such a scheduling algorithm is used,

there will be no overhearing and the messages received by a sensor will be only those it needs to relay for the previous virtual nodes. The traffic relayed by sensors in the virtual node B_i is $T_i - \lambda$. Since we assume a uniform distribution of this traffic among sensors in the virtual node B_i, the average number of receptions per sensor is given by Eq. 7.

$$\overline{R}_i^{os} = \frac{T_i - \lambda}{n_i} \tag{7}$$

Thus, the average number of operations per sensor per time unit is the sum of transmissions (Eq. 4) and receptions (Eq. 7) and is given by Eq. 8

$$\overline{O}_i = \overline{T}_i + \overline{R}_i^{os} = \frac{T_i}{n_i} + \frac{T_i - \lambda}{n_i} = \frac{2 \cdot T_i - \lambda}{n_i} \tag{8}$$

From Eq. 8, we observe that, when the number of sensors n_i increases, $\overline{O}_i$ decreases. Thus by increasing the number of sensors deployed in a virtual node, the network functioning time increases too.

3.3.3 Case 3: Practical scheduling

In a wireless network, because of multiple constraints (environment, hardware, communication technology, etc), scheduling algorithms [27–29] cannot be perfect. A sensor may receive a message addressed to another one: this is the overhearing phenomenon. Thus, if we consider a particular sensor s deployed in the virtual node B_i, in addition to the messages transmitted by previous virtual nodes and relayed by s, s may receive other transmissions in its communication range. We model this effect by a parameter α. This parameter represents the ratio of messages transmitted in the communication area of a sensor s, not addressed to s, but received by s. In the following, α will be called the overhearing ratio. To calculate the total number of messages that s can receive due to the overhearing phenomenon, the following components must be considered:

- Messages transmitted (generated or relayed) by sensors in the virtual node B_{i-1}, but relayed by sensors in the virtual node B_{i+1}

$$\overline{R}_i^{rsnr} = p \cdot T_{i-1} \tag{9}$$

- Messages transmitted (generated or relayed) by previous virtual nodes, relayed by sensors in the virtual node B_i, but a different node from s

$$\overline{R}_i^{rsr} = (T_i - \lambda) - \frac{T_i - \lambda}{n_i} = \frac{(T_i - \lambda) \cdot (n_i - 1)}{n_i} \tag{10}$$

- Messages transmitted by other sensors in the virtual node B_i:

$$\overline{R}_i^{rsc} = T_i - \frac{T_i}{n_i} = T_i \cdot \frac{n_i - 1}{n_i} \tag{11}$$

- Messages transmitted by sensors in virtual nodes B_{i+1} and B_{i+2}:

$$\overline{R}_i^{rsf} = T_{i+1} + p \cdot T_{i+2} \tag{12}$$

Thus, the number of messages received by sensor s due to overhearing is defined by Eq. 13

$$\begin{aligned} \overline{R}_i^{o} &= \alpha \cdot (\overline{R}_i^{rsr} + \overline{R}_i^{rsnr} + \overline{R}_i^{rsf} + \overline{R}_i^{rsc}) \\ &= \alpha \cdot \left[\frac{(T_i - \lambda) \cdot (n_i - 1)}{n_i} + p \cdot T_{i-1} \right. \\ &\quad \left. + T_{i+1} + p \cdot T_{i+2} + T_i \cdot \frac{n_i - 1}{n_i} \right] \\ &= \alpha \cdot \left[\frac{(2 \cdot T_i - \lambda) \cdot (n_i - 1)}{n_i} + p \cdot T_{i-1} \right. \\ &\quad \left. + T_{i+1} + p \cdot T_{i+2} \right] \end{aligned} \tag{13}$$

To obtain the total number of receptions of s, we also consider the messages transmitted from previous virtual nodes and relayed by s (see Eq. 7). Eq. 14 gives the average number of receptions per sensor deployed in a virtual node B_i.

$$\begin{aligned} \overline{R}_i^{ps} &= \overline{R}_i^{os} + \overline{R}_i^{o} = \frac{T_i - \lambda}{n_i} + \alpha \cdot \left[\frac{(2 \cdot T_i - \lambda) \cdot (n_i - 1)}{n_i} \right. \\ &\quad \left. + p \cdot T_{i-1} + T_{i+1} + p \cdot T_{i+2} \right] \end{aligned} \tag{14}$$

The average number of operations per sensor per time unit is then given by the Eq. 15.

$$\begin{aligned} \overline{O}_i &= \overline{T}_i + \overline{R}_i^{ps} \\ &= \frac{T_i}{n_i} + \frac{T_i - \lambda}{n_i} + \alpha \cdot \left[\frac{(2 \cdot T_i - \lambda) \cdot (n_i - 1)}{n_i} \right. \\ &\quad \left. + p \cdot T_{i-1} + T_{i+1} + p \cdot T_{i+2} \right] \\ &= \frac{(2 \cdot T_i - \lambda)}{n_i} + \alpha \cdot \left[\frac{(2 \cdot T_i - \lambda) \cdot (n_i - 1)}{n_i} \right. \\ &\quad \left. + p \cdot T_{i-1} + T_{i+1} + p \cdot T_{i+2} \right] \end{aligned} \tag{15}$$

Note that, the optimal scheduling corresponds to the case $\alpha = 0$, whereas the radio always on corresponds to the case $\alpha = 1$. It is also important to note that $\overline{O}_i$ decreases when n_i increases. But when n_i is very large: $\overline{O}_i \approx \overline{O}_i^* = \alpha \cdot (2 \cdot T_i - \lambda + p \cdot T_{i-1} + T_{i+1} + p \cdot T_{i+2})$. Thus, if $n_i = n_i^*$ then $\overline{O}_i \approx \overline{O}_i^*$, it is useless to deploy more than n_i^* sensors in the virtual node B_i, since this will not improve the lifetime of sensors in B_i.

3.4 Problem formulation

Given N sensors that form K virtual nodes, the goal of this paper is to find the number n_i of sensors in virtual node B_i, $1 \leq i \leq K$, such that $\sum_{i=1}^{K} n_i = N$, and which minimizes the maximum number of operations executed by a sensor. Formally, the problem can be formulated as:

$$minimize \quad \max_{1 \leq i \leq K} \overline{O}_i \tag{16}$$

subject to the following constraints:

$$\sum_{i=1}^{K} n_i = N \tag{17}$$

$$n_i \geq 1, i = 1, \ldots, N \tag{18}$$

Equations 16–18 define a mixed-integer nonlinear programming problem. In the following section, given the number N of sensors, we propose a greedy algorithm to calculate, for each virtual node B_i, the value of n_i.

