

UNIVERSITÉ DE MONTRÉAL

MÉCANISMES OPTIMISÉS DE PLANIFICATION DES ÉTATS DES  
CAPTEURS POUR LA MAXIMISATION DE LA DURÉE DE VIE DANS LES  
RÉSEAUX DE CAPTEURS SANS FIL

ALI CHAMAM

DÉPARTEMENT DE GÉNIE INFORMATIQUE ET GÉNIE LOGICIEL  
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION  
DU DIPLÔME DE PHILOSOPHIÆ DOCTOR  
(GÉNIE ÉLECTRIQUE)

AOÛT 2009

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée:

MÉCANISMES OPTIMISÉS DE PLANIFICATION DES ÉTATS DES  
CAPTEURS POUR LA MAXIMISATION DE LA DURÉE DE VIE DANS LES  
RÉSEAUX DE CAPTEURS SANS FIL

présentée par: CHAMAM Ali

en vue de l'obtention du diplôme de: Philosophiæ Doctor

a été dûment acceptée par le jury d'examen constitué de:

Mme. BOUCHENEB Hanifa, Doctorat., présidente

M. PIERRE Samuel, Ph.D., membre et directeur de recherche

M. QUINTERO Alejandro, Doct., membre

M. AJIB Wessam, Ph.D, membre

*À mes chers parents,  
pour leur soutien indéfectible*

*À ma chère épouse Lamia,  
pour sa patience et ses sacrifices*

*À mon frère Sami et à ma sœur Emna*

*À toute ma famille et ma belle-famille*

## REMERCIEMENTS

J'aimerais profiter de cette tribune pour exprimer ma gratitude envers toutes les personnes qui, de près ou de loin, m'ont apporté leur aide et leur soutien durant ces années de thèse.

Tout d'abord, j'adresse mes plus vifs remerciements à mon directeur de recherche, le professeur Samuel Pierre, pour son encadrement, son soutien continu et ses conseils judicieux. La qualité de son encadrement m'a offert une large marge de liberté intellectuelle et scientifique qui m'a permis de mener à bien ce laborieux travail de recherche.

J'aimerais ensuite exprimer toute ma reconnaissance et ma gratitude envers mes chers parents Naceur et Saloua, pour leur soutien continu, leur patience et leurs encouragements. Ils ont toujours été ma première source d'inspiration et sans leur épaulement indéfectible, cette thèse n'aurait sûrement jamais abouti.

Ma gratitude la plus sincère s'adresse également à ma chère épouse Lamia, pour son dévouement continu, sa disponibilité et ses sacrifices qui m'ont aidé à passer au travers des moments d'hésitation et de confusion, inhérents à toute thèse de doctorat. Que ce travail soit pour elle le témoignage de mon infinie reconnaissance pour ces années de compréhension, de privations et d'efforts communs.

Je tiens aussi à remercier tous mes collègues du laboratoire LARIM pour l'ambiance de travail conviviale qu'ils ont su faire partager.

Je remercie plus particulièrement mon collègue et ami Hichem Ayed Harhira pour son aide précieuse et pour la qualité des discussions scientifiques que nous avons eues lors de cette thèse et qui m'ont permis de mieux aiguiller certains aspects de mes recherches.

Je remercie enfin toutes les personnes qui m'ont aidé au sein du département de Génie Informatique et Génie Logiciel, plus particulièrement Mme Louise Longtin et Mme Madeleine Guillemette pour leur gentillesse et leur disponibilité durant toutes ces années de thèse.

## RÉSUMÉ

Depuis leur création, les réseaux de communication sans fil ont connu un engouement fulgurant qui ne cesse de croître au sein des communautés scientifiques et industrielles. Ainsi, le paradigme sans fil a vu naître, au cours de son évolution, diverses architectures dérivées, telles que les réseaux cellulaires, les réseaux locaux sans fil, les réseaux WiMax, etc. Durant la dernière décennie, un nouveau type de réseau sans fil a suscité un grand intérêt auprès de la communauté scientifique, il s'agit des réseaux ad hoc et des réseaux de capteurs sans fil (RCSF). Ce nouveau type de réseau se distingue des réseaux sans fil classiques par l'absence d'infrastructure (ou structure) préétablie et par la versatilité de ses nœuds (i.e., ces derniers peuvent s'ajouter au réseau ou en disparaître d'une manière assez aléatoire).

Un RCSF est composé d'un ensemble d'unités de traitements embarquées, appelées capteurs, communiquant via des liens sans fil et dont la fonction principale est la collecte de paramètres relatifs à l'environnement qui les entoure, telles que la température, la pression ou la présence d'objets. Les RCSF trouvent leur application dans diverses activités de la société, tels les processus industriels, les applications militaires de surveillance, l'observation et le suivi d'habitat, etc. À cause de la nature intrinsèque de leur fonctionnalité, les capteurs ont une contrainte principale : leur source d'énergie est limitée et presque jamais renouvelable. Ceci place l'optimisation de l'énergie consommée par les capteurs et la maximisation de la durée de vie des RCSF au centre des défis posés par ces réseaux.

Dans cette thèse, nous abordons le problème de la maximisation de la durée de vie des RCSF par une planification optimisée des états des capteurs, dont le but est de minimiser l'énergie totale consommée par les capteurs tout en distribuant la consommation énergétique sur l'ensemble des capteurs. Ceci implique la désactivation de certains capteurs non nécessaires à la couverture de la zone surveillée ainsi

que la formation optimale de grappes (*clusters*), tout en respectant les contraintes imposées par l'application, telle la couverture totale de la zone surveillée par le RCSF et l'existence d'une topologie de routage composée de têtes de grappes. Notre thèse se présente sous la forme de trois articles scientifiques, chacun traitant d'une problématique bien spécifique.

Dans notre premier article intitulé "On the planning of wireless sensor networks : energy-efficient clustering under the joint routing and coverage constraint" et qui a été publié dans la revue *IEEE Transactions on Mobile Computing*, nous abordons la problématique de la maximisation de la durée de vie d'un RCSF sous la triple contrainte de couverture, de structuration en grappes et de routage entre les têtes de grappes. Nous considérons un RCSF formé de nœuds déployés aléatoirement pour surveiller une zone donnée, et où chaque capteur peut être dans l'un des états suivants : éteint, allumé, ou tête de grappe. Chacun de ces états a une consommation énergétique spécifique. Notre objectif est de trouver l'allocation optimale, en termes de durée de vie du réseau, d'états aux capteurs de telle façon à garantir la couverture totale de la zone surveillée par les capteurs, la structuration en grappes du réseau et l'existence d'un arbre de couverture reliant toutes les têtes de grappes. Nous modélisons notre problème sous forme d'un problème d'optimisation en nombres entiers que nous démontrons NP-complet. Nous proposons par la suite une heuristique taboue pour trouver une solution proche de l'optimale. Nos résultats démontrent que l'heuristique proposée génère des solutions proches de la solution optimale trouvée moyennant l'outil de résolution mathématique CPLEX et qu'elle donne de meilleures performances que d'autres algorithmes distribués de formation de grappes proposés dans la littérature scientifique récente.

Dans notre deuxième article intitulé "Energy-efficient clustering with transmit power assignment under coverage constraint in wireless sensor networks", et qui a été soumis

à la revue *Wireless Communications & Mobile Computing (Wiley)*, nous traitons le problème de la maximisation de la durée de vie d'un RCSF dont les nœuds sont munis de plusieurs puissances de transmission. Un capteur peut donc atteindre un plus grand nombre de têtes de grappes potentielles en utilisant une plus grande puissance de transmission, mais à un coût énergétique plus élevé. Dans cette problématique, la puissance de transmission utilisée par chaque capteur influence la configuration globale du réseau et nous nous intéressons alors à trouver l'allocation optimale, en termes de durée de vie du réseau, des états et des puissances de transmission aux capteurs, tout en assurant la couverture totale de la zone surveillée par les capteurs "allumés" et la structuration en grappes du réseau, i.e., chaque capteur doit être soit une tête de grappe, soit connecté à une tête de grappe voisine moyennant la puissance de transmission qui lui a été allouée. Nous modélisons notre problème sous forme d'un problème d'optimisation à variables entières, que nous démontrons NP-complet. Nous proposons ensuite un algorithme à base d'heuristique taboue pour trouver une solution proche de l'optimalité. Nos résultats de simulation démontrent que notre heuristique génère de bons résultats, comparée à la solution optimale générée par CPLEX. Comparé à HEED, un protocole distribué de formation de grappes dans les RCSF munis de puissance de transmission variable, notre algorithme génère de meilleurs résultats en termes de durée de vie du réseau et de répartition de l'énergie dissipée sur les capteurs.

Dans notre troisième article intitulé "A distributed energy-efficient clustering protocol for wireless sensor networks", et publié dans la revue *Computers and Electrical Engineering (Elsevier)*, nous proposons un nouvel algorithme distribué de formation de grappes dans les RCSF. Notre algorithme diffère de la majorité des algorithmes proposés dans la littérature par la pertinence des informations récoltées par chaque capteur de la part de ses voisins. Grâce à un mécanisme d'échange de fonctions score, chaque capteur collecte des informations sur l'aptitude de ses voisins directs (*1-hop*



*neighbors*) et ceux séparés par deux sauts (*2-hops neighbors*), à devenir têtes de grappe. Ces informations seront utilisées par chaque capteur pour affiner sa prise de décision relativement à son état (devenir tête de grappe ou non). Les résultats de simulation montrent que notre algorithme donne de meilleurs résultats, en termes de durée de vie du réseau, que d'autres algorithmes répartis proposés dans la littérature scientifique récente.

Les principales contributions de cette thèse sont :

- La modélisation du problème de planification des états des capteurs dans un RCSF, dans le but de maximiser la durée de vie du réseau sous la triple contrainte de couverture, formation de grappes et routage. Les solutions du modèle proposé constituent une borne supérieure, en termes de durée de vie du réseau, pour tout protocole distribué abordant le même problème. Notre modèle peut donc servir comme référence de comparaison pour évaluer la performance d'autres protocoles distribués traitant le même problème. À notre connaissance, aucun travail préalable dans la littérature n'a proposé une telle borne pour le problème global de planification optimale des états des capteurs et sous les mêmes contraintes générales de couverture, de routage et de formation de grappes.
- La modélisation du problème d'allocation optimale des états et des puissances de transmission aux capteurs dans un RCSF muni de puissance de transmission variable. Les solutions à ce problème offrent une borne supérieure, en termes de durée de vie du réseau, aux algorithmes distribués traitant le même problème.
- La proposition d'un nouvel algorithme distribué de formation de grappes dans les réseaux de capteurs, qui génère de meilleures performances, en termes de durée de vie du réseau, que d'autres algorithmes distribués populaires dans la littérature scientifique relative au sujet. Notre algorithme constitue donc une amélioration de l'existant en ce qui a trait aux algorithmes distribués de formation de grappes dans les RCSF.

## ABSTRACT

Since their creation, wireless communication networks have witnessed a huge success that continues to grow in scientific and industrial communities. During its evolution the wireless paradigm has given birth to various derivative architectures, such as cellular, WiMax and wireless local area networks. During the last decade, a new type of wireless networks has stirred up great interest within the scientific community ; it consists in mobile ad hoc networks (MANETS) and wireless sensor networks (WSN). This new type of networks differs from conventional wireless networks by the absence of predetermined infrastructure (or structure) and by the versatility of its nodes (i.e., any node can join the network or leave it in a pretty random manner). A WSN consists of a set of embedded processing units, called sensors, communicating via wireless links, whose main function is the collection of parameters related to the surrounding environment, such as temperature, pressure or the presence/motion of objects. WSN are expected to have many applications in various fields, such as industrial processes, military surveillance, observation and monitoring of habitat, etc. Because of the intrinsic nature of their intended applications, sensors have a major constraint : their energy source is usually limited and hardly renewable. This turns energy optimization and network lifetime maximization in WSN into real challenges. In this thesis, we address the problem of maximizing the network lifetime of WSN through optimal sensor state planning, with a view to minimizing the total dissipated energy while ensuring a fair balance of energy consumption over all sensors. This involves such techniques as cluster formation and switch-off of some sensors, while conforming to the application-related constraints, such as total coverage of the area monitored by the sensors and the presence of a routing topology composed of cluster heads.

Our thesis runs through three scientific papers, each addressing a specific problem.

In our first paper entitled "On the planning of wireless sensor networks : energy-efficient clustering under the joint routing and coverage constraint", and published in *IEEE Transactions on Mobile Computing*, we address the optimal planning of sensors' states in cluster-based sensor networks. In that paper, we consider that any sensor can be in one of three states : active (turned on), turned off or cluster head, and a different power consumption is associated with each of these states. Our objective is to find an energy-optimal topology that maximizes network lifetime while ensuring simultaneously full area coverage and a fully clustered configuration in which the set of cluster heads is constrained to form a spanning tree used as a routing topology. We first formulated this problem as an Integer Linear Programming model that we proved NP-Complete. Then, we implemented a Tabu search heuristic to tackle the exponentially-increasing computation time of the exact resolution. Experimental results show that the proposed heuristic provides near-optimal network lifetime values within low computation times, which in practice is suitable for large-sized sensor networks. Our heuristic also offers better performance than other recent distributed clustering algorithms proposed in the scientific literature. In the second paper entitled "Energy-efficient clustering with transmit power assignment under coverage constraint in wireless sensor networks", submitted to the *Wireless Communications and Mobile Computing journal (Wiley)*, we address the problem of maximizing network lifetime in a WSN, under coverage and clustering constraints. We consider sensors equipped with multiple transmit powers and thus capable of reaching a wider set of cluster heads at higher energy costs, and we propose to find the optimal assignment of states and transmit powers to sensors. We first model this problem as an Integer Linear Programming (ILP) optimisation problem that we prove NP-complete. Then, we propose a Tabu search heuristic to find acceptable-quality solutions within realistic processing times. Our simulation-based results show that our proposed heuristic provides better solutions compared to the optimal solutions generated by CPLEX, i.e., it outperforms, in terms of network lifetime, HEED - the

well-recognized distributed clustering algorithm proposed in the literature. In our third paper entitled "A distributed energy-efficient clustering protocol for wireless sensor networks ", published in the *Computers and Electrical Engineering journal (Elsevier)*, we propose a new distributed clustering algorithm for WSN, called EECF. The novelty of EECF resides in the relevance of the information gathered by each sensor from its neighbors. Each sensor makes its decision to promote itself cluster head based on score functions exchanged with its one-hop neighbors, which hold information on the ability of its neighbors to become cluster heads. Simulation results show that our algorithm generates better results, in terms of network lifetime, than other distributed algorithms proposed in the recent scientific literature.

The main contributions of this thesis are :

- The modeling of the energy-optimal sensor state planning problem in WSN, under the joint coverage, clustering and routing constraint. The exact solutions of this optimization problem represent an upper bound on the network lifetime and can thus serve as a comparison basis for evaluating the performance of other distributed clustering algorithms addressing the same problem, under the same constraints. To the best of our knowledge, no prior work in the literature has proposed such a bound for the global optimal sensor state planning under the same general integrated coverage, routing and clustering constraints.
- The modeling of the problem of optimal allocation of states and transmit powers to the sensors in a WSN, under joint coverage and clustering constraints. Solutions of this problem provide an upper bound on the network lifetime and can thus serve as a comparison reference to assess the performance of other distributed algorithms addressing the same problem.
- The proposal of a novel distributed clustering algorithm for WSN, which outperforms, in terms of lifetime of the network, other popular distributed algorithms proposed in the recent scientific literature. Our algorithm is therefore an improvement of the state-of-the-art on distributed cluster formation algorithms for WSN.

## TABLE DES MATIÈRES

DÉDICACE . . . . .	iii
REMERCIEMENTS . . . . .	iv
RÉSUMÉ . . . . .	vi
ABSTRACT . . . . .	x
TABLE DES MATIÈRES . . . . .	xiii
LISTE DES FIGURES . . . . .	xvii
LISTE DES TABLEAUX . . . . .	xx
LISTE DES ALGORITHMES . . . . .	xxi
LISTE DES SIGLES ET ABRÉVIATIONS . . . . .	xxii
CHAPITRE 1 INTRODUCTION . . . . .	1
1.1 Définitions et concepts de base . . . . .	3
1.1.1 Architecture et domaines d'application des Réseaux de cap- teurs sans fil . . . . .	3
1.1.2 Anatomie des capteurs . . . . .	5
1.1.3 Contrainte d'énergie dans les réseaux de capteurs sans fil . . . .	6
1.2 Éléments de problématique . . . . .	8
1.3 Objectifs de recherche . . . . .	14
1.4 Principales contributions . . . . .	15
1.5 Plan de la thèse . . . . .	17
CHAPITRE 2 REVUE DE LITTÉRATURE . . . . .	19

2.1	Architecture des capteurs et modèle de communication en couches . . .	20
2.2	Mécanismes de conservation de l'énergie dans les RCSF au niveau des différentes couches . . . . .	26
2.2.1	Conservation de l'énergie sous contrainte de couverture . . . . .	34
2.2.2	Conservation de l'énergie par la formation de grappes ( <i>clustering</i> )	35
2.2.3	Conservation de l'énergie par l'ajustement optimisé des puis- sances de transmission . . . . .	39
2.2.4	Conservation de l'énergie par la planification optimisée des états des capteurs . . . . .	41
2.3	Conclusion . . . . .	44
CHAPITRE 3 ON THE PLANNING OF WIRELESS SENSOR NETWORKS : ENERGY-EFFICIENT CLUSTERING UNDER THE JOINT ROUTING AND COVERAGE CONSTRAINT . . . . .		
		45
3.1	Introduction . . . . .	46
3.2	Related work . . . . .	50
3.3	Problem statement and assumptions . . . . .	53
3.4	Problem modeling . . . . .	57
3.5	Proposed heuristic . . . . .	65
3.6	Simulation results . . . . .	67
3.6.1	Comparative performance evaluation : TABU-RCC with res- pect to its lower bound (CPLEX) . . . . .	68
3.6.2	Performance evaluation of TABU-RCC : Impact of the sensing range . . . . .	70
3.6.3	Performance evaluation of TABU-RCC : Impact of the maxi- mum cluster size . . . . .	71
3.6.4	Comparative performance evaluation : TABU-RCC vs EESH .	72
3.7	Conclusion . . . . .	73

CHAPITRE 4	ENERGY-EFFICIENT CLUSTERING WITH TRANSMIT POWER ASSIGNMENT UNDER COVERAGE CONSTRAINT IN WIRELESS SENSOR NETWORKS . . . . .	74
4.1	Introduction . . . . .	75
4.2	Related work . . . . .	80
4.3	Problem Statement and assumptions . . . . .	83
4.4	Problem modeling . . . . .	88
4.5	Proposed heuristic . . . . .	94
4.6	Implementation and simulation results . . . . .	95
4.6.1	Quality of the solution : T-CPACC vs its lower bound . . . . .	96
4.6.2	Measure of the network lifetime without coverage constraint : T-CPACC vs HEED . . . . .	98
4.6.3	Measure of the network lifetime with coverage constraint : T- CPACC vs Modified HEED . . . . .	100
4.7	Conclusion and future work . . . . .	101
CHAPITRE 5	A DISTRIBUTED ENERGY-EFFICIENT CLUSTERING PRO- TOCOL FOR WIRELESS SENSOR NETWORKS . . . . .	102
5.1	Introduction . . . . .	103
5.2	Related work . . . . .	105
5.3	Problem description and assumptions . . . . .	108
5.4	Description of the EECF protocol . . . . .	110
5.5	Protocol analysis . . . . .	114
5.6	Protocol implementation and Performance evaluation . . . . .	118
5.6.1	Experiment setup . . . . .	118
5.6.2	Effect of the score coefficient $\alpha$ on performance . . . . .	119
5.6.3	Comparison of EECF and EESH in terms of network lifetime . . . . .	122

5.6.4	Comparison of EECF and EESH in terms of the CH ratio (Number of CHs/Total number of sensors) . . . . .	124
5.7	Conclusion . . . . .	124
CHAPITRE 6	DISCUSSION GÉNÉRALE . . . . .	126
6.1	Synthèse des travaux . . . . .	126
6.2	Méthodologie . . . . .	128
6.3	Analyse des résultats . . . . .	129
CHAPITRE 7	CONCLUSION . . . . .	130
7.1	Sommaire des contributions . . . . .	130
7.2	Limitation des travaux . . . . .	131
7.3	Indication des travaux futurs . . . . .	133
CONTRIBUTIONS À LA RECHERCHE	. . . . .	135
BIBLIOGRAPHIE	. . . . .	137



## LISTE DES FIGURES

FIGURE 1.1	Architecture d'une réseau de capteurs sans fil . . . . .	3
FIGURE 1.2	Transmission des données en multi-sauts . . . . .	5
FIGURE 1.3	Architecture typique d'un capteur sans fil . . . . .	6
FIGURE 1.4	Sous-ensemble de couverture . . . . .	10
FIGURE 1.5	Les différents états qu'un capteur peut prendre . . . . .	11
FIGURE 1.6	Compromis entre les puissances de transmission et le nombre de TG . . . . .	12
FIGURE 2.1	Modèle de communication en couches et plan de gestion de l'énergie . . . . .	20
FIGURE 2.2	Structure en grappes d'un RCSF . . . . .	35
FIGURE 3.1	Multi-hop routing of collected data . . . . .	47
FIGURE 3.2	A full-covering sensor set . . . . .	48
FIGURE 3.3	Sensor states . . . . .	54
FIGURE 3.4	Overlay network composed of the selected cluster heads . . . . .	55
FIGURE 3.5	Shape of the score function . . . . .	59
FIGURE 3.6	Value of the objective function generated by TABU-RCC vs CPLEX . . . . .	69

FIGURE 3.7	Average consumed energy using the solution generated by TABU-RCC vs CPLEX . . . . .	69
FIGURE 3.8	Variation of the network lifetime with the sensing range, using TABU-RCC . . . . .	70
FIGURE 3.9	Variation of the network lifetime with the maximum cluster size, using TABU-RCC . . . . .	71
FIGURE 3.10	Comparison of TABU-RCC to EESH and the modified EESH in terms of network lifetime . . . . .	72
FIGURE 4.1	Multihop routing of collected data . . . . .	76
FIGURE 4.2	A full-covering sensor set with variable sensing ranges . . . . .	77
FIGURE 4.3	Two limit scenarios : low transmit power and many CHs re- quired vs high transmit power and less CHs required . . . . .	86
FIGURE 4.4	Necessary condition to have a connected overlay CH graph : $R_t^{CH} \geq 2.(R_s + R_t(P_K))$ . . . . .	93
FIGURE 4.5	Objective function generated by T-CPACC vs Cplex . . . . .	96
FIGURE 4.6	Network lifetime generated by T-CPACC vs HEED . . . . .	97
FIGURE 4.7	Standard deviation of the dissipated energy, generated by T- CPACC vs HEED . . . . .	98
FIGURE 4.8	Network lifetime generated by T-CPACC vs HEED and the Modified HEED, for $R_s=1$ . . . . .	99
FIGURE 4.9	Network lifetime generated by T-CPACC vs HEED and the Modified HEED, for $R_s=2$ . . . . .	100

FIGURE 5.1	Multi-hop routing of collected data . . . . .	103
FIGURE 5.2	Description the EECF protocol, executed by every sensor . . .	111
FIGURE 5.3	Necessary condition to have a connected overlay CH graph : $R_t^{CH} \geq 3.R_t$ . . . . .	117
FIGURE 5.4	Network lifetime generated by EECF function of $\alpha$ , for dif- ferent values of $R_t$ . . . . .	120
FIGURE 5.5	Ratio "Number of CHs/Total number of sensors" generated by EECF function of $\alpha$ , for different values of $R_t$ . . . . .	121
FIGURE 5.6	Network lifetime generated by EECF vs. EESH, for different values of $R_t$ . . . . .	122
FIGURE 5.7	Ratio "Number of CHs/Total number of sensors" generated by EECF vs. EESH, for different values of $R_t$ . . . . .	123

## LISTE DES TABLEAUX

TABLEAU 2.1	Consommation énergétique des différents états d'opération des capteurs MEDUSA II . . . . .	42
TABLEAU 2.2	Consommation énergétique, aux différents états d'opération, des capteurs WINS de Rockwell . . . . .	42
TABLEAU 3.1	Computation time of the lower bound (CPLEX) and TABU-RCC . . . . .	68

## LISTE DES ALGORITHMES

ALGORITHME 3.1	TABU-RCC : Tabu search algorithm of CH election under routing and coverage constraints . . . . .	66
ALGORITHME 4.1	T-CPACC : Tabu search algorithm for clustering and power assignment in WSN under coverage constraint . . . . .	94

## LISTE DES SIGLES ET ABRÉVIATIONS

BPSK	Binary Phase Shift Keying
CH	Cluster head
CPU	Central Processing Unit
CT	Centre de Traitement
GPS	Global Positioning System
IP	Internet Protocol
MANET	Mobile Ad hoc NETwork
O-QPSK	Offset Quadrature Phase Shift Keying
OSI	Open Systems Interconnection
PN	Processing Node
QoS	Qualité de Service
RCSF	Réseau de Capteurs Sans Fil
RF	Radio-Fréquences
SQL	Structured Query Langage
TCP	Transmission Control Protocol
TG	Tête de grappe
WiMax	Worldwide Interoperability for Microwave Access
WSN	Wireless Sensor Network

## CHAPITRE 1

### INTRODUCTION

Les récentes avancées technologiques dans le domaine des communications sans fil ont permis le développement à faible coût de minuscules systèmes micro-électromécaniques (MEMS), appelés capteurs, capables de détecter, mesurer et rapporter des données physiques liées à leur environnement. Ces capteurs sont caractérisés par de faibles ressources (énergie, capacité de calcul, mémoire, etc.) et de faible consommation énergétique, et ils ont trois fonctions principales : 1) la capture de données reliées à leur environnement physique (température, pression, vibration, lumière, mouvement, etc.); 2) le traitement des données collectées; 3) la transmission de ces données à un centre de traitement. Selon leur structure électronique, ils peuvent détecter des signaux mécaniques, acoustiques, électriques, photoniques, électromagnétiques, vibratoires, etc. Les réseaux de capteurs sans fil (RCSF) sont des réseaux sans infrastructure, généralement destinés à être déployés en grand nombre, pour couvrir des surfaces plus ou moins larges. Leurs applications potentielles diversifiées, dans les domaines militaire, industriel et domestique, sont la source de l'intérêt qu'ils suscitent, aussi bien dans la communauté scientifique qu'industrielle.