4 Deployment algorithm

This work assumes that a sensor is declared dead when it exhausts its battery. This paper assumes a uniform load distribution within a virtual node. This means that all sensors in the same virtual node will die approximately at the same time. Thus, we define the network lifetime as the time until all sensors in a virtual node die (or, equivalently, until the first sensor dies). The network lifetime is maximized by reducing as much as possible the traffic load per sensor. The problem addressed in this paper can be formulated as follows: given N sensors forming K virtual nodes, how many sensors should be deployed in each virtual node in order to maximize the network lifetime? Hereafter, the average number of operations per sensor per time unit in the virtual node B_i is denoted $\overline{O_i}$ (see Eq. 15). We propose in Algorithm 1 a greedy approach to compute the number of sensors to deploy in each virtual node.

Algorithm 1 Deployment algorithm

Require: $N, K, \lambda, \alpha, p, N \geq K$
Ensure: $n_i, 1 \leq i \leq K, n_i \geq 1, \sum_{i=1}^{K} n_i = N$
1: **for** Each virtual node B_i **do**
2: calculate T_i using Eq. 3
3: **end for**
4: $n_i \leftarrow 1, 1 \leq i \leq K$
5: $remaining\_sensors \leftarrow N - K$
6: **while** $remaining\_sensors \neq 0$ **do**
7: **for** Each virtual node B_i **do**
8: calculate $\overline{O_i}$ using Eq. 15
9: **end for**
10: $I \leftarrow \{i | \overline{O}_i = \max_j \overline{O}_j\}$
11: $i \leftarrow max\{I\}$
12: $n_i \leftarrow n_i + 1$
13: $remaining\_sensors \leftarrow remaining\_sensors - 1$
14: **end while**

Recall that, in each virtual node, we will have at least one sensor, for connectivity and sensing purposes. That is why the number N of sensors is greater than or equal to the number K of virtual nodes. Initially, one sensor is assigned to each virtual node (Line 4) and the remaining sensors are iteratively deployed. At each iteration, one sensor is added to the virtual node which has the highest $\overline{O}_i$. If more than one virtual node have the maximum number of operations per sensor per time unit, the one with the highest index is selected (Lines 10-11): it is the one which is closest to the sink. Indeed, the traffic is highest as the virtual node is closer to the sink. The consequence of adding a sensor to a virtual node B_i is the reduction of the average number of operations per sensor per time unit in that virtual node. Since our objective is to propose a deployment which minimizes the maximum number of operations per sensor, the problem can be decomposed in a series of sub-problems, which are optimally solved iteratively by the greedy Algorithm 1, resulting in an optimal general solution.

At the beginning of this algorithm, the traffic of each virtual nodes is calculated. We assume that, for a virtual node B_i, T_i is calculated in O(K). Thus, calculating this value for all virtual nodes is done in O(K^2). In the iterative part of the algorithm, at each iteration, the average number of operations per sensor per virtual node is calculated. This processing can be done in O(K). Since the goal of the iterative part of the deployment algorithm is to find the virtual node in which each of the $N - K$ (since initially one sensor is deployed per virtual node) virtual nodes will be deployed, the iterative part of our algorithm can be executed in O($N \cdot K$). Thus, the result of the deployment is computed in O($K^2 + N \cdot K$). Even for a large network with a high number of sensors, it is possible to rapidly obtain the number of sensors which forms each virtual node, since this algorithm can be executed on a computer with good performance, and not on the nodes themselves.

5 Analytical results

This section presents analytical results concerning the virtual node-based approach proposed in this paper. These results concern the greedy deployment described by Algorithm 1 and a simple uniform deployment. In the later one, the same number of sensors are deployed in all virtual nodes. Firstly (see Sect. 5.2), we compare the two deployments in terms of sensors distribution for different parameters. Secondly (see Sect. 5.3), the two deployments are compared in terms of network lifetime. The next section describes the evaluation parameters.

5.1 Evaluation parameters

Without loss of generality, it is assumed that the average output power is $P = 61.9$ mW, which is the maximum transmission power of the radio used by *Tmote Sky* [8]. It is also assumed that each sensor has a battery capacity of $E_{node} = 5$ Ah. If we assume a packet size of 128 Bytes and a data rate of 250 kbps, the time τ to transmit one packet is equal to 4.096 ms. Thus, E_i, the average energy consumed by a sensor in the virtual node B_i per time unit is defined by Eq. 19.

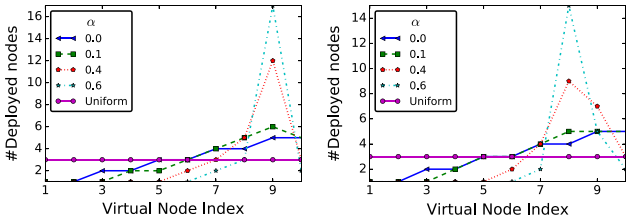
$$E_i = \overline{O}_i \cdot \tau \cdot P \tag{19}$$

Table 2 summarizes the values of parameters we use to obtain numerical results.

5.2 Spatial sensors distribution

Figure 3 presents the sensors distribution obtained by a uniform deployment and the greedy Algorithm 1. It gives the number of sensors forming each virtual node as a function of the virtual node position and the overhearing ratio (α) when we consider $N = 30$ sensors to deploy in a network of $K = 10$ virtual nodes. Note that, when the number of virtual nodes is 10, the sink is deployed alone in a particular virtual node (not represented in Fig. 3) at the position $K + 1$. Figure 3a, b are for $p = 0$ and $p = 0.5$ respectively (p being the probability of a sensor in the virtual node B_i to receive messages from B_{i-2} or B_{i+2}).

With the greedy deployment proposed by Algorithm 1, we observe that there is always a virtual node B_{i*} which receives the highest number of sensors. The position of this virtual node depends on p and is independent of α. When $p = 0.0$, B_9 is the bottleneck of the network (highest traffic) and receives the maximum number of sensors (Fig. 3a). Indeed, unlike sensors in B_{10} which receive messages only from the left side, sensors in B_9 receive messages from both the left and the right side. When $p = 0.5$ (Fig. 3b), the virtual node B_8, which receives messages from "two-hop" virtual nodes, becomes the bottleneck of the network and is assigned the maximum number of sensors. For other virtual nodes which are farther from the sink than B_{i*}, the number of sensors deployed in a virtual node increases linearly as a function of the virtual node index.



(a) $p = 0.0$ **(b)** $p = 0.5$

Fig. 3 Number of nodes deployed in each virtual node: $N = 30, K = 10$

When only transmissions are considered, like in most of the related work, the last virtual node of the network will always receive the maximum number of sensors. Since we consider transmissions and receptions of messages, our results show a different trend. In Fig. 3, we also observe that, when the overhearing ratio (α) increases, the number of sensors deployed in the virtual node B_{i*} increases too. Indeed, the traffic is highest in this virtual node and a high value of α results in a high number of receptions.

In summary, when the overhearing ratio is not taken into account ($\alpha = 0$), the number of sensors per virtual node is almost a linear increasing function of the virtual node index. When the overhearing is taken into account ($\alpha > 0$), the number of sensors deployed per virtual node increases until a virtual node B_{i*} which receives the highest number of sensors, and then drastically decreases. The position of B_{i*} depends on p. If $p = 0, B_{i*} = B_{K-1}$ and when $p > 0, B_{i*} = B_{K-2}$.

Table 2 Evaluation parameters

Parameter	Description	Value
E_{node}	The sensor battery capacity	5 Ah
P	Average transmission power	61.9 mW
τ	Time to transmit one message	4.096 ms
Δ	Time slot duration	5 s
λ	Avg. number of messages generated per virtual node per time slot	1 message

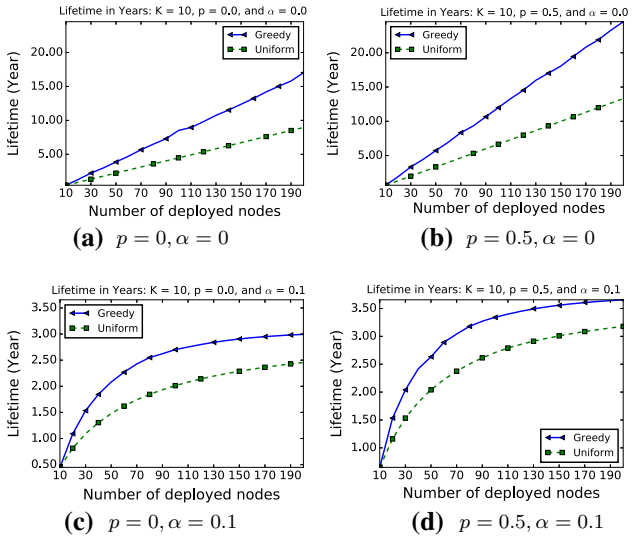
5.3 Lifetime: comparison with uniform deployment

In a uniform deployment, the same number of sensors is deployed in all virtual nodes. We consider as our baseline the network lifetime $LT^{(uniform)}$ obtained with such a deployment. Since we assume that the traffic is balanced between sensors in a virtual node, the lifetime of a virtual node is equal to the average lifetime of sensors deployed in that virtual node. We denote by $E_{max} = \max \{E_i \mid i = 1, \ldots, K\}$ the maximum energy consumed by a sensor per time unit in the network. Recall that E_i is the energy consumption rate per sensor in the virtual node B_i (see Eq. 19). If E_{node} is the initial energy of a sensor, for a given deployment, the lifetime of the network will be equal to $\frac{E_{node}}{E_{max}} \cdot \Delta$.

5.3.1 The lifetime of the greedy and uniform deployment

Figure 4 presents the lifetime of the greedy and the uniform deployment when $p = 0$ and $\alpha = 0$ (Fig. 4a), $p = 0.5$ and $\alpha = 0$ (Fig. 4b), $p = 0$ and $\alpha = 0.1$ (Fig. 4c) and $p = 0.5$ and $\alpha = 0.1$ (Fig. 4d). The results presented in this figure show that the greedy deployment, in terms of lifetime, outperforms the uniform one. Indeed, in the greedy deployment, more sensors are deployed in virtual nodes with higher load while, in the uniform deployment, the same number of sensors are deployed in all virtual nodes, no matter their relative position to the sink.

When overhearing is not taken into account (Fig. 4a, b), we observe a significant increase in lifetime (more than 20 years) compared to a network with an overhearing ratio of 0.1 (about 3.5 years). This means that if we are able to use a perfect activity scheduling which reduces the overhearing, the network lifetime can increase significantly. Of course, this perfect scheduling would most likely require some supplementary control traffic which is not considered in our study.



(a) $p = 0, \alpha = 0$ **(b)** $p = 0.5, \alpha = 0$

(c) $p = 0, \alpha = 0.1$ **(d)** $p = 0.5, \alpha = 0.1$

Fig. 4 Impact of the number of nodes on the lifetime for different values of α and p

In Fig. 4, the lifetime increases with the number of sensors. Increasing the total number of sensors in the network means to increase the number of sensors deployed per virtual node, both for the greedy and the uniform deployment. A direct consequence of increasing the number of sensors per virtual node is the reduction of the average number of transmissions per sensor, and thus prolonging the network lifetime. When the value of p increases, the lifetime increases too. Indeed, since we assume a shortest-path routing, a high value of p means a high probability for sensors in virtual node B_i to relay traffic from B_{i-2}. Thus, by increasing the value of p, the amount of traffic relayed by B_i from B_{i-1} is reduced (See Eq. 1).

When $\alpha = 0$, the average number of operations per sensor per time unit tends to 0 (see Eq. 15) when the number of sensors deployed in the virtual node is very large. That is why, in Fig. 4a, b, the lifetime is a strictly increasing linear function of the number of sensors. On the other hand, when overhearing is taken into account ($\alpha > 0$), the number of operations per node per time unit in the virtual node B_i when the number of sensors is very large tends to $\alpha \cdot (2 \cdot T_i - \lambda + p \cdot T_{i-1} + T_{i+1} + p \cdot T_{i+2})$. Thus, when this limit is reached, it is useless to deploy more sensors. That is why, in Fig. 4c, d, we observe a threshold which is asymptotically reached starting from a given number of sensors. We can also notice in this figure that the gap between the lifetime of the greedy and the uniform deployment is reduced when more sensors are deployed. Indeed, the greedy deployment rapidly reaches the maximum lifetime compared to the uniform deployment.

5.3.2 The lifetime gain of the greedy deployment

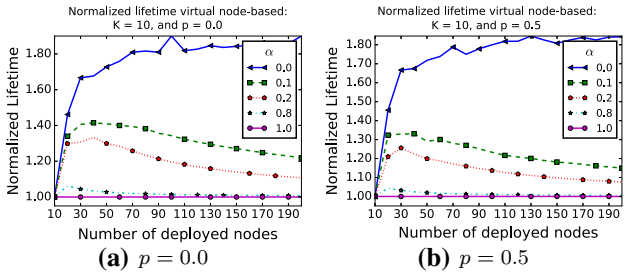
To illustrate how much the network lifetime could be extended when using the greedy Algorithm 1 rather than a simple uniform deployment, we normalize the lifetime of the greedy deployment by that of the uniform one. Thus, we use as metric the ratio r expressed by Eq. 20.

$$r = \frac{LT^{(greedy)}}{LT^{(uniform)}} \tag{20}$$

Fig. 5 shows, for different values of α, the normalized lifetime of the network as a function of the number of sensors when the number of virtual nodes (K) is equal to 10. We consider the cases $p = 0$ and $p = 0.5$ in Fig 5a, b, respectively.