Dans certaines applications domestiques ou industrielles, les RCSF peuvent constituer un pont entre les services offerts par les réseaux classiques et les données reliées à l'environnement physique. Ceci procure aux réseaux de prochaines générations, dont feront partie les réseaux de capteurs, une intelligence qui leur permet d'offrir des services plus adaptés à l'utilisateur et l'environnement qui l'entoure. Néanmoins, avant d'atteindre la phase d'implantation industrielle à grande échelle des réseaux de capteurs, beaucoup de problèmes restent encore à traiter et beaucoup de solutions embryonnaires proposées dans la recherche demeurent encore à améliorer, telles

que les solutions proposées aux problèmes d'optimisation de l'énergie dépensée, de routage des données, d'auto-organisation en cas de perte de nœuds, d'agrégation et de fusion des données, d'interrogation du réseau et de collecte des données, etc. Parmi ces problèmes, la minimisation de l'énergie dissipée et la maximisation de la durée de vie du réseau sont sans doute ceux qui ont suscité le plus d'intérêt dans la recherche. En effet, de par leurs applications potentielles et leurs propriétés physiques, les capteurs sont généralement munis d'une source d'énergie très limitée et difficilement renouvelable, dont la conservation demeure donc cruciale. Cette problématique, très spécifique aux réseaux de capteurs, est au centre de la majorité des travaux rencontrés dans la littérature scientifique. Aussi, quand elle n'est pas l'objectif principal dans un travail de recherche, elle est toujours prise en compte dans tout processus de modélisation et de conception d'outils, d'algorithmes et de solutions relatifs aux réseaux de capteurs. Dans cet ordre d'idées, nous allons nous intéresser dans cette thèse aux mécanismes d'optimisation de la durée de vie des réseaux de capteurs. Nous proposerons des solutions basées sur une gestion optimisée des rôles des capteurs durant leur cycle de vie, afin de maximiser la durée de vie de ce dernier. Dans ce chapitre, nous commencerons par donner une brève description des réseaux de capteurs, leurs applications et leurs principaux défis. Ensuite, nous présenterons quelques éléments de problématique, liés principalement à la contrainte énergétique des capteurs. Puis, nous détaillerons nos objectifs de recherche. Par la suite, nous exposerons nos principales contributions à la recherche et enfin, nous présenterons le plan de cette thèse.



## 1.1 Définitions et concepts de base

### 1.1.1 Architecture et domaines d'application des Réseaux de capteurs sans fil

Comme illustré à la Figure 1.1, un réseau de capteurs sans fil (RCSF) est composé d'un ensemble de nœuds capteurs, munis de transmetteurs sans fil. Les capteurs envoient les données collectées (traitées ou pas) à un noeud passerelle ("*sink*" ou "*gateway*") directement ou via un chemin multi-sauts formé d'autres capteurs. La passerelle transmet ensuite ces données à un centre de traitement (CT), qui pourrait être un simple ordinateur central, pour les analyser et, éventuellement, prendre des décisions. Le CT effectue certaines opérations sur les données, comme le filtrage et l'agrégation, pour en extraire l'information utile. Les capteurs sont généralement équipés de faibles ressources (CPU, mémoire, etc.) et de source d'énergie limitée. Le CT et les passerelles sont généralement munis de puissance de calcul et de source

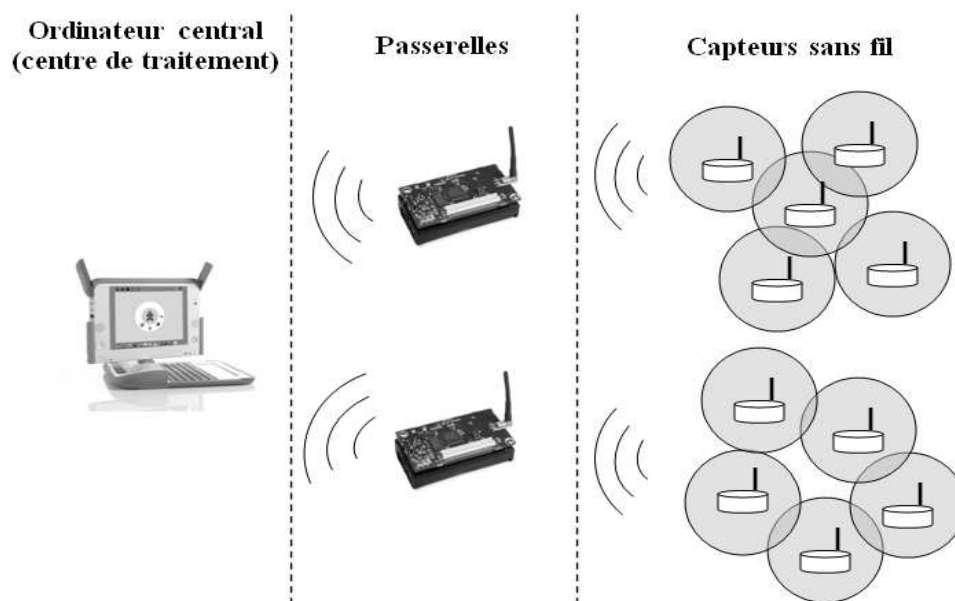


FIGURE 1.1 Architecture d'une réseau de capteurs sans fil

d'énergie supérieures à celles des capteurs. Le CT pourrait être connecté à un réseau filaire, comme Internet, et donc servir d'interface entre le réseau et l'utilisateur. À cause de leur limitation en énergie, et parce que la transmission des données est une opération qui exige une consommation énergétique relativement importante (Pottie et Kaiser, 2000), les capteurs disposent d'une portée d'émission généralement faible, limitée à quelques mètres. Les données sont donc généralement transmises au CT en multi-sauts (*multi-hop*), comme illustré à la Figure 3.1, et les capteurs doivent, dans certains cas, router les données provenant de leurs voisins, ce qui leur fait consommer encore plus d'énergie. Il est donc important de disposer de protocoles de routage efficaces en termes de conservation d'énergie.

Les RCSF sont généralement déployés en grand nombre dans une zone géographique où ils vont capter, mesurer et rapporter certains phénomènes physiques. Ils peuvent donc servir à surveiller leur environnement physique (Cayirci, 2004). Dans le domaine militaire, ils peuvent être utilisés pour surveiller le déplacement de troupes ennemies dans un espace donné ou récolter des données dans une zone dangereuse où l'envoi d'êtres humains serait à risque. À titre d'exemple, les RCSF ont été utilisés, sous une forme assez primitive, lors de la guerre froide par les Américains qui ont placé dans l'océan un système appelé SOSUS (Chong et Kumar, 2003) composé de capteurs acoustiques pour surveiller les sous-marins silencieux russes. Dans le domaine civil, les RCSF peuvent être utilisés dans la surveillance des infrastructures, des installations et des zones critiques. Ils peuvent également être utilisés pour surveiller des habitations et contribuer au confort domestique, en transformant les logements en environnements intelligents dont les paramètres (température, pression, humidité, luminosité, etc.) s'adaptent automatiquement au comportement des individus (Mainwaring *et al.*, 2002). Dans le domaine de l'environnement, ils peuvent localiser et suivre le mouvement de certains animaux, détecter des feux de forêt ou encore surveiller l'évolution de la densité moyenne de  $CO_2$  dans l'air (Akyildiz *et al.*, 2002). CORIE (CMOP, 2009) est un exemple réel d'application où des capteurs récoltent

des informations sur l'eau de la rivière Columbia (température, salinité, flux et niveau de l'eau, etc.), lesquelles sont utilisées pour gérer les opérations de secours et les interventions sur l'écosystème. Dans le domaine industriel, les capteurs peuvent être placés sur des boîtes de marchandise dans une usine ou un entrepôt pour suivre en temps réel le mouvement des marchandises, ou encore détecter le dysfonctionnement d'une machine en analysant ses émissions acoustiques et vibratoires. Actuellement, il existe quelques applications commerciales assez primitives des RCSF, par exemple pour la lecture distante des appareils commerciaux de mesure de la consommation énergétique domestique (les compteurs d'électricité, de gaz, etc.).

### 1.1.2 Anatomie des capteurs

Comme illustré à la Figure 1.3, un capteur sans fil est typiquement composé de quatre unités de base (Cayirci, 2004) : 1) une unité de capture (e.g., thermomètre dans le cas de données thermiques, caméra dans le cas de signaux vidéo, etc.) qui détecte le phénomène physique et transforme le signal détecté en signal numérique à l'aide d'un convertisseur de signal ; 2) une unité de traitement qui récupère le signal capté et effectue des opérations (généralement simples, à faible consommation énergétique) sur les données collectées ; 3) une unité de transmission radio (à faible portée) et 4) une batterie. Les capteurs peuvent être de différents types : sismique, magnétique, électromagnétique, infrarouge, vidéo, thermique, acoustique, radar, etc.

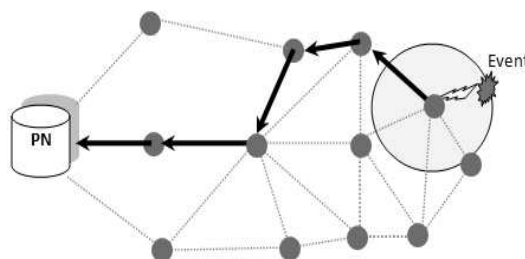


FIGURE 1.2 Transmission des données en multi-sauts

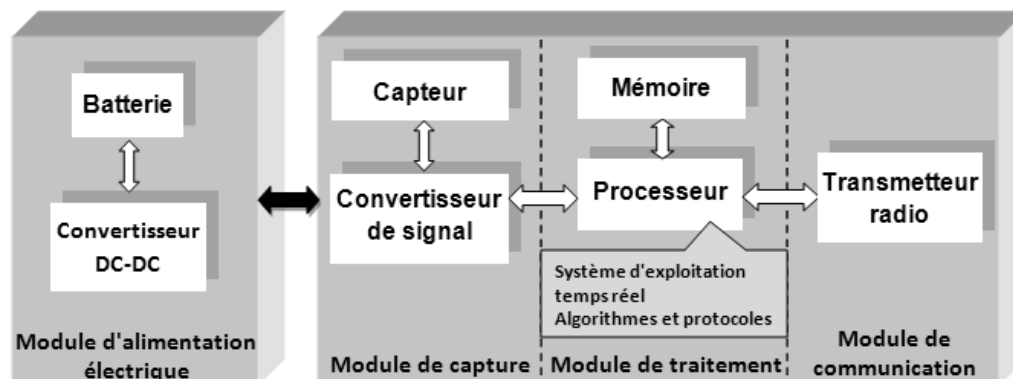


FIGURE 1.3 Architecture typique d'un capteur sans fil

(Akyildiz *et al.*, 2002), offrant ainsi un large éventail d'applications potentielles. D'autres unités, comme un localisateur, peuvent également faire partie des capteurs selon l'application envisagée. Le mode de communication inter-capteurs pourrait être optique (e.g., infrarouge), mais la communication radio, généralement dans la bande UHF, reste la plus pratique car elle supporte les obstacles physiques. Des efforts de standardisation de la couche physique et de la couche d'accès au canal des RCSF ont donné naissance à la norme IEEE 802.15.4 (Kohvakka *et al.*, 2006) qui établit des bandes fréquences et les types de modulation utilisables par les RCSF : 2.4 GHz comme fréquence globale, avec une modulation O-QBPSK, 915 MHz avec une modulation BPSK pour l'Amérique du nord et 868 MHz avec une modulation BPSK pour l'Europe. Elle définit aussi l'accès au canal pour les topologies en étoile, en maille et hiérarchiques.

### 1.1.3 Contrainte d'énergie dans les réseaux de capteurs sans fil

Les RCSF ont une contrainte importante d'énergie : la batterie des capteurs est généralement limitée pour préserver leur faible coût et leur taille réduite. Elle est également difficilement renouvelable de par la nature de certaines applications des RCSF qui nécessitent un déploiement aléatoire des capteurs, à grande échelle et dans des

zones difficiles d'accès. Ces derniers peuvent en effet être propulsés en grand nombre dans une zone à surveiller, et doivent donc pouvoir s'auto-organiser pour communiquer. Il est donc primordial que les batteries durent le plus longtemps possible. De ce fait, la conservation de l'énergie est cruciale dans un RCSF d'autant plus que le fonctionnement de tout le réseau en dépend. En effet, tout le réseau pourrait devenir non opérationnel si trop de nœuds sont perdus à cause de la dissipation complète de leur énergie, car ceci peut engendrer des pertes de communication dues à de trop grandes distances entre les capteurs. La problématique de conservation de l'énergie a été abordée par un nombre considérable de travaux de recherche (Cardei *et al.*, 2005; Raghunathan *et al.*, 2002) et elle concerne toutes les couches de communication du modèle OSI. Généralement, elle est traitée à différents niveaux que nous pouvons catégoriser sommairement comme suit (Cardei *et al.*, 2005; Chong et Kumar, 2003)

1. la planification optimale des états des capteurs pour alterner entre un état où le capteur est allumé et un autre où il est éteint, tout en préservant le fonctionnement normal du réseau et la qualité de service requise ;
2. le routage avec conservation de l'énergie dissipée, incluant des mécanismes de formation de grappes et d'agrégation des données ;
3. le contrôle efficace de la puissance de transmission afin de trouver un compromis entre la consommation énergétique et la connectivité du réseau ;
4. la compression de données, ou codage de source, pour réduire le taux de données transmises ;
5. l'accès efficace au canal afin de limiter les collisions et les retransmissions au niveau de la couche de liaison.

Plusieurs techniques, associées aux différentes couches du modèle OSI, ont été proposées dans la littérature pour limiter la consommation d'énergie des capteurs, comme le codage distribué (Slepian et Wolf, 1973; Zixiang *et al.*, 2004), le routage basé sur le facteur énergie (Al-Karaki et Kamal, 2004; Aslam *et al.*, 2002), le contrôle du taux

d'échantillonnage et de transmission (Ngan *et al.*, 2005; Ye *et al.*, 2004), l'accès au canal sans collisions (Akyildiz *et al.*, 2002; Arisha *et al.*, 2002; Ngan *et al.*, 2005), le contrôle des états des capteurs (Ye *et al.*, 2004), la topologie hiérarchique en grappes (Heinzelman *et al.*, 2000; Mhatre et Rosenberg, 2004a). Ces techniques seront présentées avec plus de détail dans le prochain chapitre.

Comme nous le verrons dans l'énoncé de nos objectifs de recherche, nous nous intéressons dans cette thèse aux trois premiers points, i.e., la planification optimale des états des capteurs, la formation optimale de grappes et l'allocation optimale des puissances de transmission.

## 1.2 Éléments de problématique

À cause de leurs propriétés intrinsèques et de la nature de leurs applications, les capteurs sont généralement équipés de source d'énergie limitée adaptée à leur taille et leur faible coût. De plus, certaines applications exigent que les capteurs soient déployés à large échelle et éparpillés aléatoirement, ce qui rend le renouvellement de leur batterie non envisageable. La conservation de l'énergie des capteurs et le maintien du réseau fonctionnel le plus longtemps possible sont donc des défis importants qui n'étaient pas posés par les réseaux sans fil classiques, ni même les réseaux ad hoc. Tous les mécanismes, algorithmes et protocoles valables pour les réseaux sans fil classiques, ou les réseaux ad hoc, doivent donc s'adapter à cette contrainte énergétique, propre aux RCSF. Comme les RCSF sont destinés à rapporter les données relatives à leur environnement, l'une des applications phares de ces réseaux se trouve être la surveillance de zones géographiques. Un RCSF serait alors déployé dans un espace, ou une zone bien définie, pour surveiller certains événements, comme le mouvement, l'intrusion ou le passage d'objets ou de personnes, et rapporter ces événements lorsque ceux-ci ont lieu (Chong et Kumar, 2003). Un capteur est capable

de détecter un événement si ce dernier a lieu dans son rayon de détection, ou rayon de couverture. Dans le domaine militaire, les zones surveillées pourraient être des terrains hostiles ou des zones de combat et, dans le domaine civil, elles pourraient être des habitats ou des environnements domestiques. Dans tous les cas, il est généralement important d'assurer une couverture importante de la zone surveillée, c'est à dire que, si un événement a lieu à n'importe quel point géographique de la zone surveillée, il doit être détecté par au moins un capteur. Généralement, quand le réseau est dense, les rayons de couverture des capteurs voisins se chevauchent, et un événement peut être détecté et rapporté par plusieurs capteurs, ce qui est inefficace du point de vue énergétique puisque certains capteurs vont dissiper de l'énergie inutilement dans la capture, la traitement et la transmission radio des données captées. Ceci va donc diminuer la durée de vie du réseau et risque de corrompre sa fonction principale, qui est la couverture totale de la zone surveillée, si certains nœuds ne sont plus opérationnels à cause de l'épuisement total de leur batterie. Une première étape d'optimisation de l'énergie consommée consiste donc à déterminer un sous-ensemble optimal de couverture, c'est à dire un sous-ensemble de capteurs suffisant pour couvrir toute la zone surveillée. Ce sous-ensemble de capteurs serait allumé alors que le reste des capteurs serait éteint pour conserver son énergie. La Figure 1.4 montre un sous-ensemble de couverture.

Le calcul du sous-ensemble optimal sera à priori basé sur les énergies résiduelles des capteurs puisque, intuitivement, il serait plus intéressant d'allumer les capteurs ayant la plus grande énergie résiduelle pour conserver le plus grand nombre de capteurs opérationnels le plus longtemps possible. Aussi, il serait plus intéressant d'allumer un plus grand nombre de capteurs ayant une énergie résiduelle élevée que moins de capteurs ayant une énergie résiduelle plus basse car le deuxième scénario, bien que consommant moins d'énergie en total, risque d'épuiser plus rapidement l'énergie de certains capteurs, ce qui diminuerait la durée de vie du réseau. Une répartition équitable de la consommation énergétique devra donc être prise en compte dans le

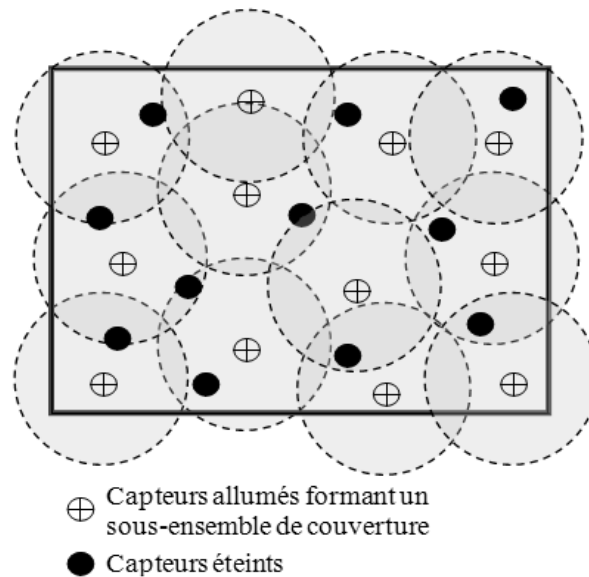


FIGURE 1.4 Sous-ensemble de couverture

processus de recherche du sous-ensemble optimal de couverture. Pour ne pas épuiser totalement l'énergie des mêmes capteurs sélectionnés et pour équilibrer la consommation d'énergie, le sous-ensemble optimal peut être re-calculé périodiquement, en tenant compte des valeurs actualisées des énergies résiduelles des capteurs.

Par ailleurs, et comme démontré au paragraphe précédent, la formation de grappes peut permettre de sauvegarder l'énergie des capteurs et prolonger la durée de vie du réseau en limitant la tâche de router les données jusqu'au CT aux têtes de grappes (TG). Un choix optimal des grappes est donc fondamental pour l'optimisation de la durée de vie du réseau. En effet, le choix et l'emplacement des TG vont influencer la consommation en énergie des capteurs élus TG ainsi que des capteurs membre de la grappe, qui doivent utiliser une puissance de transmission qui leur permet d'atteindre leur TG. La distance entre les membres de la grappe et leur TG est donc également importante.

Dans un contexte de couverture et dans un souci de trouver une configuration de capteurs optimale, le problème de formation de grappes ne peut pas être dissocié de celui du calcul d'un sous-ensemble optimal de couverture. En effet, le choix des TG



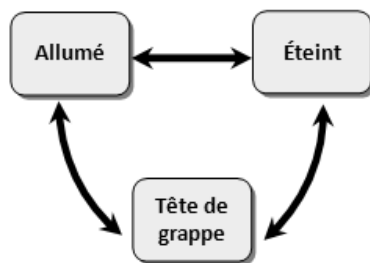


FIGURE 1.5 Les différents états qu'un capteur peut prendre

doit se faire à partir de l'ensemble des capteurs allumés, c'est à dire faisant partie du sous-ensemble optimal de couverture. Autrement dit, si ces deux problèmes sont traités séparément, la solution peut être non optimale ou même non admissible. Par exemple, si un sous-ensemble optimal de capteurs est d'abord déterminé et ensuite un ensemble optimal de TG est calculé à partir de ce sous-ensemble, la configuration a de fortes chances d'être sous-optimale car les capteurs éteints ne sont pas considérés alors qu'ils pourraient faire partie de la configuration optimale des grappes. Il faut donc considérer ces deux points dans un même problème d'optimisation global dans lequel on chercherait la meilleure configuration d'états de capteurs qui préserverait la couverture de la zone surveillée, chaque capteur pouvant être, comme le montre la Figure 1.5, dans l'un des trois états suivants : 1) éteint 2) allumé et TG 3) allumé sans être TG.

D'autre part, dans un réseau structuré en grappes, seules les TG sont habilitées à router les données. Ceci rajoute donc une contrainte dans le choix des TG potentielles : celles-ci doivent faire partie d'un arbre de recouvrement qui servira de topologie de routage. Plusieurs travaux ont abordé le routage énergétiquement efficace (Akkaya et Younis, 2005), la formation de grappes comme moyen de conservation de l'énergie (Arboleda et Nasser, 2006; Younis *et al.*, 2006) et la problématique de couverture (Cardei *et al.*, 2005; Carle et Simplot-Ryl, 2004; Inanc *et al.*, 2003; Li *et al.*, 2002) séparément. Plusieurs autres (Lu *et al.*, 2006; Xing *et al.*, 2005; Yan *et al.*, 2003)

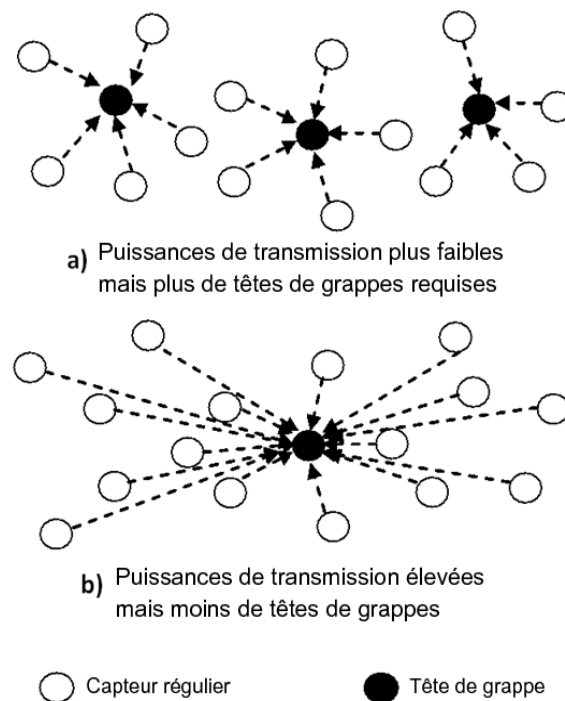


FIGURE 1.6 Compromis entre les puissances de transmission et le nombre de TG

ont abordé le problème de couverture sous contrainte de connectivité du réseau mais seulement dans le cas de réseau non hiérarchisé, et n'ont donc pas tiré profit des bénéfices énergétiques offerts par les réseaux structurés en grappes. Dans le Chapitre 3 de cette thèse, nous abordons le problème intégral de l'allocation optimale des états aux capteurs dans un RCSF sous contrainte de couverture, de structuration en grappes et d'existence d'une topologie logique de routage reliant toutes les TG.

Jusqu'ici, la formation des grappes consistait à choisir l'ensemble optimal des TG, et chaque capteur allumé devra se connecter à l'une des TG se trouvant dans son rayon de transmission. Toutefois, si les capteurs sont munis de puissance de transmission variable, l'ensemble de TG potentielles qu'ils peuvent atteindre dépendra de la puissance de transmission utilisée et aura un coût énergétique différent. Comme le montre la Figure 1.6, une puissance de transmission plus élevée exigera moins de

TG pour connecter tous les capteurs et donnera plus de flexibilité de connectivité mais consommera plus d'énergie. Autrement dit, avec des puissances de transmission élevées, l'énergie sera principalement consommée par les radios des capteurs et avec des puissances de transmission plus faibles, l'énergie sera plutôt consommée par les TG, plus nombreuses. Un compromis devra donc être trouvé entre les deux scénarios extrêmes de la Figure 1.6.

Plusieurs travaux de recherche ont abordé le problème de l'allocation des puissances de transmission aux capteurs en tenant compte de la conservation de l'énergie (Mhatre *et al.*, 2005; Sun *et al.*, 2008). Certains ont abordé le problème de l'ajustement optimal du rayon de transmission mais sans considérer la contrainte de couverture (Liu et Lin, 2005; Sun *et al.*, 2008), et d'autres (Kubisch *et al.*, 2003; Panichpapiboon *et al.*, 2006) ont abordé le problème de l'allocation optimale des puissances de transmission mais purement du point de vue de la théorie de l'information, sans considérer ni la contrainte de couverture, ni celle de la formation de grappes. Finalement, quelques travaux (Mhatre *et al.*, 2005) ont abordé le problème global de l'allocation des puissances de transmission dans un contexte de formation de grappes et sous contrainte de couverture, mais ces travaux considèrent généralement un ensemble pré-défini de capteurs habilités à devenir TG, ce qui limite le nombre et la flexibilité des configurations admissibles et ne répond donc pas aux exigences d'optimalité recherchées dans le cas de notre problématique générale. Pour ces raisons, nous nous proposons de traiter le problème d'allocation optimale des puissances et des états dans un RCSF, sous contraintes de formation de grappes et de couverture. Ce problème fera l'objet du Chapitre 4 de cette thèse.

Dans le même ordre d'idées, des algorithmes centralisés ou répartis peuvent être utilisés pour la formation efficace de grappes, qui a souvent été appliquée aux RCSF pour la conservation de l'énergie (Akkaya et Younis, 2005; Yan *et al.*, 2003; Younis *et al.*, 2006) et la ré-utilisation de la bande passante (Yao et Giannakis, 2005). Les algorithmes centralisés de formation de grappes ont l'avantage de se baser sur une vue

globale du réseau pour générer des solutions optimales. Cependant, les algorithmes répartis sont généralement plus adaptés au passage à l'échelle (*scalability*), malgré leurs solutions sous-optimales dues au fait que les décisions sont prises localement, par chaque nœud, à partir des informations recueillies de son voisinage. Dans la majorité des algorithmes répartis rencontrés dans la littérature (Chatterjee *et al.*, 2002; Ding *et al.*, 2005; Heinzelman *et al.*, 2000; Younis et Fahmy, 2004), l'élection des TG est généralement basée sur les énergies résiduelles des capteurs et leurs degrés respectifs. Les capteurs se basent également sur une vue limitée de leurs voisins directs pour prendre leur décision de s'auto-élire TG ou non. Ceci fait qu'en général, les grappes formées ne sont pas efficaces. Au Chapitre 5 de cette thèse, nous abordons la formation répartie de grappes, en proposant un nouveau mécanisme qui intègre efficacement l'information sur les nœuds de deuxième niveau (i.e., séparés de deux sauts) dans la prise de décision de chaque capteur.

### 1.3 Objectifs de recherche

Cette thèse traite principalement la problématique de l'optimisation de la consommation d'énergie et le prolongement de la durée de vie dans les réseaux de capteurs sans fil. L'objectif principal est de proposer des mécanismes efficaces de gestion des activités des capteurs afin de maximiser la durée de vie du réseau, tout en garantissant certaines contraintes fonctionnelles, telles que la couverture d'une zone surveillée et le routage des données collectées. Plus spécifiquement, dans cette thèse nous visons à :

- analyser, par une revue critique de la littérature, les mécanismes existants de conservation de l'énergie et d'optimisation de la durée de vie dans les réseaux de capteurs, afin d'y déceler certains problèmes encore ouverts ou dont les solutions proposées sont encore à améliorer ;

- proposer de nouveaux mécanismes efficaces de gestion de l'activité des capteurs, qui permettent de prolonger la durée de vie du réseau, tout en garantissant le niveau de service exigé par la fonction remplie par celui-ci ;
- évaluer la performance des mécanismes et algorithmes proposés, par rapport à des bornes supérieures connues et aux meilleurs mécanismes connus dans la littérature scientifique récente.

#### 1.4 Principales contributions

Les contributions principales de cette thèse touchent essentiellement l'optimisation de la consommation énergétique d'un réseau de capteurs et le prolongement de sa durée de vie. Plus spécifiquement, nous avons traité trois aspects de cette problématique. Dans un premier temps, nous avons abordé le problème de la maximisation de la durée de vie du réseau par une planification optimale des états des capteurs sous contrainte de couverture, de formation de grappes et de routage entre les TG. Ensuite, nous avons abordé le même problème de maximisation de la durée de vie du réseau par une planification optimale des états des capteurs, sous contraintes de couverture et de formation de grappes, mais dans un contexte de capteurs munis de plusieurs puissances de transmission. Enfin, nous avons proposé un nouvel algorithme réparti de formation de grappes dans les réseaux de capteurs, qui a prouvé son efficacité par rapport à l'existant, en termes de durée de vie du réseau. Ces contributions peuvent être détaillées comme suit :

1. **Solution au problème de planification optimale des états des capteurs sous contrainte de couverture, de formation de grappes et de routage entre les TG** : nous avons modélisé le problème de planification des états des capteurs dans un RCSF, dans le but de maximiser la durée de vie du

réseau sous la triple contrainte de couverture, formation de grappes et routage. Les solutions du modèle proposé constituent une borne supérieure, en termes de durée de vie du réseau, pour tout protocole réparti abordant le même problème. Notre modèle peut donc servir comme référence de comparaison pour évaluer la performance d'autres protocoles répartis traitant le même problème. À notre connaissance, aucun travail préalable dans la littérature n'a proposé une telle borne pour le problème global de planification optimale des états des capteurs et sous les mêmes contraintes générales de couverture, de routage et de formation de grappes.

2. **Solution au problème de planification optimale des états des capteurs, sous contraintes de couverture et de formation de grappes, dans un contexte de capteurs munis de plusieurs puissances de transmission** : nous avons modélisé le problème d'allocation optimale des états et des puissances de transmission aux capteurs dans un RCSF muni de puissance de transmission variable, sous contraintes de couverture et de structuration en grappes du réseau. La plupart des algorithmes proposés dans la littérature pour traiter ce problème sont répartis et les décisions prises par chaque nœud se basent sur des informations locales. Les configurations trouvées par ces algorithmes sont donc sous-optimales et à notre connaissance, aucun travail de recherche n'a proposé une solution optimale exacte à ce problème. Les solutions que nous proposons offrent une borne supérieure, en termes de durée de vie du réseau, aux algorithmes répartis traitant le même problème.
  
3. **Nouvel algorithme réparti de formation de grappes dans les réseaux de capteurs, dans l'optique de maximiser la durée de vie du réseau** : nous avons proposé un nouvel algorithme réparti de formation de grappes où les nœuds prennent leurs décisions en fonction des informations récupérées,

via un échange efficace de messages, de leur voisinage à deux-sauts, ce qui produit des configurations plus efficaces. Notre algorithme génère de meilleures performances, en termes de durée de vie du réseau, que d'autres algorithmes répartis populaires dans la littérature scientifique relative au sujet. Il constitue donc une amélioration de l'existant en ce qui a trait aux algorithmes répartis de formation de grappes dans les RCSF.

## 1.5 Plan de la thèse

Dans cette thèse, nous avons opté pour le format "par articles". Certains chapitres sont donc la transcription d'articles publiés dans, ou soumis à, des revues scientifiques. Suite à ce chapitre d'introduction, le Chapitre 2 présente une revue critique et sélective de la littérature sur les problèmes clés relatifs aux réseaux de capteurs, plus particulièrement les mécanismes d'optimisation de l'énergie dépensée et de prolongement de la durée de vie dans les réseaux de capteurs sans fil, et ce, aux différentes couches du modèle OSI. Les différents algorithmes et mécanismes rencontrés dans la littérature nous ont permis de faire ressortir des problèmes et défis qui ont servi comme base de recherche pour cette thèse. Ensuite, le Chapitre 3 présente le premier article (Chamam et Pierre, 2009b) intitulé "On the planning of wireless sensor networks : energy-efficient clustering under the joint routing and coverage constraint" et publié dans la revue IEEE Transactions on Mobile Computing. Dans cet article, nous modélisons le problème d'allocation optimale des états aux nœuds dans un réseau de capteurs sans fil, sous contraintes de couverture, de structuration en grappes du réseau et de routage entre les TG, et nous proposons un mécanisme à base d'heuristique taboue pour résoudre ce problème, démontré NP-complet. Le Chapitre 4 présente notre deuxième article intitulé "Energy-efficient clustering with transmit power assignment under coverage constraint in wireless sensor networks", qui a été soumis à la revue *Wireless Communications & Mobile Computing (Wiley)*. Dans cet

article, nous abordons le problème de la maximisation de la durée de vie d'un réseau de capteurs sans fil dont les nœuds sont munis de plusieurs puissances de transmission. Nous y proposons un mécanisme à base d'heuristique taboue pour résoudre ce problème, démontré NP-Complet. Le Chapitre 5 présente notre troisième article (Chamam et Pierre, 2009a) intitulé "A distributed energy-efficient clustering protocol for wireless sensor networks" et publié dans la revue *Computers and Electrical Engineering (Elsevier)*. Dans cet article, un nouvel algorithme réparti de formation de grappes dans les réseaux de capteurs sans fil est proposé. Au Chapitre 6, nous faisons une discussion générale des différents résultats obtenus ainsi qu'une synthèse de nos contributions scientifiques. Le Chapitre 7 conclut la présente thèse en mettant l'accent sur les principales contributions apportées et en dégagant les principales limitations de nos travaux. Des recommandations pour des travaux futurs y sont également apportées.



## CHAPITRE 2

### REVUE DE LITTÉRATURE

#### MÉCANISMES DE CONSERVATION DE L'ÉNERGIE DANS LES RÉSEAUX DE CAPTEURS SANS FIL

Tout comme les autres types de réseaux sans fil, les RCSF reposent sur un modèle de communication en couches. Étant donné que la limitation en énergie est la principale contrainte caractérisant les RCSF, la plupart des algorithmes et protocoles de communication proposés pour ces réseaux prennent en grande considération l'optimisation de l'énergie dissipée. Comme illustré à la Figure 2.1, la conservation de l'énergie relève de toutes les couches de communication et passe généralement par une gestion efficace des états, des rôles et activités des capteurs (Anastasi *et al.*, 2009). Par exemple, un capteur peut décider de baisser son taux de transmission pour conserver son énergie quand sa couche MAC détecte trop de collisions sur le canal. Aussi, un capteur peut décider de ne plus participer au routage quand il détecte que son énergie a baissé au dessous d'un certain seuil ou encore décider de participer au routage pour sauvegarder l'énergie de ses voisins en leur évitant des transmissions de grande portée. Dans ce chapitre, nous allons passer en revue différents mécanismes de conservation de l'énergie dans les RCSF, aux différentes couches de communication du modèle OSI. Nous commencerons tout d'abord par présenter certains protocoles et mécanismes qui régissent le fonctionnement spécifique des couches dans les RCSF. Ensuite, nous présenterons quelques protocoles pertinents intégrant la composante de conservation de l'énergie, aux différentes couches de communication. Enfin, nous pas-

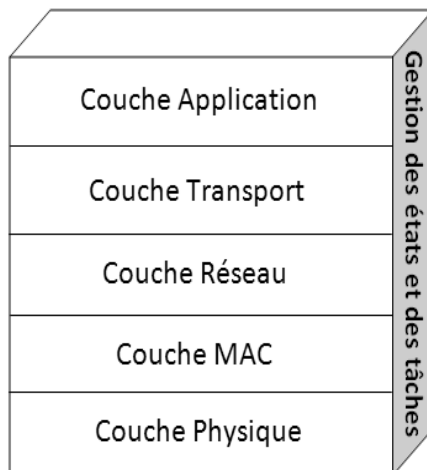


FIGURE 2.1 Modèle de communication en couches et plan de gestion de l'énergie

serons en revue certains mécanismes d'optimisation de l'énergie, basés sur le contrôle de la topologie, telles que la formation de grappes et la gestion optimisée des états des capteurs, et sur le contrôle de la puissance de transmission.

## 2.1 Architecture des capteurs et modèle de communication en couches

Les RCSF implémentent un modèle de communication réduit à 5 couches (Figure 2.1) : la couche physique qui gère la compression et le codage des données, la modulation et la gestion de la puissance de transmission ; la couche liaison qui gère l'accès au canal et les retransmissions en cas de collisions ; la couche réseau qui gère le routage de l'information ; la couche transport qui s'assure que les données parviennent de bout en bout et dont l'implémentation est tributaire de l'application en question ; et une couche application dédiée à la communication avec le RCSF.

La couche applicative doit permettre à l'utilisateur d'interroger le réseau, par exemple sur la valeur d'une certaine variable physique, ou encore l'endroit spécifique où une certaine variable satisfait à un critère donné. COUGAR (Cornell, 2009) est un exemple d'interface-usager qui permet au CT d'interroger le RCSF moyennant des

requêtes dans un langage évolué inspiré de SQL. Le réseau est alors considéré comme une base de données répartie à laquelle le CT peut émettre des requêtes relatives au flux de données. Les capteurs reçoivent ces requêtes, les exécutent et renvoient les résultats. Par analogie à COUGAR, TinyDB (Madden *et al.*, 2005) est une interface intégrée au système d'exploitation TinyOS (TinyOS, 2009) et qui représente le RCSF sous forme de table dont les colonnes sont les types de données captées (par exemple température, humidité, mouvement, etc.). La collection des données est initiée par des événements spécifiques, ce qui permet de sauvegarder l'énergie. Un langage semblable à SQL permet à l'utilisateur d'interroger le centre de traitement, qui, à son tour analyse la requête et la transmet aux TG qui la transmettent à leurs nœuds. La couche transport, quant à elle, a une importance particulière dans les RCSF car elle assure l'acheminement de bout en bout des messages, dans un environnement peu fiable à cause des conditions du canal radio et de l'instabilité de la topologie du réseau, qui peuvent provoquer la perte de messages. Cependant, relativement peu de travaux ont été proposés spécifiquement pour les RCSF. Les protocoles de la couche transport dépendent du degré critique des données (militaire, environnement, santé, etc.). En effet, le protocole TCP classique n'est pas adapté aux RCSF car il se base sur les acquittements échangés entre l'émetteur et le récepteur à chaque message reçu, et sur des retransmissions chaque fois qu'un message est perdu. Les ACKs et les retransmissions dissipent de l'énergie et doivent donc être évités lorsqu'ils ne sont pas nécessaires. Seddik-Ghaleb *et al.* (2006) ont proposé et évalué expérimentalement quelques variantes de TCP qui offrent une meilleure consommation énergétique. Au niveau de la couche réseau, le routage des données a une importance capitale dans les RCSF et il soulève beaucoup de défis à cause de plusieurs contraintes intrinsèques aux RCSF comme l'absence d'adresse ou d'identifiant propre à chaque nœud, les changements topologiques fréquents (disparition d'un nœud ou sa non participation au routage, etc.) auxquels tout protocole de routage doit s'adapter efficacement. Il existe principalement trois familles de protocoles de routage (Akyildiz *et al.*, 2002;

Al-Karaki et Kamal, 2004; Cayirci, 2004) : les protocoles orientés données, les protocoles hiérarchiques et les protocoles orientés position géographique. Dans le routage orienté données, ce ne sont pas des nœuds spécifiques qui sont interrogés sur les données mais plutôt le réseau, en tant qu'entité, qui est interrogé sur un attribut particulier du phénomène surveillé. Par exemple, le réseau pourrait être interrogé sur "les zones où la température dépasse 30 degrés Celsius". Pour cela, l'adressage, ou le nommage, est basé sur les attributs d'un phénomène (Elson et Estrin, 2001; Shen *et al.*, 2001). Ce type de routage utilise généralement la diffusion (*flooding*) pour propager l'information relative à une certaine donnée. Les nœuds intéressés par une information ou une catégorie de données spécifique en font une requête qui est propagée par *flooding* sur tout le réseau et les nœuds possédant l'information requise répondent en propageant cette information sur le réseau. Les nœuds qui ne sont pas intéressés par l'information en question devront aussi participer au routage même s'ils n'ont pas déclaré leur intérêt pour cette information. Pour éviter un trop grand volume de données acheminées des nœuds vers le CT, certains nœuds peuvent faire une fusion de données provenant de plusieurs voisins avant de router l'information vers le CT. Le *flooding* a cependant plusieurs inconvénients connus comme l'implosion, qui survient lorsque les  $N$  voisins d'un nœud  $A$  reçoivent tous de  $A$  un message  $m$  qu'ils retransmettent à un voisin commun  $B$ .  $B$  reçoit alors  $N$  copies du même message  $m$ . Pour être efficace dans le contexte des RCSF, le *flooding* doit être intégré au sein d'un protocole qui tient compte de l'énergie résiduelle des nœuds et de la corrélation spatiale. Le routage sera alors sélectif en fonction de cette énergie restante. Pour éviter l'implosion, il existe une variante du *flooding* qui est le bavardage (*gossiping*) (Hedetniemi et Liestman, 1988) et qui consiste à sélectionner au hasard un seul voisin à qui les données seront acheminées. L'inconvénient évident de ce protocole est la lenteur d'arrivée du message à son destinataire final, surtout si le réseau est dense. Une autre variante du *flooding* classique est la famille de protocoles SPIN (*Sensor Protocols for Information via Negotiation*) (Heinzelman *et al.*, 1999).

C'est une famille de protocoles de routage qui améliore la conservation de l'énergie des nœuds au niveau du routage. Dans SPIN, un nœud qui émet ou transfère des données "propose" à ses voisins une "image" de ces données ("méta-donnée") avant de les émettre, via un message ADV. Le (ou les) voisin qui serait intéressé à recevoir ces données le fera savoir au nœud émetteur par un message REQ. Ensuite, les nœuds intéressés recevront les données via un message DATA. L'avantage de SPIN est d'éviter la surcharge du réseau et de sauvegarder l'énergie des nœuds qui ne sont pas intéressés à recevoir certaines données, surtout si celles-ci sont volumineuses. Toujours dans le but d'éviter l'implosion et la surcharge du réseau par des envois systématiques à tous les voisins, la diffusion orientée (*Directed Diffusion*) (Intanagonwiwat *et al.*, 2000) est un paradigme de routage orienté données qui consiste à établir un chemin pour chaque "information d'intérêt" annoncée et diffusée par le CT. L'information d'intérêt pourrait être, par exemple, "les zones où un mouvement a été détecté". Une première étape consiste à établir le chemin associé à cette information. Ensuite, lorsqu'un capteur répond à cet intérêt (dans notre exemple, un capteur qui détecte un mouvement), il véhicule l'information via le chemin associé à cet "intérêt". Ce chemin est rafraîchi à chaque période de temps et il a une durée limitée.

Les protocoles de routage hiérarchiques (Akkaya et Younis, 2005; Arboleda et Nasser, 2006; Heinzelman *et al.*, 2000; Mhatre et Rosenberg, 2004b; Younis *et al.*, 2006), quant à eux, se basent sur la subdivision des nœuds du réseau en grappes (*clusters*). Chaque grappe dispose d'une tête de grappe (TG), élue ou désignée, qui constitue le routeur ou la passerelle pour toute la grappe, i.e. le nœud qui récolte, fusionne, filtre les données et les achemine au CT. Les TG peuvent former, elles aussi, un autre niveau de grappes ayant une tête, et ainsi de suite. Ceci donne au réseau une structure hiérarchique arborescente qui conserve l'énergie des nœuds et allège le réseau car :

- (1) les données suivent, dans un intervalle de temps donné, un seul chemin allant des nœuds vers les TG successives jusqu'au CT et
- 2) les données des nœuds d'une

grappe, qui présentent généralement une corrélation spatiale et/ou temporelle, sont fusionnées et filtrées au niveau de la TG, ce qui diminue le taux de données routées vers le CT. Comme le traitement et la transmission des données consomment de l'énergie, les nœuds d'une même grappe peuvent occuper à tour de rôle la fonction de TG, pour équilibrer la consommation d'énergie entre eux, comme proposé par Heinzelman *et al.* (2002). Nous reviendrons plus en détail sur les mécanismes de formation de grappes et leurs avantages en termes de conservation de l'énergie.