The normalized lifetime decreases when α increases. When there is no overhearing ($\alpha = 0$), the normalized lifetime increases continuously and reaches its maximum value around $N = 90$ with an improvement of the lifetime by up to 80% compared to a uniform deployment. When there is no scheduling ($\alpha = 1$), i.e. all sensors always keep their radio on, the solution obtained with the greedy algorithm is equivalent to that given by a uniform deployment. This observation is consistent with our model and hypothesis, since Eq. 6, which expresses the average number of operations per node per time unit in a virtual node, shows that, when $\alpha = 1$, $\overline{O}_i$ is independent of the number of sensors deployed in the virtual node B_i.

When $0 < \alpha < 1$, we observe, in Fig. 5, a considerable decrease of the performance compared to the case $\alpha = 0$. Nevertheless, the lifetime of the uniform deployment can be improved by up to 40% (Fig. 5a, $p = 0$) or 30% (Fig. 5b, $p = 0.5$), depending on the overhearing ratio. As expected, increasing the overhearing ratio (α) also increases the number of receptions in the network, and then reduces the lifetime of the network. We highlight that, from a given number of sensors, the performance of the greedy deployment compared to the uniform deployment starts decreasing. Indeed, as presented in the Fig. 4, from a given number of sensors, our greedy deployment reaches its maximum lifetime while the lifetime of the uniform deployment still increases, and thus, the gap between the greedy and the uniform deployment is reduced.



(a) $p = 0.0$ **(b)** $p = 0.5$

Fig. 5 Normalized lifetime, impact of the number of sensors

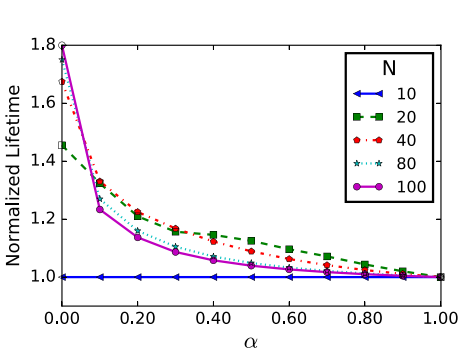


Fig. 6 Impact of α: $K = 10$ and $p = 0.5$

Figure 6 presents the normalized network lifetime as a function of α and the number of deployed sensors. When the overhearing phenomenon is taken into account (i.e. $\alpha > 0$), the gain (compared a deployment with $\alpha = 0$) is reduced. Thus, when designing a deployment strategy or when evaluating the performance of communication protocols, it is important to take into account all messages that can be received by sensors. For values of α close to 1, the greedy deployment is equivalent (or close) to a uniform deployment in terms of network lifetime.

5.3.3 Impact of the network length

Figure 7 shows the impact of the network length, in terms of the number of virtual nodes. For a given number of sensors ($N = 200$), this figure shows the impact of the network length (K) and the overhearing ratio (α) on the deployment proposed in this paper. Results presented in this figure show that the gap between the greedy deployment and the uniform one increases with K. Indeed, by increasing the K value, the overall network traffic in the network increases too, and particularly in virtual nodes close to the sink. However, considering a uniform deployment, increasing the network length while keeping a constant number of sensors (as it is the case here) reduces the number of sensors to form each virtual node. A direct consequence is a negative impact on the network lifetime of the uniform deployment. It is not the case for the greedy deployment, since this algorithm deploys sensors in virtual nodes taking into account network traffic in each area.

5.4 Residual energy

In this paper, we define the network lifetime as the time until the first sensor exhausts its energy. Considering a LWSN organized in $K = 10$ virtual nodes, we present in this section the residual energy of sensors per virtual nodes at the end of the network functioning. In Fig. 8, results are given for $N = 30, p = 0.0$ (Fig. 8a), $p = 0.5$ (Fig. 8b) and

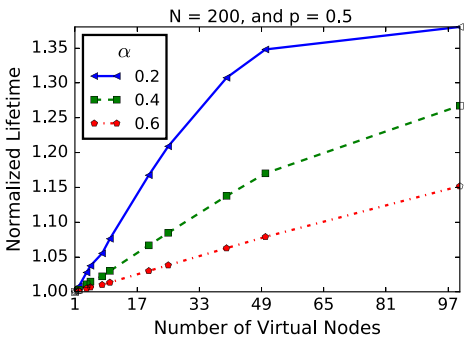
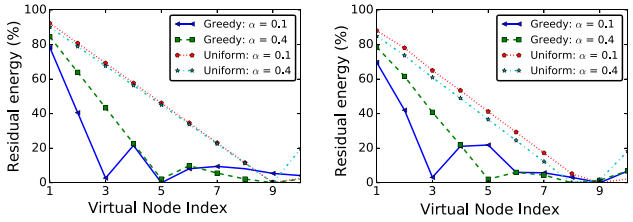


Fig. 7 Impact of the network length (K) for a network of $N = 200$ sensors

for different values of α. An important observation concerning the results presented in this figure is the highest residual energy of sensors for the uniform deployment. While the same number of sensors are deployed in each virtual nodes with the uniform deployment, our greedy deployment deploys more sensors in the area close to the sink. Thus, for virtual nodes far from the sink, the traffic per sensor is lower for the uniform deployment since it deployed more sensors compared to the greedy deployment proposed in this paper. This results in lower energy consumption for these sensors in the case of the uniform deployment compared to the greedy one. In many WSN, it is not possible to replace the sensor battery. In this case, when the network stops functioning, all the sensors are replaced. But, if only sensors with less energy where to be replaced, these results show that the battery of many sensors will be replaced for the greedy deployment than that for the uniform case. But, given the low-cost of sensors and the lifetime gain (see Fig. 5) of the greedy deployment compared to the uniform one, this is not a problem.

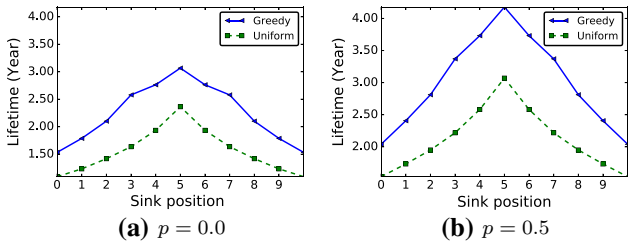
5.5 Impact of the sink position

So far, we assumed that the sink is deployed at position $K+1$. In this section, we evaluate the impact of the sink position on the network lifetime considering the uniform deployment as well as the greedy one proposed in this paper. We vary the sink position from 0 to $K+1$.



(a) $p = 0.0$ **(b)** $p = 0.5$

Fig. 8 Residual energy per sensor as function of the virtual node: $K = 10, N = 30$



(a) $p = 0.0$ **(b)** $p = 0.5$

Fig. 9 Lifetime of the uniform and greedy deployments as function of the sink position: $K = 10, \alpha = 0.1, N = 30$

Figure 9 shows the lifetime of the network considering the uniform and greedy deployments as function of the sink position. We consider a network formed of 10 virtual nodes, 30 sensors and $\alpha = 0.1$. Results are shown for $p = 0$ (Fig. 9a) and $p = 0.5$ (Fig. 9b). From the network traffic point a view, the sink deployed at position 0 or $K+1$ corresponds to the configuration considered in the previous section and gives the same network lifetime. But, by deploying the sink at a different position than the network edges, the among of traffic forwarded by the closest neighbors of the sink is reduced. And as shown in Fig. 9, the optimal position is the middle of the network.

6 Comparison to related work

In this section, we compare the deployment strategy proposed in this paper to the one proposed in [8] assuming that the sink is deployed in the virtual node $K+1$. In the following, the solution proposed in this paper will be denoted as *virtual node-based* and the one proposed in [8] as *distance-based*. Since these two approaches are quite different, it is not easy to provide a fair comparison. In the following section, we start by describing the solution proposed by [8].

6.1 Distance-based deployment

In [8], the deployment is formulated as a Mixed-Integer Linear Programming (MILP) problem. In their formulation, the authors denote by d_i the distance between sensors S_i and S_{i+1}, and R_j the communication range when the transmission power P_j is used. For each sensor S_i, m binary variables $(x_{i1}, x_{i2}, \ldots, x_{ij}, \ldots, x_{im})$ are defined to denote the assignment of power level for S_i. Thus :

$$\sum_{j=1}^{m} x_{ij} = 1 \tag{21}$$

$$d_i = \sum_{j=1}^{m} x_{ij} \cdot R_j \tag{22}$$

The authors assume that S_i relays all the messages from S_{i-1}. If μ is the number of messages generated by a sensor per time unit, the number of operations (transmissions, since in [8] receptions are not taken into account) per sensor per time unit is defined by Eq. 23.

$$\overline{O}_i^{(r)} = \mu \cdot i \tag{23}$$

If we consider receptions and the overhearing phenomenon, as in this paper, the number of operations per sensor per time unit is now equal to:

$$\begin{aligned}\overline{O}_i^{(r)} &= \mu \cdot (i-1+i+\alpha \cdot (i+1)) \\ &= \mu \cdot (2 \cdot i+\alpha \cdot (i+1)-1)\end{aligned} \tag{24}$$

The energy consumed by S_i per time slot is then equal to:

$$E_i = \overline{O}_i^{(r)} \cdot \tau \cdot \left(\sum_{j=1}^{m} x_{ij} \cdot P_j \right) \tag{25}$$

The object in [8] is to:

$$minimize \quad max\{E_i \mid i = 1, \ldots, N\} \tag{26}$$

subject to the following constraint:

$$\sum_{i=1}^{N} x_{ij} \cdot d_i \geq L \tag{27}$$

where L is the initial network length, a parameter defined in [8]. We use the CPLEX software package to find the optimal solution to this problem. It is important to note here that, in the solution provided by the solver, the total area covered by sensors is larger than L (see Eq. 27). In the following, we denote by $L_{simulated}$ the (final) length of the network provided by the solver.

In our work, we form a network of virtual nodes and we propose a greedy algorithm which calculates the number of sensors to deploy in each virtual node. The MILP problem proposed by [8] and described in this section (Eqs. 26–27) considers a flat network and expresses the deployment results in terms of distance between consecutive nodes.

6.2 Comparison framework

To provide a fair comparison, we use the same parameter values as the ones used in [8] to solve the problem described by Eq. 26 and 27. In our virtual node-based approach, without loss of generality, we assume as the average transmission power P, the maximum power level used in the distance-based approach. Thus, the area covered by a virtual node is of length R_{max}, the transmission range when the maximum transmission power is used. Since [8] assumes that S_i relays all traffic from S_{i-1}, we set $p = 0$, i.e communications are not allowed between virtual nodes B_i and B_{i+2}.

In the virtual node-based and the distance-based approaches, traffic models are different. In our evaluation, we guarantee the same average messages generation per sensor in both approaches. If, for the distance-based approach, μ is the average number of messages generated per sensor per time slot, the total number of messages generated in the network of length $L_{simulate}$ per time slot is $\mu \cdot N$, where N is the total number of sensors. Thus, in our virtual node-based approach, we assume that λ, the average number of messages generated per virtual node per time slot is equal to $\mu \cdot N \cdot \frac{R_{max}}{L_{simulated}}$. Comparison parameters are summarized in Table 3.

6.3 Comparison results

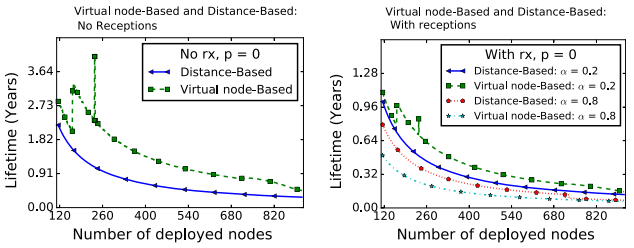
Figure 10 presents the lifetime of the network for distance-based and virtual node-based deployments. In Fig. 10a, only transmissions are considered, while Fig. 10b also takes into account receptions with and overhearing ratio of 0.2 and 0.8. In the evaluations, we assume that $\Delta = 1$ minute and $\mu = 1/\Delta$. As it might be expected, considering messages reception (Fig. 10b), the lifetime is divided by more than two compared to the model in which only transmissions are taken into account (Fig. 10a). When receptions are not considered (Fig. 10a), the virtual node-based approach gives better performance compared to the distance-based one. This advantage of the virtual node-based approach is also noticed for low and medium overhearing ($\alpha = 0.2$ in Fig. 10b). In the virtual node-based approach, all sensors use the maximum power level. Thus, sensors are more exposed to the overhearing phenomenon. That is why, in Fig. 10b, the distance-based approach gives better performance when $\alpha = 0.8$.

When the number of sensors in the network increases, the lifetime of the two approaches decreases. Indeed, deploying more sensors in the network has an impact on the two approaches. For the distance-based solution, this increases the traffic in the network (since each node generates messages) and the network length. When the network length increases, the number of virtual nodes increases too. Moreover, increasing the sensors density allows more sensors in the distance-based approach to use lower power level. That is why we observe a reduction of the gap between the two approaches in Fig. 10.