Les protocoles orientés position (Akkaya et Younis, 2005; Akyildiz *et al.*, 2002; Al-Karaki et Kamal, 2004; Cayirci, 2004; Xu *et al.*, 2001) supposent connues les positions des nœuds dans un repère fixe et le RCSF est représenté par un graphe dont les nœuds représentent les capteurs et les arcs les connexions entre les nœuds qui peuvent communiquer directement entre-eux. Le routage est alors effectué en calculant, entre chaque paire de nœuds, le chemin le plus court en utilisant comme métrique l'énergie consommée dans la communication. Le principal inconvénient des protocoles de routage orientés position géographique est la connaissance de la position de tous les capteurs. Deux moyens existent pour estimer la position d'un nœud : l'utilisation d'un GPS ou la localisation à l'aide de nœuds de référence dont la position est connue d'avance. La première solution est très coûteuse et irréaliste dans le cas d'un RCSF à cause du coût et du volume élevés d'un appareil GPS et de sa grande consommation d'énergie. La deuxième solution est réaliste mais présente des erreurs de précision parfois élevées.

Comme dans le cas des réseaux sans fil classiques, la gestion de l'accès au canal est opérée par la couche MAC, qui doit cependant être pensée différemment dans le cas des RCSF, puisqu'elle doit prendre en compte la contrainte énergétique. En effet, dans les RCSF, les protocoles de la couche MAC doivent être conçus de manière à réduire les retransmissions ainsi que les temps d'écoute active pour réduire l'énergie consommée par les capteurs (Akyildiz *et al.*, 2002; Anastasi *et al.*, 2009). Aussi, les

priorités d'accès au canal doivent être attribuées selon la nature du flux et selon l'application car, dans un RCSF, les nœuds sont coopératifs plutôt que compétitifs. Par exemple, quand un événement est détecté, la priorité d'accès devrait être attribuée au nœud qui détecte l'événement au lieu de partager le temps d'accès équitablement sur tous les nœuds, y compris ceux qui n'ont pas de flux à transmettre. La couche MAC devrait aussi tenir compte de la haute densité des nœuds qui, généralement augmente la corrélation entre les lectures des nœuds voisins, et qui rend la transmission de la totalité des données de ces nœuds inutile, car redondante. Donc, les paquets ayant subi une collision ne sont pas forcément tous retransmis.

L'accès au canal peut être géré au niveau MAC par réservation ou par contention (Akyildiz *et al.*, 2002; Cayirci, 2004; Chong et Kumar, 2003). Pour des raisons de simplicité d'implémentation, la plupart des protocoles basés sur la réservation utilisent la division temporelle, ou TDMA (Time Division Multiple Access), où il n'y pas de collision. L'unité de temps est alors divisée en intervalles de temps et chacun des capteurs partageant le même canal transmettra pendant un intervalle. Généralement, c'est la TG qui affecte les intervalles de temps aux nœuds de la grappe. Une période de réservation est donc nécessaire avant la transmission, durant laquelle les nœuds déclarent leur intention de transmettre et la priorité de leur flux. Ensuite, durant une période de transmission, les nœuds envoient leurs données lors des intervalles qui leur sont affectés ou se mettent en veille s'ils n'ont pas de données à transmettre. Par exemple, Arisha *et al.* (2002) proposent un protocole d'accès au canal basé sur TDMA et adapté au double rôle des nœuds dans un RCSF, c'est-à-dire la transmission de leurs données et le routage des données des nœuds voisins. En effet, dans ce protocole, le nombre d'intervalles de temps (*time slots*) pendant lesquels un nœud peut transmettre dépend du nombre de flux émanant de ce nœud. Par exemple, s'il doit transmettre un paquet et router deux autres paquets provenant de ses voisins, il se verra allouer trois intervalles de temps pour la transmission. Toutefois, les protocoles d'accès basés sur TDMA souffrent de latences élevées et d'une

relative incompatibilité au trafic en rafale (*bursty traffic*). De plus, une infrastructure hiérarchique basée sur les grappes est toujours nécessaire.

La technique d'accès basée sur la compétition (*contention-based*) consiste, pour les nœuds partageant le même canal, à écouter le canal et envoyer leurs données lorsque ce dernier est libre. Le protocole le plus connu est CSMA (Carrier Sense Multiple Access) qui est à la base de IEEE 802.11. Pour préserver leur énergie, les nœuds se synchronisent pour écouter le canal seulement lorsque leurs voisins sont entrain de transmettre et se mettent en veille lorsqu'il n'y a pas d'activité sur le canal. Les avantages de l'accès par compétition par rapport à TDMA sont : 1) la non-obligation d'avoir un nœud contrôleur qui affecte les intervalles de transmission aux nœuds ; 2) une meilleure adaptation aux flux en rafales et 3) de meilleurs délais. Le principal inconvénient est la forte consommation énergétique, car les nœuds dépensent beaucoup d'énergie dans l'écoute du canal et dans les retransmissions en cas de collisions, ces dernières étant plus probables lorsque le réseau est dense à cause, notamment, du problème du nœud caché. Il existe aussi des solutions hybrides qui utilisent une combinaison de techniques d'accès, où un schéma hybride TDMA-FDMA (Frequency Division Multiple Access) est utilisé (Akyildiz *et al.*, 2002; Cayirci, 2004). Cependant, l'utilisation de FDMA, ou encore CDMA (Code Division Multiple Access), par les capteurs reste encore discutable (Akyildiz *et al.*, 2002) à cause de la relative complexité des circuits associés et leur consommation énergétique.

## **2.2 Mécanismes de conservation de l'énergie dans les RCSF au niveau des différentes couches**

L'énergie dissipée par les réseaux de capteurs est répartie entre le traitement, la capture et la communication. En ce qui a trait à la communication, beaucoup de protocoles et mécanismes ont été proposés dans la littérature pour optimiser la dissipation de l'énergie, surtout au niveau des couches directement impliquées dans



la communication, i.e., les couches physique, liaison, réseau et transport. Dans ce qui suit, nous allons décrire quelques mécanismes de conservation de l'énergie au niveau de ces couches.

### *Couche physique*

Comme démontré par Pottie et Kaiser (2000), parmi toutes les opérations effectuées par un capteur (capture, traitement et transmission de données), la communication (émission et réception des données) est l'opération qui consomme le plus d'énergie. Au niveau de la couche physique, la compression, ou codage de source (Zixiang *et al.*, 2004), est une technique qui permet de réduire le taux moyen de données transmises par des capteurs dont les lectures sont corrélées selon une fonction de corrélation connue. L'idée de base découle des travaux de Slepian et Wolf (1973), qui ont établi le résultat qui suit. Soient deux sources numériques  $X$  et  $Y$  qui sont corrélées suivant une fonction de corrélation connue et qui transmettent leurs données à un récepteur  $R$ . Si  $X$  et  $Y$  ont des taux de transmission respectifs  $H(X)$  et  $H(Y)$  bits par symbole ( $H(S)$  étant l'entropie de la source  $S$ ) et ne communiquent pas entre eux, alors le récepteur  $R$  peut décoder sans perte les données de  $X$  et de  $Y$  si  $X$  (resp.  $Y$ ) transmet ses données à un taux supérieur ou égal à  $H(X/Y)$  (resp.  $H(Y/X)$ ).  $X$  et  $Y$  peuvent donc réduire leur taux de transmission initial, diminuant ainsi l'énergie moyenne consommée par la transmission. Wyner et Ziv (1976) ont étendu ce résultat au décodage avec perte. Le théorème de Slepian et Wolf (1973) et son extension par Wyner et Ziv (1976) ont été appliqués aux réseaux de capteurs où les nœuds présentent généralement une corrélation du fait de leur proximité géographique. Par exemple, Chou *et al.* (2004) proposent un algorithme de codage réparti, implanté sur un réseau de capteurs, qui permet aux nœuds de compresser d'une façon "aveugle" (i.e., sans connaître les données transmises par les nœuds voisins ni l'expression de la corrélation entre leurs données et celles de leurs voisins) leurs données à un taux spécifié par le CT. L'algorithme comprend une composante qui estime d'une façon

adaptative la corrélation entre les nœuds du réseau. La solution proposée par Chou *et al.* (2004) est adaptée aux RCSF car les nœuds, qui codent les données, sont munis d'un encodeur à faible complexité qui leur permet de consommer une énergie relativement faible dans le codage. La complexité est concentrée au décodeur, qui est implanté sur le CT ou sur tout autre nœud possédant suffisamment d'énergie et de capacité de calcul. Le problème majeur dans les solutions basées sur la compression demeure l'estimation de la corrélation entre les nœuds, qui est généralement très dynamique et varie en fonction de plusieurs facteurs comme les caractéristiques dynamiques du réseau (bruit sur le canal, type d'évanouissement du signal, etc.), l'événement rapporté par les capteurs, le mode de capture utilisé (vidéo, RF, etc.). D'autres techniques au niveau de la couche physique consistent en une planification dynamique du voltage qui alimente le processeur, en fonction des besoins de calcul de ce dernier. Cette famille de techniques s'appelle *Dyamic Voltage Scheduling (DVS)* (Pantazis et Vergados, 2007). Il s'agit de prévoir le comportement du processeur et d'adapter sa fréquence et son courant en entrée en fonction des tâches qu'il a accomplir. Pour cela, des algorithmes complexes appelés "planificateurs de voltage" (*Voltage Schedulers (VS)*) sont requis pour évaluer la charge future du processeur et adapter son courant d'alimentation et sa vitesse. Ces techniques demeurent cependant très complexes à implémenter de par la nature stochastique imprévisible, et généralement en rafale, de la distribution des tâches du processeur.

### ***Couche liaison***

La conservation de l'énergie des capteurs au niveau de la couche liaison passe principalement par deux techniques : la limitation du temps d'accès au canal, moyennant des mises en veille des capteurs, et une gestion plus efficace des transmissions afin de limiter les collisions. Ces deux techniques sont en fait liées car une limitation de l'accès au canal tend à baisser le taux de contention sur le canal et donc baisser la probabilité de collision et de retransmission, ce qui va sauvegarder l'énergie

des nœuds émetteurs. Généralement, les protocoles d'accès efficace au canal se divisent en deux catégories principales (Anastasi *et al.*, 2009) : l'accès sans collision basé sur TDMA (Time Division Multiple Access), l'accès avec contention ou une combinaison hybride de ces deux types. Dans la catégorie d'accès basé sur TDMA, l'axe du temps est divisé en trames temporelles, chacune subdivisée en intervalles (*time slots*) de communication indépendants. Chaque nœud se voit attribuer un ou plusieurs intervalles par trame pour transmettre/recevoir des données. Par exemple, Rajendran *et al.* (2003) proposent TRAMA, un protocole qui divise chaque trame de temps en deux périodes : l'une à accès aléatoire (avec contention) dans laquelle chaque capteur réserve sa période de transmission sans collisions, et l'autre à accès planifié (TDMA) où chaque capteur a accès au canal dans la période qu'il a réservée auparavant. Une fois la réservation faite, chaque nœud envoie un message *Synch* à tous ses interlocuteurs potentiels. De cette manière, les capteurs peuvent connaître les intervalles pendant lesquels ils doivent être actifs et peuvent se mettre en veille tout le reste du temps, jusqu'à la trame de temps suivante, ce qui va sauvegarder leur énergie.

Dans la catégorie des protocoles d'accès avec contention, Polastre *et al.* (2004) proposent le protocole B-MAC, qui est intégré au système d'exploitation TinyOS (TinyOS, 2009). Dans B-MAC, les capteurs effectuent, entre chaque deux périodes de veille successives, des écoutes périodiques du canal pendant un intervalle de temps fixe pour vérifier si des communications ont lieu. La fréquence des "réveils" pour écoute peut être adaptée aux besoins de l'application. Cerpa et Estrin (2004) proposent ASCENT, un mécanisme réparti qui permet aux nœuds d'adapter leurs états aux conditions du réseau et de changer de mode ("Actif", "En veille" ou "Passif en écoute") selon le besoin. Ainsi, quand un nœud détecte un taux élevé de perte de paquets, il en informe ses voisins qui décideront alors de passer de l'état "En veille" à l'état "Actif" et participer au routage. Un nœud peut aussi réduire son taux de transmission si le taux de perte de paquets est élevé et conserver ainsi son énergie.

Le fonctionnement d'ASCENT est affecté par deux paramètres : le taux de perte toléré et le degré de connectivité moyen du réseau. ASCENT reste cependant dépendant du protocole de routage dont les performances peuvent être dégradées par la lenteur d'action d'ASCENT, ce qui constitue l'un des inconvénients de ce dernier. En effet, si, pour des raisons de fiabilité et de qualité de service, un protocole de routage désire réserver deux routes : l'une pour router les données et l'autre pour la relève en cas de panne de la première, ASCENT peut empêcher la seconde route de se former en poussant certains nœuds à repasser à l'état "En veille" s'il juge qu'ils sont redondants. Ye *et al.* (2004) proposent S-MAC, un protocole adaptatif d'accès au canal qui introduit une modification sur la couche MAC afin de l'adapter aux contraintes énergétiques des RCSF. S-MAC favorise l'accès au canal aux nœuds ayant beaucoup de trafic à transmettre. S-MAC permet aussi aux nœuds de passer en mode "En veille" (*Sleep*) périodiquement lorsqu'il n'y pas de trafic à acheminer pour éviter l'écoute excessive en l'absence de trafic. Pour réduire les délais et les trames de contrôle engendrés par le "réveil" provoqué de certains nœuds suite à leur veille périodique, S-MAC utilise l'écoute adaptative qui consiste à provoquer un changement d'état des nœuds, passant de l'état "En veille" à "Écoute", selon le type et l'importance du trafic. Ci *et al.* (2005) proposent un mécanisme de conservation de l'énergie qui consiste à réduire les transmissions en évitant les collisions. Pour ce faire, la taille des trames est modifiée et continuellement adaptée aux conditions du canal. Quand le canal présente un bruit élevé, la taille de la trame est diminuée pour réduire le risque de collision et donc de retransmission. La taille des trames est prédite moyennant un filtre de Kalman étendu.

### *Couche réseau*

Certains travaux s'intéressent à minimiser et à équilibrer la consommation d'énergie des nœuds au niveau de la couche réseau. Ainsi, un bon nombre de protocoles de routage qui optimisent l'énergie dissipée (Al-Karaki et Kamal, 2004) ont été proposés

dans la littérature. Par exemple, dans la catégorie des protocoles de routage orientés données, SPIN (Heinzelman *et al.*, 1999) intègre la conservation de l'énergie grâce à la négociation. En effet, avant d'envoyer toutes les données, qui pourraient être volumineuses, un nœud diffuse d'abord dans tout le réseau une méta-donnée décrivant les données à transmettre. Seuls les nœuds qui manifestent leur intérêt pour cette méta-donnée recevront les données en entier, ce qui sauvegarde considérablement la bande passante et l'énergie des nœuds.

Dans le protocole présenté par Intanagonwiwat *et al.* (2000), et qui est un protocole basé sur la diffusion orientée (*Directed Diffusion*), la conservation de l'énergie est réalisée grâce à l'agrégation de données. En effet, les nœuds se trouvant autour d'un événement stockent l'information sur cet événement avec un certain gradient. Quand l'utilisateur, via le CT, manifeste un intérêt pour une certaine information (par exemple : les zones où un mobile en mouvement a été détecté à un instant donné), les capteurs dont les données répondent à cet intérêt vont propager leur information avec le gradient associé. Un nœud routeur intermédiaire effectue la fusion et l'agrégation de ces données en fonction de leurs gradients respectifs avant de les transmettre, ce qui élimine considérablement les redondances et diminue par la suite le taux de données transmises, conservant ainsi les ressources (énergie et bande passante). Shah et Rabaey (2002) proposent de répartir la charge de routage sur un ensemble de chemins sous-optimaux pour ne pas épuiser l'énergie des mêmes nœuds faisant partie du chemin optimal et ainsi prolonger la durée de vie du réseau. Le choix de ces chemins se fait via une fonction de probabilité qui exprime la consommation énergétique de chacun d'eux. La survivabilité du réseau est la métrique principale dans cette approche. Le protocole suppose que chaque nœud possède une adresse (ce qui est assez discutable comme hypothèse) et il se déroule en trois phases : 1) une phase d'initialisation durant laquelle les routes sont calculées et leurs coûts sont estimés en fonction de l'énergie résiduelle des nœuds qui les composent. Aussi, chaque nœuds se fait attribuer une probabilité de participer au routage, inversement

proportionnelle à son énergie résiduelle ; 2) une phase de communication où chaque nœud route ses paquets en choisissant aléatoirement parmi ses nœuds voisins, selon leurs probabilités respectives de participer au routage ; 3) une phase de maintenance des routes, pour garder à jour les tables de routage et les coûts des routes. Ce protocole offre une nette amélioration par rapport à la diffusion orientée. La conservation de l'énergie dans les protocoles de routage hiérarchiques est présentée en détail à la section 2.2.2.

Quant aux protocoles de routage orientés position, ils nécessitent la connaissance de l'emplacement géographique des nœuds pour optimiser le choix de la route. Par exemple, Xu *et al.* (2001) proposent le protocole *Geographical Adaptive Fidelity (GAF)* qui conserve l'énergie du réseau en éteignant les capteurs qui ne sont pas nécessaires pour le fonctionnement du réseau. Chaque capteur détermine sa position à l'aide d'un GPS et s'associe à un point d'une grille virtuelle regroupant des capteurs géographiquement proches. Les capteurs associés au même point de la grille sont considérés équivalents et ne sont donc pas requis de fonctionner simultanément. Certains nœuds peuvent donc se mettre en veille, conservant ainsi leur énergie résiduelle.

### *Couche transport*

Au niveau de la couche transport, quelques protocoles ont été proposés pour assurer une communication fiable de bout en bout dans les RCSF. Par exemple, *Reliable Multi-Segment Transport Protocol (RMST)* (Stann et Heidemann, 2003) est un protocole qui tire profit de la corrélation des données transmises et utilise certains nœuds du réseau comme mémoire cache pour garantir l'acheminement de l'information des capteurs vers le CT. L'arrivée des données au CT ne suppose pas l'arrivée de tous les paquets transmis mais plutôt de l'information globale qui est extraite de la corrélation et des données en cache, chaque fois que des paquets sont perdus. Bien que ce mécanisme contribue à conserver l'énergie en minimisant les retransmissions,

il demeure peu adapté aux RCSF car il nécessite de la mémoire cache et une certaine capacité de calcul au niveau de certains nœuds pour le calcul de la corrélation. Contrairement à RMST, *Event-to-Sink Reliable Transport protocol (ESRT)* (Akan et Akyildiz, 2005) est un protocole qui assure un transport fiable des données des capteurs vers le CT sans avoir recours à une mémoire cache au sein du réseau, ce qui minimise l'énergie dépensée par rapport à RMST. De plus, ESRT est adapté aux RCSF car il ne repose pas sur l'identification des nœuds mais plutôt sur celle de l'événement détecté, quelle que soit sa source. ESRT permet aussi de contrôler le flux des données allant des nœuds vers le CT et les algorithmes de contrôle de flux sont implantés au niveau du CT. Comme RMST, ESRT tire aussi profit de la corrélation entre les événements détectés. Cette corrélation est analysée au niveau du CT, qui ordonne aux nœuds de diminuer leurs flux respectifs dès qu'il détecte l'événement souhaité à partir des différentes lectures corrélées, sans garantir les arrivées individuelles des données envoyées par chacun des capteurs. Contrairement au chemin montant (*upstream*), le chemin descendant (*downstream*) nécessite des protocoles de transport plus fiables qui garantissent l'arrivée des données aux capteurs. De plus, grâce à la puissance d'émission élevée du CT, les données émises par ce dernier ne suivent pas nécessairement un chemin multi-sauts mais peuvent plutôt être envoyées en multicast ou passer par beaucoup moins de nœuds intermédiaires que le chemin montant. *Pump Slowly, Fetch quickly (PSFQ)* (Wan *et al.*, 2002) est un protocole de transport basé sur une garantie locale d'arrivée des données, c'est-à-dire que chaque nœud intermédiaire doit conserver une copie du paquet transmis en cache jusqu'à s'assurer que ce dernier est bien arrivé au nœud suivant sur le chemin. Ceci évite de devoir transmettre le paquet perdu du CT vers le nœud destinataire, mais plutôt de corriger localement une perte de paquet. Cependant, ce protocole ne permet pas de corriger localement la perte des paquets due à la congestion ou la saturation des tampons.

### 2.2.1 Conservation de l'énergie sous contrainte de couverture

L'optimisation de l'énergie dissipée sous contrainte de couverture dans les RCSF se base généralement sur l'alternance des états des capteurs entre un état d'activité et un autre de mise en veille. Par exemple, Xing *et al.* (2005) font une analyse géométrique de la relation entre la couverture de la zone surveillée et la connectivité du réseau. Ils proposent *Coverage Configuration Protocol* (CCP), un protocole qui permet de configurer dynamiquement le réseau pour garantir différents degrés de couverture en fonction des requis de l'application. Dans CCP, chaque capteur choisit son état (*Actif* ou *En Veille*) en fonction du degré de couverture des points d'intersection de son cercle de couverture avec celui de ses voisins. Il permet ainsi de garantir une  $K$ -couverture, c'est à dire la couverture de chaque point de surface par au moins  $K$  capteurs.

Lu *et al.* (2006) présentent l'algorithme *Scalable COverage Maintenance* (SCOM), un protocole localisé de maintien de couverture, comme celui proposé par Xing *et al.* (2005), et dans lequel les capteurs utilisent une politique de redondance pour décider s'ils doivent s'allumer ou se mettre en veille. SCOM propose d'implémenter, pour chaque capteur, un compteur d'attente (back-off timer) proportionnel à son énergie résiduelle. Ce mécanisme de compteur permet aux capteurs ayant une énergie résiduelle faible de décider de leur état avant ceux ayant une énergie résiduelle plus élevée, ce qui leur permet, quand ils se retrouvent redondants, de se mettre en veille avant les capteurs à énergie résiduelle plus élevée, favorisant ainsi la prolongation de la durée de vie du réseau.

Chamam et Pierre (2007a) proposent une heuristique centralisée qui calcule dynamiquement un ensemble presque-optimal de capteurs qui garantit un taux de couverture pré-défini, tout en assurant la connectivité du réseau quand le rayon de transmission est supérieur ou égal au double du rayon de couverture.

Yan *et al.* (2003) proposent une planification des états (*Actif/En veille*) des capteurs



de manière à ce que chaque point de la zone surveillée est couvert à tout instant. Les capteurs voisins échangent un temps de référence aléatoire  $T_{ref}$  à chaque période  $T$  et prennent la décision de se mettre ou non à l'état activé durant la prochaine période  $T$ . Bien que ce mécanisme de planification permette de répartir l'énergie dissipée sur l'ensemble des capteurs, il ne prend pas en compte l'énergie résiduelle de ces derniers dans le calcul du temps d'activation de chaque capteur, ce qui rend les nœuds avec une énergie résiduelle faible plus enclins à disparaître.