In Fig. 10, we observe a significant increase of the performance of our virtual node-based approach at some particular values of N. This is an artifact of the discrete power level model used in [8], where 6 power levels are possible. By increasing the number of sensors, smaller power levels are used with the distance-based approach. Thus, when a sufficient number of sensors is deployed, all sensors use the next lowest power level [8]. At these points, the resulting network length $L_{simulated}$ is almost equal to L, which reduces the number of virtual nodes and favors our solution.

Table 3 Virtual node-based and distance-based comparison parameters

Parameter	Distance-based	Virtual node-based
Communication range	R_{min} to R_{max}	R_{max}
Number of virtual nodes	–	$K = \lceil \frac{L_{simulate}}{R_{max}} \rceil$
Messages generation rate	μ per node	$\mu \cdot N \cdot \frac{R_{max}}{L_{simulated}}$ per virtual node



(a) No Receptions **(b)** With Receptions

Fig. 10 Network lifetime of distance-based and virtual node-based deployment: $L = 5$ km

7 Investigating the case of heterogeneous sensors

In the previous sections, we assumed we have a large number of homogeneous sensors. We divided the network into virtual nodes and then we proposed a greedy algorithm to compute the number of sensors to deploy in each virtual node in order to extend the network lifetime. In this section, the network is still divided virtual nodes, but only one sensor is deployed into each virtual node. We assume that sensors are heterogeneous in terms of battery capacity, and our goal is to find the battery capacity to assign to a given virtual node, taking into account its traffic, in order to extend the network lifetime. We also assume in this section that the sink is deployed in the virtual node $K + 1$.

7.1 Problem formulation

We assume that we have an energy budget E_{net}, and our goal is to distribute this energy to all sensors, in order to maximize the network lifetime. We assume in this section that one sensor is deployed in each virtual node (i.e. $K = N$). If we replace $n_i = 1$ in Eq. 15, we will have a new model of the average number of operations per sensor per time unit, defined by the Eq. 28:

$$\overline{O}_i = (2 \cdot T_i - \lambda) + \alpha \cdot [p \cdot T_{i-1} + T_{i+1} + p \cdot T_{i+2}] \quad (28)$$

We denote by ξ_i the energy assigned to virtual node B_i. Since E_i (Eq. 19) is the energy consumption rate of the virtual node B_i, the lifetime of this virtual node, $LT_i^{(energy)}$, is defined by Eq. 29:

$$LT_i^{(energy)} = \frac{\xi_i}{E_i} \quad (29)$$

If we define the network lifetime (LT) as the time until the first virtual node exhausts its energy, it will be defined by Eq. 30:

$$LT^{(energy)} = min \quad \{LT_i^{(energy)} | i = 1, \ldots, K\} \quad (30)$$

Our goal is to find $\xi_{i,1 \le i \le K}$ which maximizes $LT^{(energy)}$. Formally, the problem can be presented as:

$$maximize \quad LT^{(energy)} \quad (31)$$

subject to the following constraints:

$$\sum_{i=1}^{i=K} \xi_i = E_{net} \quad (32)$$

$$\xi_i > 0 \quad (33)$$

Equations 31–33 define a linear optimization problem. We find the optimal solution to this problem using the CPLEX software package.

Intuitively, the energy budget E_{net} could be distributed among virtual nodes proportionally to their traffic. Thus, alternatively, we propose Eq. 34 to model the energy assigned to virtual node B_i.

$$\xi_i = \frac{\overline{O}_i}{\sum_{j=1}^{j=K} \overline{O}_j} \cdot E_{net} \quad (34)$$

In Eq. 34, $\overline{O}_i$ is the average number of operations per sensor per time unit in the virtual node B_i. Figure 11 presents the optimal solution (provided by CPLEX software) to the problem described by Eqs. 31–33 and the solution

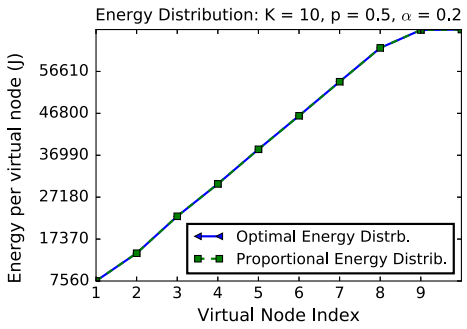


Fig. 11 Optimal and proportional energy distribution per virtual node for an energy budget $E_{net} = 408240$ J

given by Eq. 34. This figure shows that the two solutions are almost identical.

Recall that $E_i = \overline{O}_i \cdot \tau \cdot P$. Thus, $LT_i^{(energy)}$ is defined by Eq. 35:

$$LT_i^{(energy)} = \frac{\xi_i}{E_i} = \frac{\overline{O}_i}{\sum_{j=1}^{j=K} \overline{O}_j} \cdot E_{net} \cdot \frac{1}{\overline{O}_i \cdot \tau \cdot P} = \frac{E_{net}}{(\sum_{j=1}^{j=K} \overline{O}_j) \cdot \tau \cdot P} \tag{35}$$

Note in Eq. 35 that all sensors have the same lifetime and thus, will exhaust their battery at the same moment.

7.2 Uniform versus proportional energy distribution

For a uniform energy distribution, the same amount of energy is assigned to each virtual node. We define $\xi_i^{(u)}$, the energy assigned to each node, given by Eq. 36.

$$\xi_i^{(u)} = \frac{E_{net}}{K} \tag{36}$$

And $LT_i^{(uniform)}$, the lifetime of the virtual node B_i for a uniform energy distribution, is defined by Eq. 37.

$$LT_i^{(uniform)} = \frac{\xi_i^{(u)}}{E_i} = \frac{E_{net}}{K} \cdot \frac{1}{\overline{O}_i \cdot \tau \cdot P} = \frac{E_{net}}{K \cdot \overline{O}_i \cdot \tau \cdot P} \tag{37}$$

If the lifetime of the optimal energy distribution is normalized by the lifetime given by the uniform energy distribution, we obtain:

$$\frac{LT_i^{(energy)}}{LT_i^{(uniform)}} = \frac{E_{net}}{(\sum_{j=1}^{j=K} \overline{O}_j) \cdot \tau \cdot P} \cdot \frac{K \cdot \overline{O}_i \cdot \tau \cdot P}{E_{net}} = \frac{\overline{O}_i \cdot K}{\sum_{j=1}^{j=K} \overline{O}_j} \tag{38}$$

Thus, if different energy capacities can be assigned to sensors, the optimal capacity assignment (Eq. 34) can improve the lifetime of the virtual node B_i by a factor of $\frac{\overline{O}_i \cdot K}{\sum_{j=1}^{j=K} \overline{O}_j}$ compared to a uniform energy distribution. If we denote by $\overline{O}_{max} = \max \{\overline{O}_i \mid i = 1, \ldots, K\}$, the maximum number of operations per sensor per time unit in the network, we obtain: $\overline{O}_{max} \cdot K > \sum_{j=1}^{j=K} \overline{O}_j$ and $\frac{\overline{O}_{max} \cdot K}{\sum_{j=1}^{j=K} \overline{O}_j} > 1$.

7.3 Deployments based on heterogeneous and homogeneous sensors

Our goal here is to compare the deployment based on energy distribution solution proposed in this section to the one based on virtual nodes described by Algorithm 1. Recall that, given N sensors, this greedy algorithm finds the number of sensors to deploy in each virtual node. To allow a fair comparison, we assume that the energy unit is equal to E_{node}, and an energy budget of E_{net} can be discretized to have $N = \frac{E_{net}}{E_{node}}$ homogeneous sensors. Thus, given the energy budget, we can fairly compare the energy distribution deployment to the sensors distribution one by discretizing the energy budget and implementing the two solutions.

Figure 12 presents $\frac{LT^{(energy)}}{LT^{(greedy)}}$ as a function of $\frac{E_{net}}{E_{node}}$ (recall that $LT^{(greedy)}$ is the network lifetime obtained by Algorithm 1). This figure shows that the energy distribution solution outperforms the sensor distribution one and the gap increases with the ratio $\frac{E_{net}}{E_{node}}$. It is also important to note that, by deploying a single sensor in each virtual node, the routing problem is simplified and there is no load balancing problem of the traffic per virtual node as in the case of the solution proposed in the previous sections. However, it may not be economically or practically feasible for industrials to produce sensors with different batteries.

8 Conclusion

In this paper, we studied the problem of sensors deployment to maximize the lifetime of WSN with linear topology. To simplify the deployment, we divided the network into virtual nodes. By considering messages transmission and/or reception as energy operations consuming, we proposed an analytical model of the number of operations per virtual node. Unlike existing works, our model considers the message reception cost. It also takes into account messages received due to the overhearing phenomenon.

Considering homogeneous sensors in terms of battery capacity, we proposed a greedy algorithm to compute the

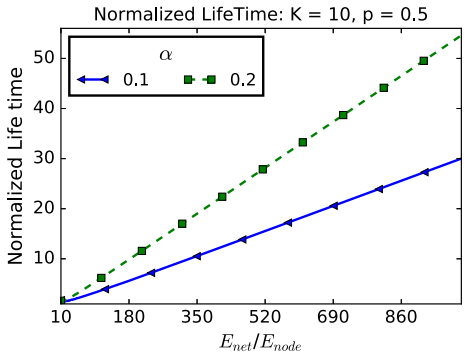


Fig. 12 Energy distribution versus virtual nodes-based: impact of the energy budget when $K = 10, p = 0.5$

number of sensors to deploy in each virtual node. Performance evaluation shows that, depending on the number of sensors, the simple virtual node-based deployment proposed can improve the network lifetime by up to 40%, when compared to the uniform deployment. Moreover, it outperforms the distance-based approach when a scheduling algorithm which reduces the overhearing phenomenon is used.

Finally, we studied an alternative solution. We considered a network in which the batteries of sensors are nonuniform, i.e. the battery of a sensor is proportional to its traffic. Results show that, by properly selecting the battery to assign to each sensor, such an approach significantly improves the network lifetime.

Our model assume an average transmission power used by all sensors in a virtual node. However, existing radio modules provide many transmission power levels that can be configured by programming. In our future work, we plan to integrate this into our model by defining the output power level to be used by each sensor deployed in a virtual node.

Acknowledgements Funding was provided by Agence Universitaire de la Francophonie, Institut National des Sciences Appliquées de Lyon and Universite de Yaounde I.