### 2.2.2 Conservation de l'énergie par la formation de grappes (*clustering*)

La formation de grappes est une technique qui consiste à subdiviser le réseau en sous-ensembles de nœuds, appelés grappes. Les capteurs sont alors groupés en sous-ensembles ayant chacun un capteur désigné comme TG et des nœuds membres connectés à la TG. Comme illustré à Figure 2.2, les nœuds membres transmettent leurs données à la TG qui, à son tour, les achemine jusqu'au centre de traitement, soit directement (Heinzelman *et al.*, 2000), soit en multi-sauts via des passerelles ou d'autres TG voisines (Mhatre et Rosenberg, 2004b). Les TG peuvent effectuer

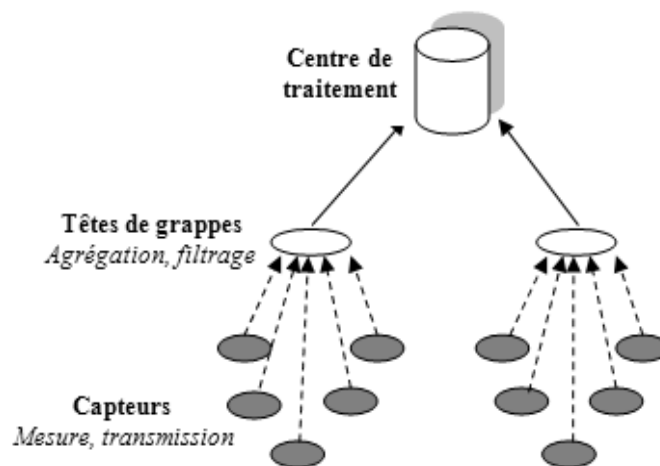


FIGURE 2.2 Structure en grappes d'un RCSF

certaines opérations comme le filtrage et l'agrégation des données collectées, ce qui permet de limiter la quantité de données transmises au CT et ainsi alléger la bande passante, diminuer la charge de traitement sur le CT et sauvegarder l'énergie des nœuds collecteurs. Dans certains cas, la formation de grappes offre des économies d'énergie importantes (Akkaya et Younis, 2005; Karl et Willig, 2005; Younis *et al.*, 2006) puisque seules les TG sont impliquées dans les tâches de collecte, routage et agrégation des données. De plus, la structuration hiérarchique en grappes allège la bande passante et permet sa réutilisation, ce qui contribue à augmenter la capacité du réseau (Yao et Giannakis, 2005).

Par ailleurs, l'organisation en grappes permet de mieux coordonner et ordonnancer les tâches des capteurs dans un RCSF. En effet, les TG peuvent ordonnancer l'activité des capteurs membres de leur grappe en ordonnant à certains capteurs de se mettre en veille ou de limiter leur puissance de transmission, par exemple dans le cas où il y a une redondance des événements rapportés ou des zones couvertes. Les TG peuvent également organiser les instants de transmission des capteurs de leurs grappes respectives afin d'éviter les collisions et par conséquent les retransmissions (Younis et Abbasi, 2007). Cela dit, comme les TG consomment plus d'énergie que les capteurs réguliers, il est important de disposer d'un mécanisme efficace de rotation des TG (Heinzelman *et al.*, 2000; Yao et Giannakis, 2005).

Plusieurs approches peuvent être suivies pour la formation de grappes : les TG peuvent être élues par les capteurs ou désignées d'avance ; elles peuvent faire partie de l'ensemble des capteurs ou d'un ensemble séparé de capteurs munis de plus grandes ressources ; elles peuvent demeurer fixes dans le temps ou dynamiques. Plusieurs protocoles de formation de grappes ont été proposés dans la littérature et ont généralement le prolongement de la durée de vie du réseau comme objectif principal. Pour cela, la formation de grappes est typiquement basée sur l'énergie résiduelle des nœuds et la proximité de ces derniers des TG potentielles (Arboleda et Nasser, 2006; Li *et al.*, 2002). Par exemple, *Low-Energy Adaptive Clustering Hierarchy*

(*LEACH*)(Heinzelman *et al.*, 2000) est l'un des premiers protocoles de formation de grappes proposés pour les RCSF, il est réparti, proactif et dynamique. Dans *LEACH*, la formation des grappes se base sur la puissance du signal reçu pour évaluer le degré d'un nœud et sa proximité de ses voisins. Chaque nœud prend sa décision de devenir ou non TG en se basant sur le nombre de fois qu'il a joué ce rôle dans le passé et sur une valeur pré-définie du pourcentage optimal des TG dans le réseau. Les TG sont les seuls nœuds qui peuvent router les données vers le CT, ce qui sauvegarde l'énergie des autres capteurs. *LEACH* offre une distribution équitable de la consommation énergétique sur les nœuds, via une rotation aléatoire des TG. Cependant, les TG envoient directement les données vers le CT, ce qui peut engendrer une forte consommation énergétique.

*Power-Efficient Gathering in Sensor Information Systems (PEGASIS)* (Lindsey et Raghavendra, 2002) et sa variante *Hierarchical-PEGASIS* sont deux améliorations de *LEACH*; plutôt que de former plusieurs grappes, *PEGASIS* forme une chaîne de capteurs voisins telle que chacun de ces capteurs envoie et reçoit à partir de ses voisins directs et un seul de ces capteurs est sélectionné pour envoyer directement les données au CT.

*Weighted Clustering Algorithm (WCA)* (Chatterjee *et al.*, 2002) est un algorithme réactif de formation de grappes, où l'élection des TG est basée sur l'évaluation, pour chaque capteur, d'une fonction score appelée "pondération combinée" (*combined weight*). Cette fonction score est une combinaison linéaire pondérée du degré, du niveau de mobilité, de la puissance de transmission et de l'énergie résiduelle d'un nœud. Chaque capteur diffuse sa *pondération combinée* à ses voisins et le nœud ayant le score le plus bas est élu TG.

*Hybrid Energy-Efficient Distributed Clustering (HEED)* (Younis et Fahmy, 2004) est un autre protocole réparti de formation de grappes qui utilise une combinaison hybride de l'énergie résiduelle et du coût de communication intra-grappe comme critère de sélection d'une TG. *HEED* favorise une distribution plus ou moins uniforme des

TG sur le réseau et ajuste la probabilité de sélection d'une TG pour garantir la connectivité du graphe formé par les TG. Dans la phase d'initialisation de HEED, chaque capteur  $i$  calcule sa probabilité  $P_i^{TG}$  de devenir TG. Cette probabilité est proportionnelle à son énergie résiduelle et à un taux pré-déterminé de têtes de grappes dans le réseau. Ensuite, durant une phase de répétition, chaque capteur cherche la meilleure TG dans son voisinage, pour se connecter à elle. Si un capteur ne trouve aucune TG dans son voisinage, il double sa probabilité  $P_i^{TG}$  et la diffuse de nouveau à ses voisins. Puis, il cherche la meilleure TG dans son voisinage et ainsi de suite. Le capteur  $i$  arrête le processus soit lorsque  $P_i^{TG} = 1$  (le capteur se promet lui-même TG), soit lorsqu'il trouve une TG dans son voisinage.

*Distributed Weight-based Energy-Efficient Hierarchical Clustering (DWEHC)* (Ding *et al.*, 2005) est un autre algorithme réparti de formation de grappes similaire à HEED. Dans DWEHC, chaque capteur calcule son score, fonction de son énergie résiduelle et de sa proximité de ses voisins. Dans un voisinage, le capteur ayant le score le plus élevé est élu TG et tous ses voisins pourront alors se connecter à lui. Ensuite, chaque membre non-TG du voisinage cherchera le chemin multi-sauts ayant le coût cumulé le plus bas, pour atteindre une TG, en passant par ses voisins non-TG. Pour cela, ces voisins diffusent leurs coûts respectifs pour atteindre la TG et le capteur en question pourra alors décider s'il se connecte directement à la TG ou s'il passe par le chemin de moindre coût.

*Energy-Efficient Hierarchical Clustering (EEHC)* (Bandyopadhyay et Coyle, 2003) est un autre algorithme réparti de formation de grappes à un ou plusieurs niveaux de hiérarchie. Dans le cas d'un seul niveau de hiérarchie, chaque capteur s'annonce à ses voisins éloignés de  $k$  sauts, comme TG potentielle, avec une probabilité  $p$ . Ensuite, chaque capteur se connecte à la TG la plus proche. Les capteurs qui ne reçoivent aucune annonce et qui ne sont pas TG deviennent alors des "TG forcées". Dans le cas de la formation de grappes à plusieurs niveaux de hiérarchie, le même processus est appliqué, à l'unique exception que seules les TG désignées participent au processus

d'élection des TG de niveau supérieur.

*Energy-Efficient Strong Head clustering (EESH)* (Zhou *et al.*, 2007) est un autre protocole de formation de grappes dans lequel les TG sont élues en fonction de leurs coûts, lesquels sont calculés à partir de leurs énergies résiduelles respectives, leurs degrés respectifs, les distances les séparant de leurs voisins et les énergies résiduelles de ces derniers. Le capteur ayant le coût le plus élevé est élu TG. Le processus d'élection se termine lorsque chaque capteur est soit élu TG, soit connecté à une TG voisine.

### **2.2.3 Conservation de l'énergie par l'ajustement optimisé des puissances de transmission**

Dans un RCSF où les capteurs sont équipés de plusieurs puissances de transmission ou d'une puissance ajustable, l'énergie consommée dépend directement de l'énergie de transmission, comme nous pouvons le noter aux Tableaux 2.1 et 2.2 pour le cas des capteurs MEDUSA-II et WINS. La conservation de l'énergie dissipée est donc tributaire d'une optimisation de la puissance de transmission utilisée, tout en respectant certaines contraintes du réseau comme, par exemple, la contrainte de connectivité de chaque capteur à une TG, ce qui implique l'utilisation d'une puissance minimale permettant d'atteindre au moins une TG voisine.

Certains travaux se sont penchés sur l'optimisation de la puissance de transmission des capteurs. Par exemple, Sun *et al.* (2008) considèrent un RCSF hétérogène où un ensemble pré-défini de capteurs avec des ressources "illimitées" est assigné le rôle de TG et où chacun des autres capteurs doit être connecté à une TG moyennant un lien direct ou multi-sauts. Les auteurs calculent un rayon de transmission optimal pour chaque ensemble de capteurs situés à égale distance de la TG la plus proche. Plus les capteurs sont proches de la TG, plus ils vont dissiper de l'énergie dans le routage des données provenant des nœuds plus lointains de la TG, mais moins ils vont dissiper de l'énergie dans la transmission. La dissipation de l'énergie dans le routage est

compensée par l'ajustement de la puissance de transmission. Le disque centré en la TG et contenant tous les capteurs connectés à cette TG est divisé en anneaux de largeur variable. Le rayon optimal de chaque anneau est calculé de manière à minimiser l'énergie totale dépensée par le réseau. Toutefois, le fait que l'ensemble des TG est pré-défini et fixé à l'avance a tendance à produire des configurations sous-optimales. De plus, la contrainte de couverture n'est pas abordée.

Liu et Lin (2005) proposent un protocole réparti qui permet d'ajuster dynamiquement la puissance de transmission de chaque capteur pour contrôler la taille des grappes, indépendamment de la distribution des nœuds. L'idée de base du protocole proposé est que chaque capteur ajuste sa puissance de transmission en fonction du nombre de ses voisins. Ensuite, un ensemble de TG est élu en fonction des énergies résiduelles. Ce protocole donne de bonnes performances en terme de durée de vie du réseau, mais n'aborde pas la couverture. De plus, le choix des TG, qui détermine le rayon de transmission des capteurs est sous-optimal puisqu'il se fait d'une manière relativement intuitive basée uniquement sur les énergies résiduelles.

Kubisch *et al.* (2003) présentent deux algorithmes répartis pour l'ajustement des puissances de transmission. Pour cela, le premier se base sur un seuil minimal du nombre de voisins et le second sur un seuil du nombre moyen des voisins. Les algorithmes proposés sont localisés : chaque capteur calcule une fonction score et décide soit d'augmenter sa puissance de transmission par un certain facteur si le score calculé est au dessous d'un certain seuil minimal, soit de la diminuer si le score calculé est au dessus d'un autre seuil maximal. Les protocoles proposés garantissent la connectivité du réseau mais n'abordent ni la couverture ni la formation de grappes. De plus, ils sont sous-optimaux à cause de la valeur aléatoire des seuils.

Panichpapiboon *et al.* (2006) évaluent analytiquement la puissance de transmission optimale dans le cas d'une topologie ordonnée et d'une manière expérimentale (par des simulations) dans le cas d'une topologie aléatoire. Ils font une étude de la corrélation entre la puissance de transmission optimale, le débit de données et la densité

spatiale des nœuds. De plus, ils étudient : 1) l'impact des différents facteurs d'évanouissement exponentiel du canal (sur différents liens d'une route multi-sauts) sur la performance de leur mécanisme de gestion de puissance et 2) la corrélation entre la puissance de transmission, la connectivité et la durée de vie du réseau. De plus, les auteurs font une analyse rigoureuse de la puissance de l'interférence et du taux d'erreur sur bit (BER) en utilisant une approche basée sur la théorie de la détection. La connectivité du réseau est alors abordée sous l'angle de la théorie de la communication : un lien multi-sauts existe entre deux nœuds si le BER moyen sur les différents segments qui composent la route est au dessous d'une valeur théorique  $BER_{th}$ . Cependant, cet article n'aborde ni le problème de couverture, ni la formation optimale de grappes.

Mhatre *et al.* (2005) considèrent un RCSF hétérogène de surveillance avec deux types de nœuds, capteur régulier et TG, répartis sur un espace à deux dimensions selon un processus de Poisson ayant des densités moyennes respectives de  $\lambda_1$  et  $\lambda_0$ . Les auteurs calculent les valeurs optimales des densités moyennes ainsi que les énergies initiales  $E_1$  et  $E_0$  des deux types de nœuds, qui assurent une durée de vie d'au moins  $T$  unités, tout en garantissant la connectivité de chaque capteur à une TG et un haut taux de couverture de la zone surveillée.

#### **2.2.4 Conservation de l'énergie par la planification optimisée des états des capteurs**

Beaucoup de travaux se sont intéressés à la planification optimale des états des capteurs et à l'ordonnancement optimal des activités de ces derniers, dans le but de minimiser l'énergie dissipée par le réseau. Il s'agit, d'une façon générale, d'alterner entre un (ou des) état(s) où le capteur est activé, et donc consomme un certain niveau d'énergie, et un autre où il est éteint et donc consomme une quantité d'énergie négligeable. Les tableaux 2.1 et 2.2 (Raghunathan *et al.*, 2002), illustrent les consom-

mations d'énergie respectives de deux types de capteurs : MEDUSA-II, développés au laboratoire de réseautique et des systèmes embarqués de l'Université de Californie à Los Angeles (UCLA), et WINS, développés par la compagnie Rockwell. Sur ces

TABLEAU 2.1 Consommation énergétique des différents états d'opération des capteurs MEDUSA II

État du capteur	État du module radio	Schéma de Modulation	Taux de transmission(kb/s)	Puissance consommée(mW)
Actif	Tx (0.7368 mW)	OOK	2.4	24.58
	Tx (0.0979 mW)	OOK	2.4	19.24
	Tx (0.7368 mW)	OOK	19.2	25.37
	Tx (0.0979 mW)	OOK	19.2	20.05
	Tx (0.7368 mW)	ASK	2.4	26.55
	Tx (0.0979 mW)	ASK	2.4	21.26
	Tx (0.7368 mW)	ASK	19.2	27.46
	Tx (0.0979 mW)	ASK	19.2	22.06
Actif	Rx	tout	tout	22.20
Actif	Écoute passive	tout	tout	22.06
Actif	Éteint	tout	tout	9.72
Éteint	Éteint	tout	tout	0.02

TABLEAU 2.2 Consommation énergétique, aux différents états d'opération, des capteurs WINS de Rockwell

État du capteur	État du module radio	Puissance consommée (mW)
Actif	Tx (36.3 mW)	1080.5
	Tx (19.1 mW)	9.86
	Tx (13.8 mW)	9426
	Tx (3.47 mW)	815.5
	Tx (2.51 mW)	807.5
	Tx (0.96 mW)	787.5
	Tx (0.30 mW)	773.9
	Tx (0.12 mW)	771.1
Actif	Rx	751.6
Actif	Écoute passive	727.5
Actif	Éteint	383.3
Écoute passive	Éteint	64



deux tableaux, nous pouvons noter la différence de consommation entre un état de veille et un état actif. Beaucoup de travaux proposent des mécanismes qui mettent en veille des capteurs, autant que possible, pour préserver leur énergie. Par exemple, Yao et Giannakis (2005) proposent un algorithme d'ordonnancement des transmissions des capteurs dans une architecture hiérarchisée avec un seul niveau de TG, où les capteurs disposent de séquences de données de différentes longueurs à transmettre durant une période de temps  $T$ . L'algorithme proposé trouve, pour chaque capteur, un ensemble optimal d'intervalles temporels (*time slots*) de transmission, qui minimise l'énergie dissipée par tout le réseau durant une période  $T$ . Cependant, durant une période  $T$ , tous les nœuds sont activés et aucune planification spatiale traitant la contrainte de couverture n'est considérée.

D'autres travaux se sont intéressés à la planification spatiale des états des capteurs, basée sur une activation sélective d'un sous-ensemble de capteurs, de manière à maximiser la durée de vie du réseau. Par exemple, Tian et Georganas (2002) proposent un algorithme réparti pour trouver un sous-ensemble optimal de couverture ou, si ce n'est pas possible, un ensemble de couverture offrant le minimum de points géographiques non couverts. Dans cet algorithme, si la portée entière d'un capteur est couverte par un sous-ensemble de ses voisins directs, alors ce capteur se met en veille. Un temps de *back-off* aléatoire garantit que deux capteurs ne prendront pas la décision de se mettre en veille simultanément. Cependant, ce protocole n'est pas optimal, à cause de la redondance dans la couverture, due à l'aspect aléatoire des mises en veille des capteurs et à l'aspect localisé des décisions prises par ces derniers.

Dans le mécanisme proposé par Ahmed *et al.* (2005), les rayons de transmission des capteurs suivent une distribution basée sur le modèle d'évanouissement exponentiel du canal. Un événement qui a lieu en dehors d'un certain rayon de détection d'un capteur est capté avec une certaine probabilité. Les probabilités de détection cumulatives augmentent le taux de couverture moyen, ce qui, par conséquent, réduit le nombre de capteurs activés et diminue donc l'énergie moyenne dissipée. Toute-

fois, l'algorithme proposé n'est pas optimal puisqu'il ne prend pas en considération l'énergie résiduelle des capteurs, ce qui peut amener les capteurs de moindre énergie résiduelle à disparaître avant les autres, provoquant ainsi un dysfonctionnement du réseau.

Hwang *et al.* (2006) proposent un mécanisme d'allocation des états aux capteurs dans un réseau hiérarchisé en grappes, avec garantie de couverture. Ce mécanisme réparti d'auto-organisation se base sur un réseau dense et affecte les états (*Actif, En veille, TG*) aux capteurs. Il commence par diviser les capteurs en grappes et définit, pour chaque grappe, des sous-ensembles de capteurs qui peuvent être mis à l'état *Actif* simultanément pour assurer la couverture totale de la zone surveillée. Un seul parmi ces sous-ensembles sera activé et sera maintenu jusqu'à l'épuisement complet de son énergie alors que les autres seront mis en veille. Quand un sous-ensemble n'est plus fonctionnel, un autre sera choisi.

### 2.3 Conclusion

La conservation de l'énergie dans les RCSF est un défi indéniable qui continue à attirer l'intérêt de la communauté scientifique. De nombreux algorithmes, mécanismes et protocoles ont été proposés dans la littérature scientifique pour traiter les problématiques de la minimisation de l'énergie dissipée par les capteurs et la maximisation de la durée de vie du réseau. L'approche de planification optimale des états, qui passe par la mise en veille de certains capteurs redondants ainsi que la formation de grappes, est particulièrement intéressante pour les nombreux avantages cités tout au long de ce chapitre. Un nombre considérable de travaux scientifiques a été publié dans ce sens et nous comptons, dans la suite de cette thèse, apporter notre modeste contribution à la recherche dans ce domaine.

**CHAPITRE 3**

**ON THE PLANNING OF WIRELESS SENSOR  
NETWORKS : ENERGY-EFFICIENT CLUSTERING  
UNDER THE JOINT ROUTING AND COVERAGE  
CONSTRAINT**

Ali Chamam and Samuel Pierre

Department of Computer & Software Engineering,

École Polytechnique de Montréal,

C.P 6079, succ. Centre-Ville, Montreal, H3C 3A7, Quebec, Canada.

*E-mail : {ali.chamam, samuel.pierre}@polymtl.ca*

**ABSTRACT**

Minimizing energy dissipation and maximizing network lifetime are important issues in the design of applications and protocols for sensor networks. Energy-efficient sensor state planning consists in finding an optimal assignment of states to sensors in order to maximize network lifetime. For example, in area surveillance applications, only an optimal subset of sensors that fully covers the monitored area can be switched on while the other sensors are turned off. In this paper, we address the optimal planning of sensors' states in cluster-based sensor networks. Typically, any sensor can be turned on, turned off or promoted cluster head, and a different power consumption level is associated with each of these states. We seek an energy-optimal topology that maximizes network lifetime while ensuring simultaneously full area coverage and sensor connectivity to cluster heads, that are constrained to form a

spanning tree used as a routing topology. First, we formulate this problem as an Integer Linear Programming model that we prove NP-Complete. Then, we implement a Tabu search heuristic to tackle the exponentially-increasing computation time of the exact resolution. Experimental results show that the proposed heuristic provides near-optimal network lifetime values within low computation times, which is, in practice, suitable for large-sized sensor networks.

**Key words :**

Wireless Sensor Networks (WSN), coverage, clustering, routing, network lifetime, energy efficiency, optimization, mathematical programming, Tabu search heuristic.

### **3.1 Introduction**

Wireless sensor networks (WSN) consist of a large number of limited-capability (power and processing) Micro-ElectroMechanical Systems (MEMS) capable of measuring and reporting physical variables related to their environment. In surveillance applications, sensors are deployed in a certain field to detect and report events like presence, movement or intrusion in the monitored area (Chong et Kumar, 2003). As depicted in Figure 3.1, data collected by sensors is transmitted to a special node equipped with higher energy and processing capabilities called "processing node" (PN) or "*sink*" (Chong et Kumar, 2003). The PN collects, filters and compiles data sent by sensors in order to extract useful information. Due to their energy constraints, wireless sensors usually have a limited transmission range making multi-hop data routing toward the PN more energy-efficient than direct transmission (one-hop). Energy conservation in WSN is critical and has been addressed by substantial research (Cardei *et al.*, 2005; Raghunathan *et al.*, 2002).

Generally, energy conservation is dealt with on five different levels (Cardei *et al.*,

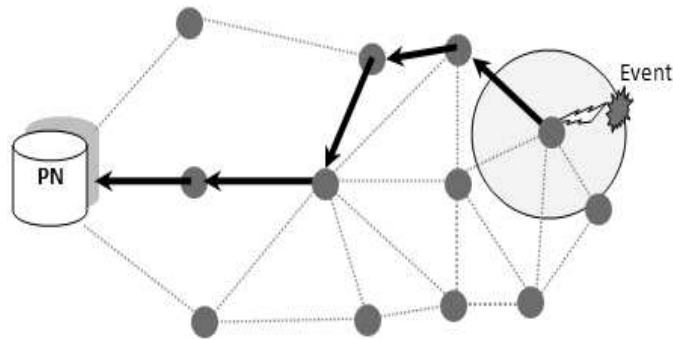


FIGURE 3.1 Multi-hop routing of collected data

2005; Chong et Kumar, 2003) :

1. efficient scheduling of sensor states to alternate between sleep and active modes ;
2. energy-efficient routing, clustering and data aggregation ;
3. efficient control of transmission power to ensure an optimal trade-off between energy consumption and connectivity ;
4. data compression (source coding) to reduce the amount of uselessly transmitted data ;
5. efficient channel access and packet retransmission protocols on the Data Link Layer.

The scope of this paper includes both the first and the second levels. We address the global problem of maximizing network lifetime under the joint clustering, routing and coverage constraint. We consider a sensor network that is deployed in a certain area  $A$  to monitor some given events. When the network is dense, sensing ranges of neighbor sensors usually overlap. That means that when an event occurs at a point  $P$  of  $A$ , it will be detected and reported by all the sensors whose sensing range encompasses  $P$ . This redundant transmission results in useless energy consumption. To save network energy and increase its lifetime, we propose to switch on only a

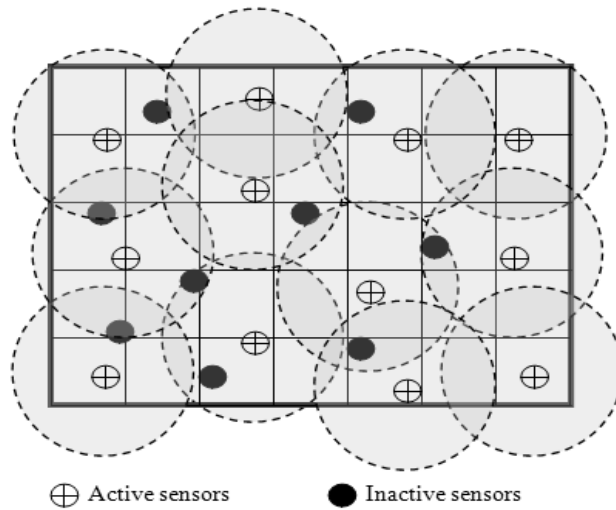


FIGURE 3.2 A full-covering sensor set

subset of sensors that covers  $A$  while all other sensors are turned off. Figure 3.2 depicts an example of full-covering sensor set.

On the other hand, clustering has been proven energy-efficient in WSN (Akkaya et Younis, 2005; Younis *et al.*, 2006). In cluster-based WSN, sensors are organized in clusters having, each, one sensor promoted as CH. All non-CH nodes transmit their data to their CH, which routes it to the remote PN. Clustering can provide for substantial energy saving (Akkaya et Younis, 2005; Arboleda et Nasser, 2006; Heinzelman *et al.*, 2000; Karl et Willig, 2005; Mhatre et Rosenberg, 2004b; Younis *et al.*, 2006) since only CH sensors are involved in routing and relaying data. Moreover, clustering alleviates bandwidth, enables its reuse and can, thus, increase system capacity (Yao et Giannakis, 2005). Besides, the fact that only the CH is transmitting information out of the cluster helps avoid collisions between the sensors inside the cluster and helps avoid the uncovered hole problem (Hwang *et al.*, 2006). However, since CHs consume more energy in aggregating and routing data, it is important to have an energy-efficient mechanism for CHs' election and rotation (Heinzelman *et al.*, 2000; Yao et Giannakis, 2005). In flat networks, sensors route data to the

PN through their peer sensors using one of the many routing protocols proposed in the literature (Akkaya et Younis, 2005). In contrast, in clustered networks, CHs transmit aggregated data to the PN, either directly (one-hop) (Heinzelman *et al.*, 2000) or in multi-hop (Mhatre et Rosenberg, 2004b). For the sake of minimizing energy consumption, both optimal number and optimal placement of CHs have to be sought.

In a cluster-based monitoring sensor network, any energy-efficient sensor scheduling mechanism has to guarantee a certain area coverage rate. Besides, the connectivity of every sensor to a CH has to be ensured at any time. Furthermore, for data to be routed from any CH to the PN, all CHs have to belong to a single connected graph. Hence, for sensors' states allocation to be optimal, coverage, connectivity of sensors to CHs and routing have to be taken into account within the same global planning process. When coverage and connectivity are dealt with separately, the obtained configuration may not be optimal. For example, an optimal covering subset of sensors can fail to guarantee network connectivity because some nodes are switched off or the optimally-designated CHs may belong to the set of switched-off sensors.

Many papers addressed separately energy-efficient routing (Akkaya et Younis, 2005), clustering (Arboleda et Nasser, 2006; Younis *et al.*, 2006) and area coverage (Cardei *et al.*, 2005; Carle et Simplot-Ryl, 2004; Inanc *et al.*, 2003; Li *et al.*, 2002). Many other works (Xing *et al.*, 2005; Yan *et al.*, 2003) addressed the integrated problem of maintaining area coverage and network connectivity but only on flat networks and did not take advantage of the potential energy saving and ease of manageability of cluster-based networks (Chong et Kumar, 2003; Karl et Willig, 2005; Younis *et al.*, 2006). To the best of our knowledge, the problem of maximizing sensor network lifetime under the integrated constraint of clustering, coverage and routing has not been addressed within the same global optimization process. In this paper, we address the optimal planning of cluster-based WSN under the joint routing and coverage constraint. In our architecture, any sensor can be active, switched off or upraised as CH and only

CHs can route data. We seek an optimal allocation of states to sensors, that maximizes network lifetime, while ensuring simultaneously full area coverage, connectivity of every sensor to a CH and connectivity of the overlay network composed of CHs. This paper is organized as follows : in the next section, we present some related work. In Section 5.3, we outline our problem and enumerate our assumptions. In Section 5.4, we mathematically formulate our problem as an Integer Linear Programming (ILP) model. In Section 5.5, we present our resolution method based on a Tabu search algorithm. In Section 5.7, we discuss our simulation results. Finally, Section 3.7 concludes this paper and points out some future directions.

### 3.2 Related work

The problem of maintaining both area coverage and network connectivity under energy constraint in WSN has been extensively addressed in the literature and many protocols were proposed to alternate sensor states between active and sleep in order to maximize network lifetime. For example, Xing *et al.* (2005) provide a geometric analysis of the relationship between coverage and connectivity and propose the CCP (Coverage Configuration Protocol) protocol that dynamically configures the network to guarantee different degrees of coverage depending on the application requirements. In CCP, every node decides its state (*Active* or *Sleep*) based on the coverage degree of the intersection points of its sensing circle with those of its neighbors. When coupled with any connectivity maintenance protocol, CCP offers connectivity and  $K$ -coverage. Lu *et al.* (2006) present SCOM (Scalable Coverage Maintenance), a localized coverage maintenance algorithm where sensors use the same redundancy eligibility rule as in (Xing *et al.*, 2005) to decide whether to turn on or off. SCOM implements, for each sensor, a back-off timer proportional to its residual energy. The back-off timer allows sensors with lower residual energies to decide about their states before sensors with more energy, making them more likely to turn off than the other



sensors, if they find themselves redundant. Chamam et Pierre (2007a) propose a centralized heuristic which dynamically calculates a near-optimal subset of sensors that guarantees a pre-defined coverage rate while ensuring network connectivity when the transmission range is greater than or equal to twice the sensing range. Yan *et al.* (2003) propose to schedule sensors' activities (*Active/Sleep*) so that every point in a grid monitored area is covered at any time. Neighbor sensors exchange a random reference time  $T_{ref}$  within cyclic rounds of constant duration  $T$  and decide to be active for a certain time duration within  $T$ . The round period  $T$  is equally shared among all the neighbor sensors that cover a common grid. Even though the proposed schedule balances consumed energy over neighbor nodes, it does not take into account the residual energy of sensors when calculating the activation time period of every node, which makes nodes with less residual energy more prone to expiration. However, all the works cited above do not address cluster-based architectures. Cluster formation is typically based on the energy reserve of sensors and sensor's proximity to the cluster head (Arboleda et Nasser, 2006; Younis *et al.*, 2006). Energy-efficient cluster-based routing algorithms for WSN have been widely addressed in literature (Akkaya et Younis, 2005). For instance, Low-Energy Adaptive Clustering Hierarchy (LEACH) (Heinzelman *et al.*, 2000), one of the most popular hierarchical routing algorithms for WSN, proposes to form clusters of sensor nodes based on the received signal strength and to use local cluster heads as routers to the sink. This saves energy since the transmissions will only be operated by cluster heads rather than all sensor nodes. Even though LEACH is completely distributed, it uses single-hop communication between cluster-heads and the sink, which is energy-consuming and not applicable to networks deployed in large regions. PEGASIS (Power-Efficient Gathering in Sensor Information Systems) (Lindsey et Raghavendra, 2002) and its variant Hierarchical-PEGASIS are two improvements of LEACH. Rather than forming multiple clusters, PEGASIS forms chains of sensor nodes so that each node transmits to and receives from a neighbor and only one node is selected from that chain to

transmit to the PN. But still, communication between the elected CH and the PN is made in one hop, that is not suitable for large networks. Energy-efficient sensor state scheduling mechanisms in cluster-based WSN also raised much interest in the research community. For example, Yao et Giannakis (2005) propose a scheduling algorithm for the one-level clustered WSN, where sensors have different data sequence lengths to transmit within a period of time  $T$ . The authors propose an Inverse-Log algorithm that finds, for every sensor, a set of optimal time allocations that minimizes the dissipated energy of the whole network over the period  $T$ . However, all sensors are activated during the time period  $T$  and no coverage constraint is considered. Besides temporal scheduling, other publications propose a spatial scheduling scheme based on the selective activation of sensors to maximize network lifetime (Ahmed *et al.*, 2005; Cardei *et al.*, 2005; Tian et Georganas, 2002). Tian et Georganas (2002) propose a localized algorithm that finds an optimal subset of sensors ensuring full area coverage or, if not possible, the least uncovered points. If the whole range covered by a sensor is covered by a subset of its neighbors, then the sensor decides to turn off. A random back-off time ensures that two nodes do not make the decision to turn off at the same time. When implemented over LEACH (Heinzelman *et al.*, 2000), the protocol proposed in Tian et Georganas (2002) shows some energy saving. However, this protocol is not optimal because of the uncontrolled coverage redundancy due to the random aspect of switch on/off of sensors. In Ahmed *et al.* (2005), sensors' sensing ranges follow a certain distribution derived from the channel characteristics and the log-normal path loss. An event occurring outside a certain range is still detected with a corresponding probability. The cumulative detection probabilities are shown to increase the mean area coverage which obviously decreases the number of sensors activated within the covering subset, thus reducing the consumed energy. However, the proposed algorithm is not optimal because it only takes into account each sensor's range and not its residual energy making nodes with very little residual energies prone to expiration. Hwang *et al.* (2006) propose a

cluster-based coverage-preserved node scheduling scheme. This mechanism assumes a dense network and assigns states (*Active*, *Sleep*, *Cluster head*) to sensors in a distributed and self-organized manner. The algorithm starts by dividing sensors into clusters and defines, for each cluster, a number of sponsor sensor sets that may be turned-on simultaneously. Only one among the computed sponsor sets is elected to be turned-on until it completely runs out of energy. Even though Hwang *et al.* (2006) provide an efficient coverage and clustering mechanism, they assume that sensed data is routed to the sink in one-hop, that may be energy-intensive for the relay nodes.

### 3.3 Problem statement and assumptions

In this paper, we consider a WSN deployed in an area  $A$  to monitor certain critical activities or events. As shown in Raghunathan *et al.* (2002) for the case of Rockwell's WINS seismic sensors, a sensor's radio can be in one of the following four activity modes, characterized by their respective power consumptions : *Transmit* (0.38-0.7 W), *Receive* (0.36 W), *Idle* (0.34 W) and *Sleep* (0.03 W). We note that when sensors are transmitting, receiving or idle, they have roughly the same energy consumption and can then be associated to a same state, *Active*, in which the sensor's radio is switched on.

Moreover, in this paper, we consider a cluster-based topology in which CHs route the data they receive from the non-CH sensors of their cluster to the PN through an overlay network solely composed of CHs. CH election and cluster formation are very important issues that deeply affect network lifetime of WSN and different approaches exist to implement these stages. For example, it is possible to use a fixed distribution of the sensor nodes and CHs, or use a dynamic algorithm for CH election. If CHs were chosen a priori and fixed throughout the system lifetime, they would quickly exhaust all their energy making them no longer operational. In this work, we propose to dynamically designate the set of CHs according to their residual energies, their

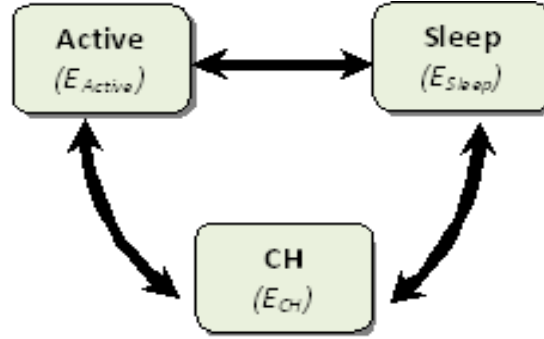


FIGURE 3.3 Sensor states

distance to their neighboring non-CH active nodes and their position within the graph formed by CHs.

As depicted in Figure 3.3, we will consider, without loss of generality, that each sensor can be in one of the three states : *Sleep*, *Active* and *Cluster Head (CH)* having respectively power consumptions  $E_{Sleep}$ ,  $E_{Active}$  and  $E_{CH}$  per time unit, where  $E_{Sleep} \ll E_{Active} < E_{CH}$ .

To control energy dissipation of the sensors that perform data relaying, we restrict the routing task to CHs. For that, there must exist a route from any CH to the PN. The most straightforward solution for that is to have a connected graph linking all the CHs. In our problem modeling, we propose that any admissible configuration must exhibit a spanning tree connecting all CHs, as shown in Figure 3.4. This makes the overlay network formed by the CHs sufficient to route data from any sensor toward the PN. The spanning tree construction is explained in the next section.

On the other hand, we assume that all sensors have the same sensing range  $R_s$  and that their detection model follows a binary probability function, also called *Disk model*(Cardei *et al.*, 2005; Xing *et al.*, 2005), in which an *Active* sensor  $i$  detects any event that occurs within its sensing range  $R_s$  with the same probability  $P_d = 1$ , whatever the distance from the event is. And, any event occurring outside  $R_s$  is not detected by  $i$ . Let us note that this binary probability sensing model is a simplified

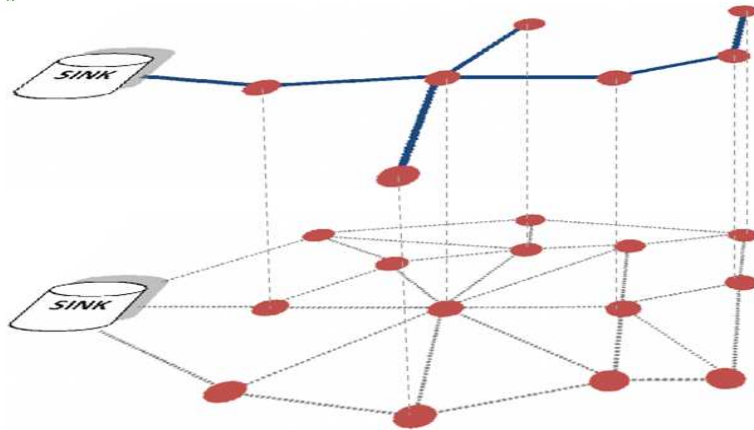


FIGURE 3.4 Overlay network composed of the selected cluster heads

representation of a sensor's detection probability and it does not necessarily reflect the received signal loss model on a real wireless channel. In fact, on a real wireless channel, because of signal fading and interference, events occurring far away from a sensor are less likely to be detected than events occurring close to it. In other words, an event occurring outside a sensor's range is still detected with a certain probability which decreases when the distance to the sensor increases. Many papers in the literature addressed probabilistic coverage in wireless sensor networks (Ahmed *et al.*, 2005; Hefeeda et Ahmadi, 2007) and proposed coverage protocols based on this model. However, to simplify the representation of the sensing model, substantial work in the literature (Cardei *et al.*, 2005; Xing *et al.*, 2005; Yen *et al.*, 2007) make the assumption of disk model. In this paper, we also make this simplifying assumption but we will show later that our optimization model and our proposed heuristic can be easily adapted to handle probabilistic coverage.

**Definition 1.** A set of sensors  $S_c$  is a covering set of area  $A$  if and only if :  $\forall$  point  $P \in A$ ,  $\exists i \in S_c$  such as  $i$  covers  $P$ .

In critical surveillance applications, it is important to guarantee that the monitored area is fully covered by sensors at every instant of the network lifetime. Hence, in

our problem, the optimal network configuration must :

- contain a full-covering set of active sensors ;
- contain a set of CHs so that every sensor is connected to a CH ;
- ensure that all CHs belong to a spanning tree over which data will be routed toward the PN.

Our objective is to find the network-lifetime-optimal allocation of sensors' states (*Active, Sleep, CH*) that meets these three conditions. Before modeling our problem, we make the following assumptions :

1. Each sensor has a unique ID, known to the PN and to the sensor itself.
2. The position of each sensor is fixed and known to the PN. The location information can be obtained either through a GPS (Global Positioning System), as assumed by Cardei *et al.* (2005) (but this technique is still expensive due to the high cost of placing a GPS on each sensor), or using one of the many GPS-free localization techniques proposed in the literature (Chu et Jan, 2007; Patwari *et al.*, 2005; Rudafshani et Datta, 2007). However, in this paper, we do not specifically address any localization technique and we assume that, whatever the localization technique used, sensors' location information is available at the PN.
3. Active sensors capture events occurring in their sensing range and transmit data associated with these events straightaway, without any buffering, because sensors are usually not equipped with large (and costly) buffers.
4. All sensors have the same sensing range  $R_s$  and the same transmission range  $R_t$ . All CHs have the same transmission range  $R_t^{CH} > R_t$ .
5. Only the CHs can perform data routing. Routing over the overlay network composed of CHs can be performed using one of the energy-efficient routing protocols for WSN proposed in the literature (Akkaya et Younis, 2005). However, we do not address any specific routing protocol, we only guarantee the existence of a routing topology.

6. Each sensor has an initial energy  $E_0$ . The PN has no energy limitation. Besides, we assume that, when a sensor is *Active*, it has a constant energy dissipation during a unit of time, no matter how the events' distribution is.
7. The network is dense enough so that when all the sensors are *Active*, the monitored area is fully covered. Besides, we assume that the graph representing the sensor network is connected (two sensors being connected when they are within the transmission range of each other).
8. Network lifetime is defined as the time separating the instant the network starts operating and the instant at which the network cannot be covered anymore because of the expiration of some nodes.
9. We assume ideal MAC layer conditions, i.e., perfect transmission of data on a node-to-node wireless link.
10. We assume that sensors have ideal sensing capabilities, i.e., that inside the sensing range, the quality of sensing does not depend on the distance from the sensor.

### 3.4 Problem modeling

Our problem consists of finding the optimal allocation of states to sensors, that maximizes network lifetime under the integrated constraint of coverage, clustering and routing. We call this problem **OPT-ALL-RCC**. To maximize network lifetime, we need a trade-off between total energy consumption and energy balancing among sensors. For example, to ensure area coverage, we would prefer to activate more sensors having higher residual energy (and consuming a higher total energy) than few sensors having little residual energy (consuming less total energy but more prone to expiration). On the other hand, any admissible solution of our model has to ensure full coverage of the monitored area and the existence of a spanning tree connecting

all CHs. To model this problem, we first define the following sets and constants :

- Let  $S = \{1..N\}$  be the set of sensors.
- Let  $C = \{1..M\}$  be the set of cells composing the monitored area.
- Let,  $\forall i = 1..|S|, \forall j = 1..|S|, \forall c = 1..|C|,$

$$d_{ij} = \begin{cases} 1 & \text{if sensor } i \text{ can reach sensor } j \text{ in one hop} \\ 0 & \text{else} \end{cases}$$

$$D_{ij} = \begin{cases} 1 & \text{if CH } i \text{ can reach CH } j \text{ in one hop} \\ 0 & \text{else} \end{cases}$$

$$\rho_{ic} = \begin{cases} 1 & \text{if sensor } i \text{ covers cell } c \\ 0 & \text{else} \end{cases}$$

Then, we define our binary decision variables  $\forall i = 1..|S|, \forall j = 1..|S| :$

$$X_i = \begin{cases} 1 & \text{if sensor } i \text{ is } \textit{Active} \\ 0 & \text{else,} \end{cases}$$

$$Y_i = \begin{cases} 1 & \text{if sensor } i \text{ is a CH} \\ 0 & \text{else,} \end{cases}$$

$$Z_{ij} = \begin{cases} 1 & \text{if sensor } i \text{ is connected to CH } j \\ 0 & \text{else.} \end{cases}$$

$$W_{ij} = \begin{cases} 1 & \text{if CH } i \text{ is connected to CH } j \text{ within a spanning tree} \\ 0 & \text{else.} \end{cases}$$



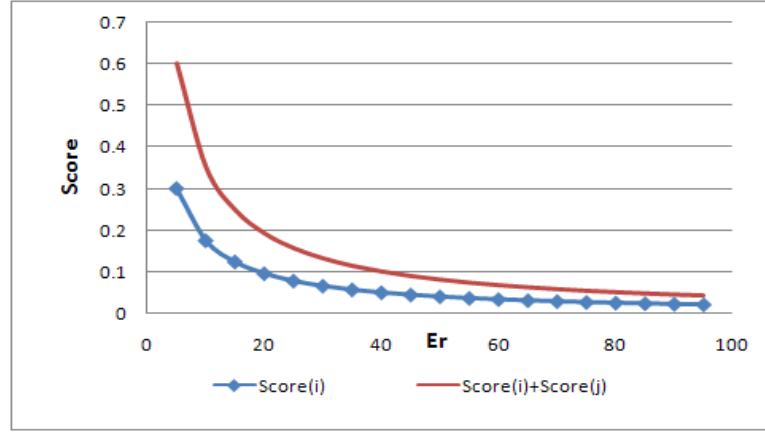


FIGURE 3.5 Shape of the score function

To balance energy consumption among nodes, we choose to minimize an objective function that is a linear combination of sensors scores. The score of a sensor  $i$  is defined by :

$$Score(i) = Log \left( 1 + \frac{Ed_i}{Er_i} \right)$$

where :

$$Ed_i = \begin{cases} E_{Active} & \text{if sensor } i \text{ is Active but not CH} \\ E_{CH} & \text{if sensor } i \text{ is CH} \\ E_{Sleep} \approx 0 & \text{else} \end{cases}$$

As shown in Figure 3.5, the logarithmic nature of this score function will tend to accentuate the importance of the residual energy when this latter is small and will give more importance to the dissipated energy when the residual energy is high enough. In other terms, when the residual energy is low, sensors will be selected essentially according to their residual energies, favoring the activation of sensors having relatively high residual energy and when the residual energy is relatively

high, the optimal solution will tend to activate as less sensors as possible. Thus, we can model our problem by the following optimization system :

**Minimize :**

$$\sum_{i=1}^{|S|} Y_i \cdot \text{Log} \left( 1 + \frac{E_{CH}}{Er_i} \right) + (X_i - Y_i) \cdot \text{Log} \left( 1 + \frac{E_{Active}}{Er_i} \right). \quad (3.1a)$$

**Subject to :**

$$\forall c = 1..|C|, \sum_{i=1}^{|S|} X_i \cdot \rho_{ic} \geq 1, \quad (3.1b)$$

$$\forall i = 1..|S|, Y_i \leq X_i, \quad (3.1c)$$

$$\sum_{i=1}^{|S|} Y_i \cdot d_{i0} \geq 1, \quad (3.1d)$$

$$\forall i = 1..|S|, \forall j = 1..|S|, j \neq i, Z_{ij} \leq X_i - Y_i, \quad (3.1e)$$

$$\forall i = 1..|S|, \forall j = 1..|S|, j \neq i, Z_{ij} \leq Y_j, \quad (3.1f)$$

$$\forall i = 1..|S|, \forall j = 1..|S|, j \neq i, Z_{ij} \leq d_{ij}, \quad (3.1g)$$

$$\forall i = 1..|S|, \sum_{\substack{j=1 \\ j \neq i}}^{|S|} Z_{ij} + Y_i = X_i, \quad (3.1h)$$

$$\forall j = 1..|S|, \sum_{\substack{i=1 \\ j \neq i}}^{|S|} Z_{ij} \leq N_{max}, \quad (3.1i)$$

$$\forall i = 1..|S|, \forall j = 1..|S|, j \neq i, W_{ij} \leq Y_i, \quad (3.1j)$$

$$\forall i = 1..|S|, \forall j = 1..|S|, j \neq i, W_{ij} \leq Y_j, \quad (3.1k)$$

$$\forall i = 1..|S|, \forall j = 1..|S|, j \neq i, W_{ij} \leq D_{ij}, \quad (3.1l)$$

$$\forall H \subseteq S, \forall s \in H, \sum_{j \in H} \sum_{\substack{k \in H \\ k > j}} W_{jk} \leq \left( \sum_{i \in H} Y_i \right) - Y_s, \quad (3.1m)$$

$$\sum_{j \in |S|} \sum_{\substack{k \in S \\ k > j}} W_{jk} = \sum_{j \in S} Y_j - 1, \quad (3.1n)$$

$$\mathbf{X}, \mathbf{Y} \in \{0, 1\}^{|S|}, \mathbf{Z}, \mathbf{W} \in \{0, 1\}^{|S|^2}. \quad (3.1o)$$

The expression of the objective function (3.1a) aims at balancing the energy consumption over the network. Equations (4.1d)-(3.1o) are the model constraints. Constraint (4.1d) guarantees a full coverage of the monitored area, such that every elementary cell is covered by at least one *Active* sensor. Let us note here that this coverage constraint is valid under the assumption of binary probability (disk) sensing model, where every cell is either covered by a sensor (if it is located inside its sensing range) or not covered at all (if it is outside that sensing range). With probabilistic coverage, the coverage constraint must ensure that every cell is covered up to a certain predefined coverage rate. Hence, the constraint (4.1d) should be replaced with :

$$\forall c = 1..|C|, \text{Prob}(c \text{ is covered}) \geq \text{cov\_rate}, \quad (3.2)$$

where *cov\_rate* is a predefined threshold coverage rate. Let  $P = \{P_{ic}\}$  be a coverage probability matrix known a priori, where  $P_{ic}$  = probability that sensor  $i$  covers cell  $c$ . Then, Equation (3.2) is equivalent to :

$$\forall c = 1..|C|, 1 - \prod_{i=1..|S|} (1 - X_i \cdot P_{ic}) \geq \text{cov\_rate}. \quad (3.3)$$

Once linearized, the new coverage constraint (3.3) will replace constraint (4.1d) in the above model to have it handle probabilistic coverage.