References

1. Jawhar, I., Mohamed, N., & Agrawal, D. P. (2011). Linear wireless sensor networks: Classification and applications. *Journal of Network and Computer Applications*, *34*(5), 1671–1682.
2. Stajano, F., Hoult, N., Wassell, I., Bennett, P., Middleton, C., & Soga, K. (2010). Smart bridges, smart tunnels: Transforming wireless sensor networks from research prototypes into robust engineering infrastructure. *Ad Hoc Networks*, *8*(8), 872–888.
3. Fisher, W., Camp, T., & Krzhizhanovskaya, V. (2016). Crack detection in earth dam and levee passive seismic data using support vector machines. In *Proceedings of ICCS*, San Diego, CA, USA.
4. Komguem, R. D., Stanica, R., Tchuente, M., & Valois, F. (2014). WARIM: wireless sensor networks architecture for a reliable intersection monitoring. In *Proceedings of IEEE ITSC 2014*, Quingdao, China.
5. Sun, Z., Wang, P., Vuran, M., Al-Rodhaan, M., Al-Dhelaan, A., & Akyildiz, I. (2011). BorderSense: Border patrol through advances wireless sensor networks. *Ad Hoc Networks*, *9*(3), 468–477.
6. Perillo, M., Cheng, Z., & Heinzelman, W. (2004). On the problem of unbalanced load distribution in wireless sensor networks. In *IEEE global telecommunications conference workshops*, Dallas, USA (pp. 74–79).
7. Noori, M., & Ardakani, M. (2008). Characterizing the traffic distribution in linear wireless sensor networks. *IEEE Communications Letters*, *12*(8), 554–556.
8. Guo, Y., Kong, F., Zhu, D., Tosun, A. Ş., & Deng, Q. (2010). Sensor placement for lifetime maximization in monitoring oil pipelines. In *Proceedings of the 1st ACM/IEEE international conference on cyber-physical systems* (pp. 61–68).
9. Olariu, S., & Stojmenovic, I. (2006). Design guidelines for maximizing lifetime and avoiding energy holes in sensor networks with uniform distribution and uniform reporting. *INFOCOM*, *2006*, 1–12.
10. Ok, C., Thadakamalla, H., Raghavan, U., Kumara, S., Kim, S. G., Zhang, X., & Bukkapatnam, S. (2007). Optimal transmission power in self-sustainable sensor networks for pipeline monitoring. In *IEEE international conference on automation science and engineering (CASE)* (pp. 591–596).
11. Liu, X., & Mohapatra, P. (2007). On the deployment of wireless data back-haul networks. *IEEE Transactions on Wireless Communications*, *6*(4), 1426–1435.
12. Aberer, K., Hauswirth, M., & Salehi, A. (2006). A middleware for fast and flexible sensor network deployment. In *Proceedings of the 32nd international conference on Very large data bases*.
13. Raveendranathan, N., Galzarano, S., Loseu, V., Gravina, R., Giannantonio, R., Sgroi, M., et al. (2012). From modeling to implementation of virtual sensors in body sensor networks. *IEEE Sensors Journal*, *12*(3), 583–593.
14. Madria, S., Kumar, V., & Dalvi, R. (2014). Sensor cloud: A cloud of virtual sensors. In *IEEE Software* (Vol. 31, No. 2, pp. 70–77).
15. Rappaport, T. S. (1996). *Wireless communications: Principles and practise*. Upper Saddle River: Prentice Hall.
16. Yetgin, H., Cheung, K. T. K., El-Hajjar, M., & Hanzo, L. H. (2017). A survey of network lifetime maximization techniques in wireless sensor networks. *IEEE Communications Surveys & Tutorials*, *19*(2), 828–854.
17. Vicaire, P., He, T., Cao, Q., Yan, T., Zhou, G., Gu, L., et al. (2009). Achieving long-term surveillance in vigilnet. *ACM Transaction Sensor Networks*, *5*(1), 1–39.
18. Sevgi, C., & Koçyiğit, A. (2014). Optimal deployment in randomly deployed heterogeneous WSNs: A connected coverage approach. *Journal of Network and Computer Applications*, *46*, 182–197.
19. Bhuiyan, M. Z. A., Wang, G., Cao, J., & Wu, J. (2015). Deploying wireless sensor networks with fault-tolerance for structural health monitoring. *IEEE Transactions on Computers*, *64*(2), 382–395.
20. Parrado-Garcia, F. J., Vales-Alonso, J., & Alcaraz, J. J. (2017). Optimal planning of WSN deployments for in situ lunar surveys. *IEEE Transactions on Aerospace and Electronic Systems*, *53*(4), 1866–1879.
21. Boubrima, A., Bechkit, W., & Rivano, H. (2017). Optimal WSN deployment models for air pollution monitoring. *IEEE Transactions on Wireless Communications*, *16*(5), 2723–2735.
22. Kulkarni, N., Prasad, N. R., & Prasa, R. (2018). A novel sensor node deployment using low discrepancy sequences for WSN. *Wireless Personal Communications*, *100*(2), 241–254.
23. Potthuri, S., Shankar, T., & Rajesh, A. (2018). Lifetime improvement in wireless sensor networks using hybrid differential evolution and simulated annealing (DESA). *Ain Shams Engineering Journal*, *9*(4), 655–663.
24. Domga, K.R., Stanica, R., Tchuente, M., & Valois, F. (2017). Nodes ranking in wireless sensor network with linear topology. In *IEEE WD'2017*, Porto, Portugal.
25. Zahid, A. M., Kamalrulnizam, A. B., Muhammad, A., & Hafiz, M. (2018). An overview of routing techniques for road and pipeline monitoring in linear sensor networks. Wireless *Networks*, *24*(6), 2133–2143.
26. Plancoulaine, S., Bachir, A., & Barthel, D. (2006). *WSN node energy dissipation*. Technical report, France Telecom R&D, Internal Report.
27. Ye, W., Heidemann, J., & Estrin, D. (2004). Medium access control with coordinated adaptive sleeping for wireless sensor networks. *IEEE/ACM Transactions on Networking*, *12*(3), 493–506.

28. Zhao, Y., Wu, J., Li, F., & Lu, S. (2012). On maximizing the lifetime of wireless sensor networks using virtual backbone scheduling. *IEEE Transactions on Parallel and Distributed Systems*, *23*(8), 1528–1535.
29. Chen, Y., Zhao, Q., Krishnamurthy, V., & Djonin, D. (2007). Transmission scheduling for optimizing sensor network lifetime: A stochastic shortest path approach. *IEEE Transactions on Signal Processing*, *55*(5), 2294–2309.

Rodrigue K. Domga received his Bachelor and Master degree in Computer Sciences from the University of Yaounde I, Cameroon, in 2007 and 2011, respectively. He is currently a Ph.D. student both in the department of Computer Sciences at the University of Yaounde I, Cameroon, and in the CITI Laboratory at INSA Lyon, France. He is currently an assistant lecturer in the Department of Computer Sciences at the University of Yaounde I, Cameroon. His research interests include self-configuration and self-organizaion in wireless sensor networks.

Razvan Stanica received the M.Eng. and Ph.D. degrees in computer science from Institut National Polytechnique de Toulouse, France, in 2008 and 2011, respectively. He also received the MEng degree from the Polytechnic University of Bucharest, Romania. He is an associate professor with INSA Lyon, France, and a research scientist with the INRIA Agora Team, CITI Laboratory. His research interests include wireless networking in dynamic environments and network data analytics.

Maurice Tchuente is Board Chairman University of Yaounde II, Professor of computer science, University of Yaounde I, Member of the Pan African University Council and the Scientific Council of IRD, France, and Chairman of the Scientific Council of the international laboratory LIRIMA. He is docteur d'état es sciences mathématiques, Université Joseph Fourier, Grenoble, HDR in Computer Science, INPG Grenoble, and has served as research fellow CNRS, France and Directeur de Recherche IRD, France. Among his former administrative positions, he was Minister of higher education (2002–2004), Pro-Chancellor University of Buea (2009–2017), coordinator (interim rector) Pan African University (2012), chairman of the National Agency for Information and Communication Technologies (2006–2009), Rector of the universities of Dschang (1996–1998), Ngaoundere (1998–2000) and Douala (2000–2002). His research topics are parallel algorithms, dynamics of automata networks, and data mining. He is the co-author of more than one hundred research papers and author of the book "Parallel Computation on Regular Arrays", Manchester University Press. He has supervised more than 20 Ph.D. thesis He is a founding member of CARI, an African Network for Research in mathematics and Computer Science, and his research team is a component of UMMISCO (Unité de Modélisation Mathématique et Informatique des Systèmes Complexes) He received the Boutros Boutros Ghali Award of the United Nations University in 1995 and the CNR Rao Prize for Scientific Research, TWAS, 2007.

Fabrice Valois is full professor in INSA Lyon (France), since 2008. In january 2000, he received a Ph.D. in computer science about Performance Evaluation of Hierarchical Cellular Networks from University of Versailles (France). In november 2007, he received the Habilitation à Diriger des Recherches from University of Lyon I and INSA Lyon about Self-organisation of wireless multi-hop networks. Fabrice Valois is also a member of the CITI laboratory and he is involved in the Agora Inria research team. His research interests are in the area of dynamic, dense and autonomous wireless networks (e.g. networks for IoT, wireless sensor networks, wireless mobile ad hoc networks). More particularly, his research works are focused on networking protocols design and performance evaluation. He is co-author of 3 chapter books, 95+ international publications and he holds 4 international patents.