The rest of our constraints are as follows : constraint (3.1d) ensures that there exist at least a CH located one hop away from the PN. Constraints (3.1e) to (3.1h) ensure that every *Active* and non-CH sensor is connected to at least one CH within its

range. Constraint (3.1i) gives an upper bound on clusters' sizes. Equations (3.1j) to (3.1n) describe the routing constraint ensuring that the overlay network composed of CHs is connected and hence, there exist a tree-like partial subgraph. (3.1o) are the integrality constraints. To ensure that a spanning tree connecting all the CHs exists in any solution, constraints (3.1m) and (3.1n) require the enumeration of all the subsets of  $S$ . Even though these constraints represent the theoretical conditions to have a spanning tree in any graph (no cycles and a connected graph), they quickly result in a combinatorial explosion of the number of constraints due to the exponentially-increasing number of subsets of  $S$ . To circumvent this problem, we will proceed differently : we represent the routing constraint of our problem as a multi-flow routing problem. We consider that a virtual flow has to be routed from any CH to, at least, one CH which is one hop from the sink. Indeed, the optimal graph configuration that allows a flow to be routed between any pair of nodes of a connected graph where links have infinite capacity is the minimal-cost spanning tree. To model this virtual flow routing problem, we define a binary variable representing the use of the wireless link  $kl$  to convey a flow  $(i, j)$ , where  $i, j, k$  and  $l$  are CHs and  $i, j$  are respectively the source and destination of the flow. Let :

$$\begin{aligned} &\forall i, j, k, l \in \{1..|S|\}, i \neq j, k \neq l, \\ &V_{ij}^{kl} = \begin{cases} 1 & \text{if flow } (i, j) \text{ passes through the link } kl \\ 0 & \text{else.} \end{cases} \end{aligned} \tag{3.4}$$

The following constraints ensure that the network contains a spanning tree connecting all CHs :

$$\begin{aligned} & \forall i = 1..|S|, \forall j = 1..|S|, j \neq i, \forall k = 1..|S|, \\ & \sum_{\substack{l \in S \\ l \neq k}} V_{ij}^{kl} - \sum_{\substack{l \in S \\ l \neq k}} V_{ij}^{lk} = \begin{cases} 0 & \text{if } k \neq i \text{ and } k \neq j \\ Y_i \cdot Y_j & \text{if } k = i \\ -Y_i \cdot Y_j & \text{if } k = j \end{cases} \end{aligned} \quad (3.5a)$$

$$\forall i = 1..|S|, \forall j = 1..|S|, j \neq i, \forall k = 1..|S|, \forall l = 1..|S|, V_{ij}^{kl} \leq Y_i \quad (3.5b)$$

$$\forall i = 1..|S|, \forall j = 1..|S|, j \neq i, \forall k = 1..|S|, \forall l = 1..|S|, k \neq l, V_{ij}^{kl} \leq Y_j \quad (3.5c)$$

$$\forall i = 1..|S|, \forall j = 1..|S|, j \neq i, \forall k = 1..|S|, \forall l = 1..|S|, l \neq k, V_{ij}^{kl} \leq Y_k \quad (3.5d)$$

$$\forall i = 1..|S|, \forall j = 1..|S|, j \neq i, \forall k = 1..|S|, \forall l = 1..|S|, l \neq k, V_{ij}^{kl} \leq Y_l \quad (3.5e)$$

$$\forall i = 1..|S|, \forall j = 1..|S|, j \neq i, \forall k = 1..|S|, \forall l = 1..|S|, l \neq k, V_{ij}^{kl} \leq D_{kl} \quad (3.5f)$$

$$\mathbf{V} \in \{0, 1\}^{|S|^4}, \mathbf{Y} \in \{0, 1\}^{|S|}. \quad (3.5g)$$

Equation (3.5) is the flow constraint ensuring that a feasible path exists between any pair of CHs to convey an elementary unit of flow. Remaining constraints (3.5b) to (3.5e) limit the relevance of this virtual flow problem to the overlay network. Constraint (3.5f) ensures that CHs  $k$  and  $l$  are neighbors for a flow to pass on the link  $kl$ . Finally, (3.5g) are the integrality constraints.

In (3.5), we have a non-linear term that we need to linearize. For that, we define :

$$\forall i = 1..|S|, \forall j = 1..|S|, U_{ij} = Y_i \cdot Y_j \quad (3.6a)$$

To have a logical equivalence between  $U_{ij}$  and  $X_i.Y_j$ , we add the following constraints :

$$\forall i = 1..|S|, \forall j = 1..|S|, U_{ij} \leq Y_i, \quad (3.6b)$$

$$\forall i = 1..|S|, \forall j = 1..|S|, U_{ij} \leq Y_j, \quad (3.6c)$$

$$\forall i = 1..|S|, \forall j = 1..|S|, U_{ij} \geq Y_i + Y_j - 1. \quad (3.6d)$$

Taking into account the virtual-flow-related constraints (3.5) to (3.5f) and the linearized constraints (3.6a), (3.6b), (3.6c) and (3.6d), we end up with a linearized model of *OPT-ALL-RCC*.

**Proposition 3.4.1.** *OPT-ALL-RCC is NP-Complete.*

**Proof.** To prove the NP-Completeness of *OPT-ALL-RCC*, we will derive a polynomial reduction to the set covering problem which is known to be NP-Complete (Cormen *et al.*, 2001, pp. 1033-1038). For that, we propose to define an instance  $P$  of the set covering problem, build an instance  $I$  of *OPT-ALL-RCC* and then show that any algorithm that resolves  $I$  is able to resolve  $P$ .

Any instance  $P$  of the set covering problem is defined by a set of nodes, a set of node subsets and a cost for each node subset. Let  $S = \{s_1..s_N\}$  be a set of  $N$  sensors (nodes) and let  $J = \{S_1..S_N\}$  be a set of  $N$  node subsets ( $S_j \subseteq S$ ,  $j = 1..N$ ) such that  $\forall j = 1..N$ ,  $S_j$  is built from the empty set, as follows :

- The node  $s_j$  is inserted into  $S_j$  ;
- All nodes  $s_m \in S - \{s_j\}$  that are one-hop neighbors of  $s_j$  (i.e :  $d_{s_j s_m} = 1$ ) are added to  $S_j$ .

For each  $j = 1..N$ , we define the cost of subset  $S_j$  by :  $\alpha(j) = \text{Log}(1 + E_{CH}/Er_j)$ .

The instance  $P$  being defined, let us consider the instance  $I$  of *OPT-ALL-RCC*, with the following input :

- $N_{max} = \text{Infinity}$ .

- Each cell of the monitored area is covered by one and only one sensor.
  - The residual energy of nodes  $s_i \in S$  is equal to  $E_0 > E_{CH} > E_{Active}$ .
  - The transmission power of a CH is such that any potential CHs can reach each other, that is,  $\forall j, m \in \{1..N\}, j \neq m$ , if  $s_j$  and  $s_m$  are elected CH then  $D_{s_j s_m} = 1$ .
- Since each sensor covers a single cell, any admissible solution will have all its sensors turned on to satisfy the full coverage constraint. We clearly see that any algorithm that is able to resolve the above instance  $I$  of *OPT-ALL-RCC* can resolve any instance  $P$  of the set covering problem. Indeed, for every sensor  $s_k$ , ( $1 \leq k \leq N$ ) designated as cluster head in the solution of  $I$ ,  $S_k$  is a member of the minimum-cost covering set. Such a reduction proves that *OPT-ALL-RCC* is NP-Complete.

### 3.5 Proposed heuristic

As the considered problem is NP-Complete, we propose a Tabu search heuristic, called TABU-RCC, to tackle to exponentially-increasing processing time of the exact solution. TABU-RCC will be run by the PN to find a near-optimal sensor state configuration. As shown in Algorithm 3.1, TABU-RCC starts with an admissible solution and iteratively performs movements that consist of changing the state of one sensor at a time. The best solution found after the predefined number of iterations is transposed on sensors to form the new network configuration. The network will operate with this configuration for a predefined period  $T$  during which residual energies of active nodes and CHs will decrease, then TABU-RCC is run again to find a new configuration based on the new values of residual energies. This new configuration will be kept for another period  $T$  and so fourth. The periodic execution of TABU-RCC by the PN requires sensors-related information (e.g. residual energies) to be transmitted periodically to the PN (upstream communication) and the newly-computed sensor states to be transmitted to the sensors (downstream communication). In our architecture, sensor-related data will be collected exactly in the same manner as the sensed data,

---

**Algorithm 3.1** TABU-RCC : Tabu search algorithm of CH election under routing and coverage constraints

---

- **Initial solution** : the Tabu algorithm starts with a configuration where all sensors are activated as cluster heads. This configuration is obviously admissible ;
  - **Admissible configuration** : a configuration  $S$  is defined by the states of its sensors (*Sleep*, *Active* or *CH*). Only feasible configurations (i.e : satisfying model constraints (4.1d) to (3.1i)) are considered ;
  - **Score function** : a configuration is evaluated using the score function given by Equation (3.1a) ;
  - **Neighborhood investigation** : a search movement  $M < i, u, v >$  consists of changing the state (*Sleep*, *Active* or *CH*) of a single sensor  $i$  from state  $u$  to state  $v$ , such that the model constraints (4.1d) to (3.1i) are satisfied ;
  - **Aspiration criterion** : Tabu movements are allowed when the score of the resulting configuration is lower than the score of the best solution  $s^*$  found so far over the whole search process ;
  - **Stop criterion** : The search algorithm stops after a predefined number of iterations.
- 

i.e., using the cluster-based hierarchical structure of the network. *Active* sensors that have data to report will send it to the PN via their respective CHs. They will append the value of their respective residual energies to the data packets they are sending. When they have no data to send, they will synchronize their energy information with the PN periodically. As for the turned-off sensors, their energy consumption is constant and very low and can therefore be estimated by the PN whenever needed, as long as they are idle. As far as state assignment to sensors is concerned, the PN will broadcast a notification message holding associations between sensors and their newly-computed states (*Sensor ID*, *Sensor State*). Only the sensors which state has changed will have an entry in the notification message while all other sensors will keep their ongoing state. The notification message will be routed to the sensors via



their respective CHs.

### 3.6 Simulation results

In this paragraph, we will evaluate the quality of TABU-RCC with respect to the exact (optimal) solution which provides a lower bound of the objective function. To find the optimal solution, we implement our linear integer-variable model using ILOG CPLEX (ILOG, 2009). CPLEX is a mathematical programming and optimization tool that solves linear problems with continuous, integer or mixed variables, using the branch-and-bound method.

Then, we draw the variation of TABU-RCC's network lifetime when, respectively, the sensing range and the cluster size vary. These measures will show how worth it is to invest in sensors with higher sensing range or higher processing capacity.

Finally, we compare our Tabu search algorithm to EESH (Zhou *et al.*, 2007), a clustering algorithm recently proposed in the literature. EESH elects the set of cluster heads and ensures network connectivity. However, EESH does not handle optimal area coverage. To tackle this issue, we propose an enhanced version of EESH that computes an optimal area-covering subset of nodes (among the non-cluster head nodes) by resolving the linear system proposed by Chamam et Pierre (2007a). The remaining nodes are then switched off. Our choice of EESH as a comparison reference for our algorithm was motivated by the fact that EESH is one of the most recent clustering algorithms proposed in the literature and that EESH has been shown to outperform, in terms of network lifetime (which is our main objective to maximize), two of the most popular and well recognized clustering algorithms for WSN : LEACH (Heinzelman *et al.*, 2000) and HEED (Younis et Fahmy, 2004).

For all the simulations made, TABU-RCC was run with a tabu list size of 15 and a number of iterations of 10000.

### 3.6.1 Comparative performance evaluation : TABU-RCC with respect to its lower bound (CPLEX)

To evaluate the quality of our heuristic, we considered different network sizes of 9, 25, 49, 64, 100 and 225 sensors and compared the objective function provided by TABU-RCC to the lower bound provided by CPLEX. The measure of the objective functions has no practical interest but is only useful to assess the quality of the heuristic with respect to the exact solution (lower bound). It was difficult to run simulations on bigger network sizes because of the exponentially-increasing processing time of CPLEX, as shown on Table 3.1. In fact, we can clearly see that, while TABU-RCC always converges in few seconds, CPLEX's computation time grows exponentially, which deeply justifies the use of a heuristic. Let us note here that for the network sizes of 9, 25 and 49 sensors, the lower bound used is the value of the objective function given by the solution of the integer-variable system 4.3, while for network sizes of 64, 100 and 225 sensors, the resolution of the integer-variable system could not be made in a reasonable time, so we made a continuous relaxation of the system 4.3 and we used its solution to evaluate the solution provided by TABU-RCC. In fact, the relaxed system provides an obvious lower bound and it takes much less time to be resolved than the original integer-variable system. This explains why the lower

TABLEAU 3.1 Computation time of the lower bound (CPLEX) and TABU-RCC

Network size	Computation time	
	CPLEX	TABU-RCC
9	0.023	0.011
25	155	0.082
49	39987	0.21
64	362	2.33
100	873	6.47
225	1475	10.89

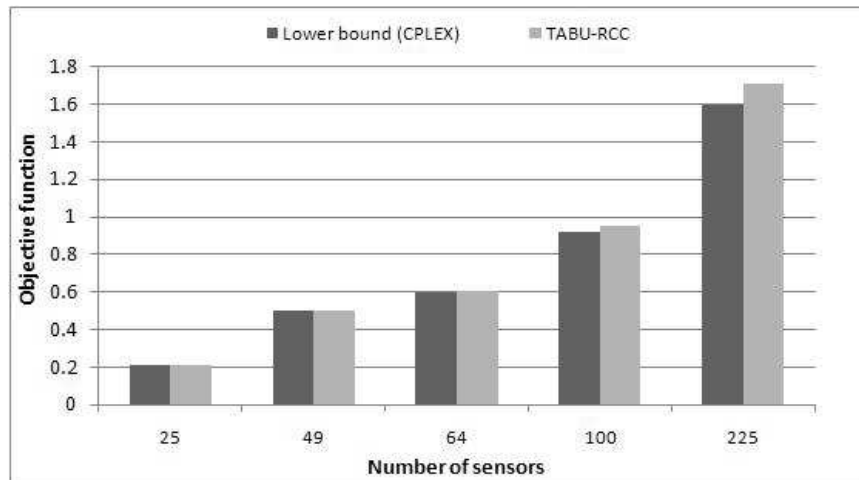


FIGURE 3.6 Value of the objective function generated by TABU-RCC vs CPLEX

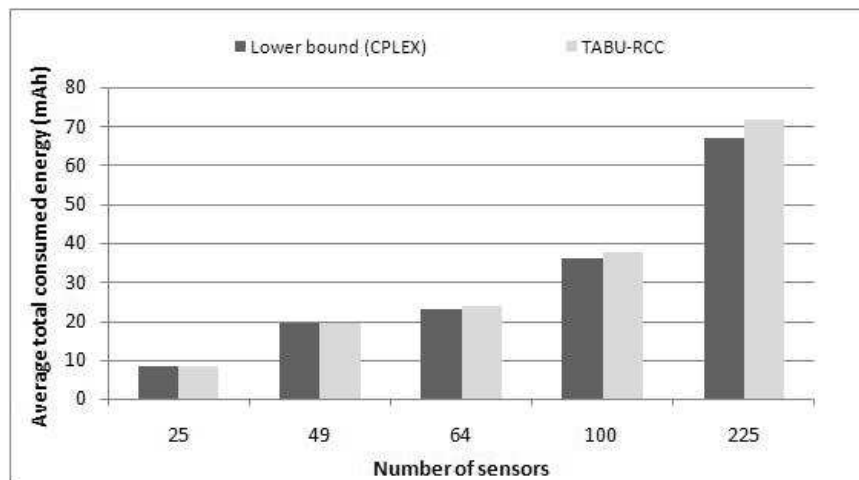


FIGURE 3.7 Average consumed energy using the solution generated by TABU-RCC vs CPLEX

bound processing time for 64 sensors reported on Table 3.1 seems lower than the processing time of the lower bound for 49 sensors.

As shown in Figure 3.6, the objective function of TABU-RCC's solutions is very close to that of the lower bound. Also, Figure 3.7 shows that the average energy consumed by TABU-RCC's solutions is very close to that of the lower bound. Thus, we can infer that TABU-RCC provides good solutions, close to their lower bound.

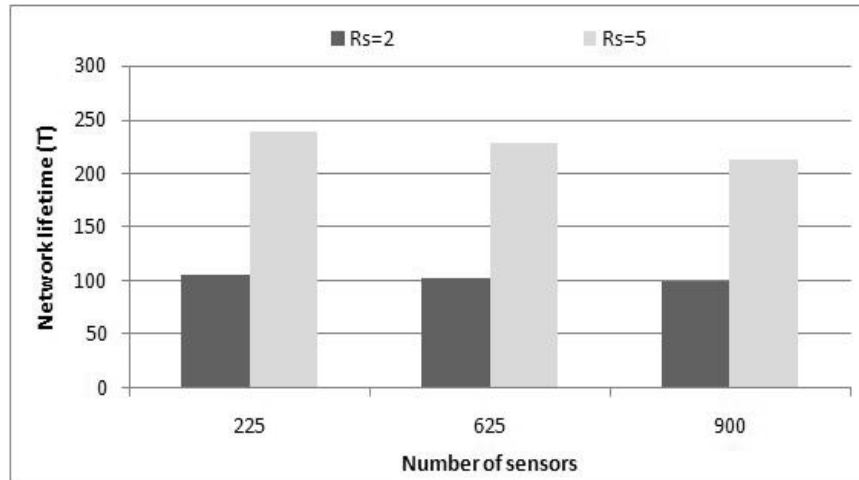


FIGURE 3.8 Variation of the network lifetime with the sensing range, using TABU-RCC

Even though it was predictable that an exact ILP resolution would not be suitable for large-sized networks, we computed the objective function of CPLEX's solutions for small and medium-sized networks only to assess the quality of the solutions provided by TABU-RCC with respect to optimality.

### 3.6.2 Performance evaluation of TABU-RCC : Impact of the sensing range

In practice, the value of the sensing range could vary depending on the physical properties of the sensors and on the type of signal they are sensing. To measure the impact of the sensing range, we considered three network sizes of respectively 225, 625 and 900 sensors. To measure the network lifetime, we run TABU-RCC iteratively and at every iteration, a sensor's energy is decremented by  $E_{CH}$  if it is designated as CH and by  $E_{Active}$  if it is assigned the state *Active*. Iterations stop when the monitored area cannot be covered anymore because of the energy exhaustion of some sensors. Figure 3.8 shows the variation of the network lifetime with the sensing range, for the considered networks. We see that, for all network sizes, when  $R_s$  increases, the

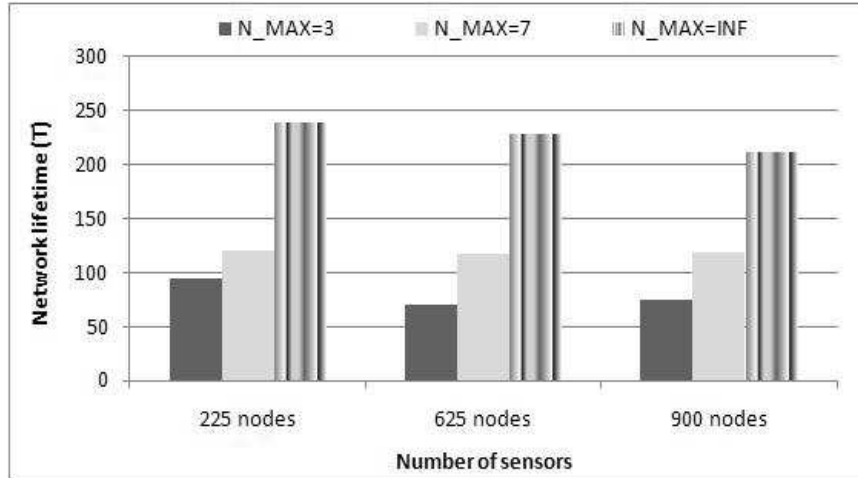


FIGURE 3.9 Variation of the network lifetime with the maximum cluster size, using TABU-RCC

network lifetime increases as well. This is due to the fact that, when  $R_s$  increases, less sensors are activated to ensure the full area coverage and the connectivity. Hence, less energy is dissipated, which, in average, increases network lifetime.

### 3.6.3 Performance evaluation of TABU-RCC : Impact of the maximum cluster size

In practice, a CH does not have an infinite capacity due to many physical factors like interference, limited CH's processing capabilities, collisions on the MAC layer, etc. To evaluate the impact of the maximum cluster size on the performance of our heuristic, we measured the network lifetime for different network sizes of respectively 225, 625 and 900 sensors. In Figure 3.9, we observe that, for all configurations, when  $N_{max}$  increases, the network lifetime increases as well. This is due to the fact that, when  $N_{max}$  increases, less CH are needed to ensure connectivity since a CH can serve more sensors. The leap-up in network lifetime becomes salient when the cluster size tends to infinity. In fact, when  $N_{max} \rightarrow \infty$ , one single CH can ensure the connectivity of all the sensors that can reach it.

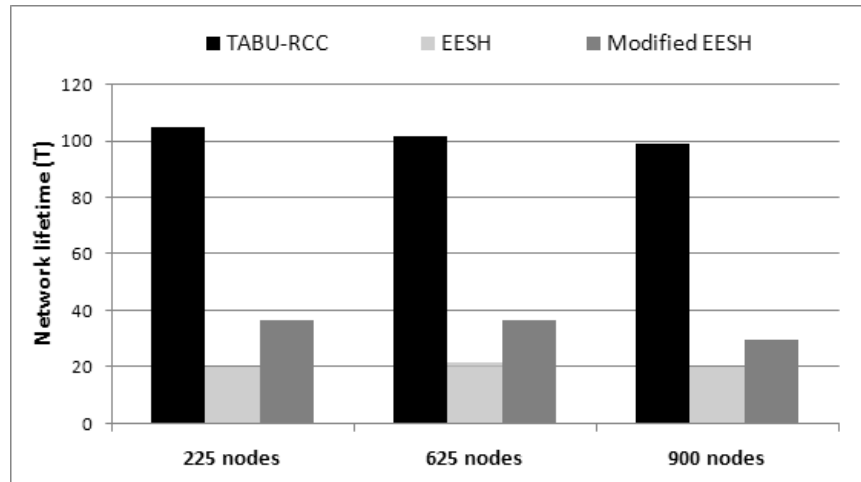


FIGURE 3.10 Comparison of TABU-RCC to EESH and the modified EESH in terms of network lifetime

#### 3.6.4 Comparative performance evaluation : TABU-RCC vs EESH

To evaluate our TABU-RCC in terms of network lifetime, we compared it to EESH (Zhou *et al.*, 2007). EESH functions as follows : nodes are promoted cluster heads according to their respective residual energies, their respective degrees, the distance to their neighbors and the residual energies of these neighbors. For that, EESH evaluates a cost function for every sensor in the network and iteratively elects the node having the greatest cost as CH. This process terminates when all non-CH sensors in the network are connected to at least one cluster head. As EESH does not consider optimal area coverage, we made a small modification of EESH that consists of computing, in each cluster provided by EESH, an optimal subset of non-CH sensors that will be activated while the other sensors are turned off. Figure 3.10 depicts the network lifetime provided respectively by TABU-RCC, EESH and the modified version of EESH, for different network sizes. We clearly see that TABU-RCC outperforms the modified version of EESH by providing three times longer network lifetime, for all network sizes. This considerable gain is one of the main benefits of our centralized heuristic.

### 3.7 Conclusion

In this paper, we proposed a novel centralized mechanism for near-optimal state assignment to sensors in large-scale cluster-based monitoring wireless sensor networks. Our mechanism is based on a tabu algorithm that computes a near-optimal network configuration in which each sensor can be activated, put in sleep mode or promoted as cluster head. Our mechanism maximizes network lifetime while ensuring the full coverage of the monitored area and the connectivity of the obtained configuration. Connectivity is fulfilled through an optimally-computed spanning tree connecting all the cluster heads. Simulations show that our mechanism provides for acceptable results with respect to the exact solutions of the derived ILP model, within low computation times. Despite its centralized aspect, our mechanism exhibits low complexity and low computation times making its practical implementation adaptable for large-scale networks.

As future research directions, we intend to develop a more sophisticated heuristic to improve the network lifetime. Furthermore, we intend to consider distance-dependent probabilistic event detection, where the probability that a sensor detects an event is function of the distance of that sensor from the event. Furthermore, we intend to work on distributed algorithms that address energy-efficient clustering under the joint coverage and routing constraint.

## CHAPITRE 4

# ENERGY-EFFICIENT CLUSTERING WITH TRANSMIT POWER ASSIGNMENT UNDER COVERAGE CONSTRAINT IN WIRELESS SENSOR NETWORKS

Ali Chamam and Samuel Pierre

Department of Computer & Software Engineering,

École Polytechnique de Montréal,

C.P 6079, succ. Centre-Ville, Montreal, H3C 3A7, Quebec, Canada.

*E-mail : {ali.chamam, samuel.pierre}@polymtl.ca*

### ABSTRACT

Because a sensor's energy is usually limited and hardly renewable, maximizing network lifetime is one of the fundamental issues in wireless sensor networks. In this paper, we address the problem of maximizing network lifetime in wireless sensor networks, under coverage and clustering constraints. We consider a dense network with overlapping sensing ranges, so that only a subset of sensors can be turned on to guarantee a full coverage of the monitored area. We also restrict our network to be clustered, e.g., every sensor is either a cluster head or connected to a neighboring cluster head. Besides, we consider sensors equipped with multiple transmit powers and thus capable of reaching a wider set of cluster heads with higher energy costs of higher transmit powers. We propose to find the optimal assignment of states (*Active, Sleep, Cluster Head*) and transmit powers to sensors, that maximizes network lifetime under the joint coverage and clustering constraint. We first model this pro-



blem as an Integer Linear Programming (ILP) optimization problem that we prove NP-complete. Then, we propose a Tabu search heuristic to tackle the exponential processing time of the exact solution. Our results show that in terms of network lifetime, our heuristic outperforms HEED, a well-recognized distributed clustering algorithm proposed in the literature.

**Key words :**

Wireless sensor networks, energy conservation, clustering; coverage, power assignment, linear programming, Tabu search.

#### 4.1 Introduction

Sensors are tiny MEMS (*Micro-Electro-Mechanical Systems*) capable of sensing, measuring and reporting physical variables related to their environment, such as temperature, pressure, motion, etc. In surveillance applications, sensors can be deployed in a monitored field to detect and report events such as object/human presence, movement, or intrusion (Chong et Kumar, 2003). As shown in Figure 5.1, wireless sensor networks (WSN) consist of a large number of limited-capability (power and processing) sensors and a special node equipped with higher energy and processing capabilities, called "processing node" (PN) or "*sink*". The PN collects data sent by sensors, filters it and aggregates it to extract useful information. The PN could be used as an interface between users and the sensor network. In the most general scenario, sensors may have variable sensing ranges, as depicted in Figure 4.2, and they can be equipped with several transmit powers so that they can reach a wider set of neighbors using a higher transmit power. However, due to their energy constraint, wireless sensors usually have a limited transmission range making multihop data routing toward the PN (Figure 5.1) more energy-efficient than direct transmission

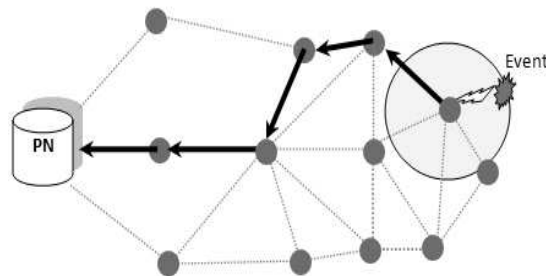


FIGURE 4.1 Multihop routing of collected data

(one-hop). Energy conservation in WSN is critical and has been addressed by substantial research (Cardei *et al.*, 2005; Raghunathan *et al.*, 2002). Generally, it is dealt with on five different levels (Cardei *et al.*, 2005; Chong et Kumar, 2003) :

1. efficient scheduling of sensor states to alternate between sleep and active modes ;
2. energy-efficient routing, clustering and data aggregation ;
3. efficient control of transmit power to ensure an optimal trade-off between energy consumption and connectivity ;
4. data compression (source coding) to reduce the amount of data transmitted ;
5. efficient channel access and packet retransmission protocols on the Data Link Layer.

The scope of this paper includes the first, the second and the third levels. We address the global problem of maximizing network lifetime under the joint coverage and clustering constraint, with optimal assignment of transmit powers. We consider a dense sensor network that is deployed in a certain area  $A$  to monitor some given events. When the network is dense, sensing ranges of neighbor sensors usually overlap, meaning that when an event occurs at a point  $P$  of  $A$ , all sensors that have  $P$  within their sensing range will detect and report the same event. This redundant

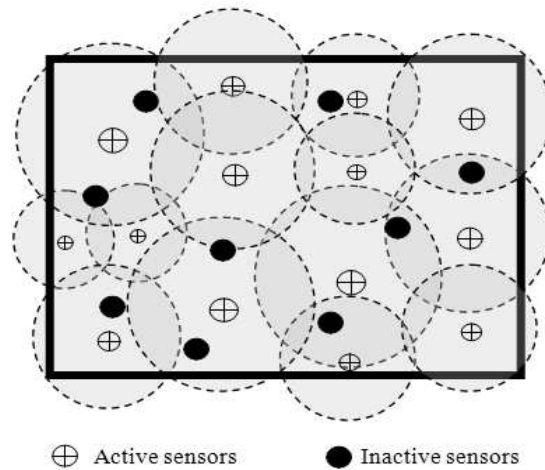


FIGURE 4.2 A full-covering sensor set with variable sensing ranges

transmission results in useless energy consumption. To save network energy and increase its lifetime, we propose to switch on only a subset of sensors that fully covers  $A$  while all other sensors are turned off. Figure 4.2 depicts an example of a full-covering sensor set.

On the other hand, clustering has been proven energy-efficient in WSN (Akkaya et Younis, 2005; Younis *et al.*, 2006). In cluster-based WSN, sensors are grouped in clusters, having each, one sensor promoted as cluster head (CH) and a set of cluster members connected to the CH. All non-CH nodes transmit their data to their CH, which routes it to the remote PN. In flat networks, sensors route data to the PN through their peer sensors, possibly using one of the many routing protocols proposed in the literature (Akkaya et Younis, 2005), whereas in cluster-based WSN, CHs transmit aggregated data collected from non-CH cluster members to the PN, either directly (one-hop) (Heinzelman *et al.*, 2000) or in multihop (Mhatre et Rosenberg, 2004b). Clustering provides substantial energy savings (Akkaya et Younis, 2005; Yan *et al.*, 2003) since only CH sensors are involved in routing and relaying data. Moreover, clustering alleviates bandwidth, enables its reuse, and can thus increase system capacity (Yao et Giannakis, 2005). Besides, the fact that only the CH transmits

information out of the cluster helps avoiding collisions between cluster-member sensors and helps avoiding the uncovered hole problem (Hwang *et al.*, 2006). However, since CHs consume more energy in aggregating and routing data, it is important to have an energy-efficient mechanism for CHs' election and rotation (Heinzelman *et al.*, 2000; Yao et Giannakis, 2005). When sensors are equipped with many transmit powers, they can reach a wider set of potential CHs by using higher transmit powers, at higher energy costs. When higher transmit powers are used, less CHs will be required to connect all non-CH sensors, therefore, less energy will be spent by the network in data processing and relaying (CH role) while more energy will be dissipated in transmission. Hence, for the sake of minimizing the average total energy consumption of the network, both optimal number and optimal placement of CHs have to be sought. Besides, since a substantial part of the total dissipated energy is spent in radio transmission (Raghunathan *et al.*, 2002), we address optimal assignment of transmit powers to sensors (Mhatre *et al.*, 2005; Sun *et al.*, 2008) in the context of clustering. Moreover, when the coverage constraint is addressed, it has to be considered along with the clustering constraint within the same global planning process. This is necessary to ensure the optimality of state and transmit power assignment to sensors. Indeed, when coverage and transmit power assignment are dealt with separately, the obtained configuration may not be optimal. For instance, an optimal covering subset of sensors could fail to guarantee network connectivity and an optimally clustered configuration may fail to guarantee coverage.

Numerous papers have addressed energy-efficient clustering (Arboleda et Nasser, 2006; Younis *et al.*, 2006) and area coverage (Cardei *et al.*, 2005; Carle et Simplot-Ryl, 2004; Inanc *et al.*, 2003; Li *et al.*, 2002) separately. Many others (Xing *et al.*, 2005; Yan *et al.*, 2003) addressed the integrated problem of maintaining area coverage and network connectivity but only considered flat networks and did not take advantage of the potential energy saving and ease of manageability of cluster-based networks (Chong et Kumar, 2003; Karl et Willig, 2005; Younis *et al.*, 2006). On the

other hand, many publications addressed the optimal assignment of transmit power to sensors (Mhatre *et al.*, 2005; Sun *et al.*, 2008). However, some of them addressed optimal transmit radii in a cluster-based architecture without addressing coverage (Liu et Lin, 2005; Sun *et al.*, 2008), and some others (Kubisch *et al.*, 2003; Panichpapiboon *et al.*, 2006) addressed optimal transmit radii from a pure information-theoretical aspect without address neither coverage, nor clustering. Finally, a narrow range of works addressed the optimal assignment of transmit power under coverage and clustering constraints but usually considered a predefined set of cluster heads, as in Mhatre *et al.* (2005), which limits the flexibility of the sought optimal configuration.

To the best of our knowledge, the problem of energy-efficient assignment of states (*Sleep*, *Active* or *ClusterHead*) and transmit powers to sensors under the integrated constraint of clustering and coverage has not been addressed within the same global optimization process. In this paper, we address the optimal planning of cluster-based WSN under the joint clustering and coverage constraint. In our architecture, any sensor can be *Active*, switched off (*Sleep*) or upraised CH, and each of these states is characterized by an average energy consumption per time unit. When *Active*, a sensor uses one of the transmit powers available to reach a neighboring CH and only CHs can process and route data. We seek an optimal assignment of states and transmit powers to sensors, that maximizes network lifetime, while ensuring simultaneously full area coverage and connectivity of every sensor to a CH.

This paper is organized as follows : in the next section, we present some related works. In Section 5.3, we outline our problem and enumerate our assumptions. In Section 5.4, we mathematically formulate our problem as an Integer Linear Programming (ILP) model that we prove NP-Complete. In Section 5.5, we present our resolution method based on a Tabu search heuristic. In Section 5.6, we discuss our simulation results. Finally, Section 5.7 concludes this paper and points out some future directions.

## 4.2 Related work

The problem of optimizing energy dissipation under area coverage and network connectivity constraints in WSN has been extensively addressed in the literature. Many protocols were proposed to alternate sensor states between *Active* and *Sleep* in order to maximize network lifetime. For example, Xing *et al.* (2005) provide a geometric analysis of the relationship between coverage and connectivity and propose the Coverage Configuration Protocol (CCP) protocol that dynamically configures the network to guarantee different degrees of coverage depending on the application requirements. In CCP, every node chooses its state (*Active* or *Sleep*) based on the coverage degree of the intersection points of its sensing circle with those of its neighbors. When coupled with any connectivity maintenance protocol, CCP offers connectivity and  $K$ -coverage. Lu *et al.* (2006) present the Scalable COverage Maintenance algorithm (SCOM), a localized coverage maintenance algorithm where sensors use the same redundancy eligibility rule as in Xing *et al.* (2005) to decide whether to turn on or off. SCOM implements, for each sensor, a back-off timer proportional to its residual energy. The back-off timer allows sensors with lower residual energies to decide about their states before sensors with higher residual energy. This makes them more likely to turn off than the sensors with higher residual energy, if they find themselves redundant. Chamam et Pierre (2007a) propose a centralized heuristic which dynamically calculates a near-optimal subset of sensors that guarantees a pre-defined coverage rate while ensuring network connectivity when the transmission range is greater than or equal to twice the sensing range. Yan *et al.* (2003) propose to schedule sensors' activities (*Active/Sleep*) so that every point in a grid monitored area is covered at any time. Neighbor sensors exchange a random reference time  $T_{ref}$  within cyclic rounds of constant duration  $T$  and decide to be active for a certain time duration within  $T$ . Even though the proposed schedule balances consumed energy over neighbor nodes, it does not take into account the residual energy of sensors

when calculating the activation time period of every node, which makes nodes with less residual energy more prone to expiration. All the works cited above consider flat networks and do not address clustering. Energy-efficient cluster formation and cluster-based routing for WSN have been widely addressed in literature (Akkaya et al., 2005). Cluster formation is typically based on the energy reserve of sensors and sensors' proximity to the cluster head (Arboleda et al., 2006; Li et al., 2002). For instance, in Low-Energy Adaptive Clustering Hierarchy (LEACH), one of the most popular hierarchical routing algorithms for WSN, Heinzelman et al. (2000) propose to form clusters of sensor nodes based on the received signal strength and to use local cluster heads as routers to the sink. This saves energy since the transmissions are only operated by cluster heads rather than all sensor nodes. Power-Efficient Gathering in Sensor Information Systems (PEGASIS) (Lindsey et al., 2002) and its variant Hierarchical-PEGASIS are two improvements of LEACH. Rather than forming multiple clusters, PEGASIS forms chains of sensor nodes so that each node transmits to and receives from a neighbor and only one node is selected from that chain to transmit to the PN. Hybrid Energy-Efficient Distributed Clustering (HEED) (Younis et al., 2004) is a distributed clustering protocol that uses a hybrid combination of the residual energy and the intra-cluster communication cost as attribute for CH selection. HEED ensures a uniform distribution of CHs across the network and adjusts the probability of CH-selection to ensure inter-CH connectivity. On the other hand, many works addressed the optimization of sensors' transmission radii to minimize energy dissipated in transmission. For example, Sun et al. (2008) consider a heterogeneous sensor network where a predefined subset of sensors with unlimited resources is assigned the role of CH while the other sensors are regular sensors. Communication between regular sensors and the closest CH can be one-hop or multihop. Authors propose an optimal transmission range for every set of sensors located at an equal distance from the closest CH. Sensors that are close to the CH will dissipate more energy than sensors that are farther from it because they will

serve as data relay toward the CH, so the gap in energy dissipation is compensated by adjusting the transmission radii of sensors located at equal distance from the CH. The disk having as center the CH and encompassing all the sensors connected to that CH is then divided into rings of variable widths. The optimal radius of each ring is derived so that the average dissipated energy is minimized. Even though in Sun *et al.* (2008), variable transmit powers are considered for clustering, the fact that the set of CHs is predefined will produce sub-optimal configurations. Moreover, coverage is not addressed. Liu et Lin (2005) propose a distributed protocol to dynamically adjust sensors' transmission ranges to control clusters' sizes in an energy-constrained sensor network, regardless to the nodes distribution. The basic idea of the proposed protocol is that every node adjusts its transmit power based on the number of its neighbors. The proposed protocol elects a set of cluster heads based on the residual energies. Even though the proposed protocol addresses variable transmission radii and shows good performance in terms of network lifetime and average dissipated energy, it does not address area coverage. Moreover, the choice of the cluster degree, which determines the transmission radius of each node, does not derive from an optimization process and thus, the clustered configurations are sub-optimal. Kubisch *et al.* (2003) present two distributed algorithms for transmit power control. The first one is based on a threshold on the number of neighbors and the second one is based on a threshold on the mean number of neighbors. The proposed algorithms are localized : every node decides to increase its transmit power by a certain factor if the computed metric is below the threshold and to decrease it if the computed metric is above the threshold. Even though the proposed algorithms insure network connectivity, they do not address neither area coverage, nor clustering. Furthermore, they are not optimal because of the random choice of the threshold value. Panichpapiboon *et al.* (2006) evaluate the optimal transmit power in the case of both regular topology (analytically) and random topology (via simulations). They also investigate the interrelation between optimal transmit power, data rate, and node spatial density. Besides, they



investigate 1) the impact of different propagation path loss exponents (on different links of a multihop route) on the performance of a common transmit power control scheme and 2) the interrelation between transmit power, connectivity, and network longevity. Furthermore, they provide a rigorous analysis of exact interference power and Bit Error Rate (BER) using a detection theory approach. In particular, network connectivity is viewed from a communication-theoretic angle : a network is said to be connected if any source node can communicate with a BER lower than a prescribed value  $BER_{th}$  to a destination node placed at the end of a multihop route with an average number of hops. Although this paper presents a very interesting communication-theoretic aspect of connectivity, it does not address neither area coverage nor clustering. Mhatre *et al.* (2005) consider a heterogeneous monitoring sensor network where two types of nodes, cluster heads and regular sensors, are distributed over an area using two-dimensional homogeneous Poisson processes of average densities  $\lambda_1$  and  $\lambda_0$ , respectively. Regular sensors use multihopping to communicate with their closest CHs. Authors investigate the optimum node intensities ( $\lambda_0$  and  $\lambda_1$ ) and node energies ( $E_0, E_1$ ) that guarantee a lifetime of at least  $T$  units, while ensuring connectivity (of every regular node to a CH) and coverage of the monitored area with a high probability. Although this work is very interesting in the sense that it addresses both coverage and clustering in the same optimization problem, it has some limitations. In fact, even if the density of cluster heads is optimized, their number and positions are still limited to the nodes that have the capabilities (energy and processing) to be CHs. That limits the flexibility of the configuration and thus may limit the total average lifetime of the network.

### 4.3 Problem Statement and assumptions

In this paper, we consider a homogeneous WSN deployed in an area  $A$  to monitor certain critical activities or events. As shown in Raghunathan *et al.* (2002) for the

case of Rockwell's WINS seismic sensors, a sensor's radio can be in one of the following four activity modes, characterized by their respective power consumptions : *Transmit* (0.38 - 0.7 W), *Receive* (0.36 W), *Idle* (0.34 W) and *Sleep* (0.03 W). We note that when sensors are transmitting, receiving or idle, they have roughly the same energy consumption and can then be associated with the same state, *Active*, in which the sensor's radio is switched on. When the sensor is in *Sleep* mode, it does not perform any sensing and its radio is turned off. When sensors are equipped with many transmit powers  $\{P_1..P_K\}$ , the dissipated energy depends on the level of transmit power used. We denote by  $Active(P_i), \forall i = 1..K$ , the state in which a sensor is *Active* and using transmission power level  $P_i$ . Moreover, we consider a cluster-based topology in which any sensor can be designated CH and every *Active* sensor has to be either a CH or connected to a CH. CHs consume more energy than regular sensors because of their energy-intensive tasks (data collection, aggregation, filtering and routing). As sensors are equipped with multiple transmit powers, they can reach further CHs using higher transmit powers, but at a higher energy cost. Less CHs would then be required to cluster the network (and thus less energy will be spent by the CHs) but more energy would be dissipated by *Active* sensors in transmission, as shown in Figure 4.3. Hence, a trade-off between the number of CHs and the transmit powers used by non-CH sensors has to be sought. Furthermore, as sensors' residual energies are dynamic, optimal configurations (sensor states and assigned transmit powers) have to be recalculated periodically, because if CHs were chosen a priori and fixed throughout the system's lifetime, they would quickly exhaust all their energy and become non-operational. In this paper, we propose to dynamically designate the set of CHs according to an objective function that reflects their residual energies and their relative position with respect to their neighboring non-CH *Active* nodes. Without loss of generality, we consider that each sensor is equipped with  $K$  transmit power levels  $P_1..P_K$  ( $P_1 < P_2 < \dots < P_K$ ) that it can use when it is *Active* and non-CH, and an additional transmit power  $P_{CH}$  used by CH sensors for inter-cluster communication.

Every sensor can be in one of the following states : *Sleep*,  $Active(P_i)$  and *Cluster Head (CH)*, having respectively energy consumptions  $E_{Sleep}$ ,  $E_{Active}(P_k)$  and  $E_{CH}$  per time unit, where  $E_{Active}(P_k)$  is the energy dissipated, during a time unit, by a sensor using transmission power  $P_k$ ,  $k = 1..K$ , and  $\forall k = 1..K$ ,  $E_{Sleep} \ll E_{Active}(P_k) < E_{CH}$ .

On the other hand, we assume that all sensors have the same sensing range  $R_s$  and that their detection model follows a binary probability function, called *Disk model* (Cardei *et al.*, 2005; Xing *et al.*, 2005), in which an *Active* sensor  $i$  detects any event that occurs within its sensing range  $R_s$  with the same probability  $Prob(detection) = 1$ , whatever the distance from the event is. And, any event occurring outside  $R_s$  is not detected by  $i$ . Note that this binary probability sensing model is a simplifying representation of a sensor's detection probability and it does not necessarily reflect the received signal loss model on a real wireless channel. In fact, on a real wireless channel, because of signal fading and interference, events occurring far away from a sensor are less likely to be detected than events occurring close to it. In other words, an event occurring outside a sensor's range is still detected with a certain probability which decreases when the distance to the sensor increases. In the literature, many have addressed probabilistic coverage in wireless sensor networks (Ahmed *et al.*, 2005; Hefeeda et Ahmadi, 2007) and proposed coverage protocols based on this model. However, to simplify the representation of the sensing model, most work in the literature (Cardei *et al.*, 2005; Xing *et al.*, 2005; Yen *et al.*, 2007) make the disk model assumption. In this work, we make this assumption too.

We define a set of sensors  $S_c$  as a covering set of area  $A$  if and only if :  $\forall$  point  $P \in A$ ,  $\exists i \in S_c$  such as  $i$  covers  $P$ . As stated earlier, when sensors' sensing ranges overlap, we propose to switch on only a subset of sensors that fully covers the monitored area, while the other sensors are switched off (*Sleep*) to save their energy. In critical surveillance applications, it is important to guarantee that the monitored area is fully covered by sensors at every instant of the network lifetime. Hence, our

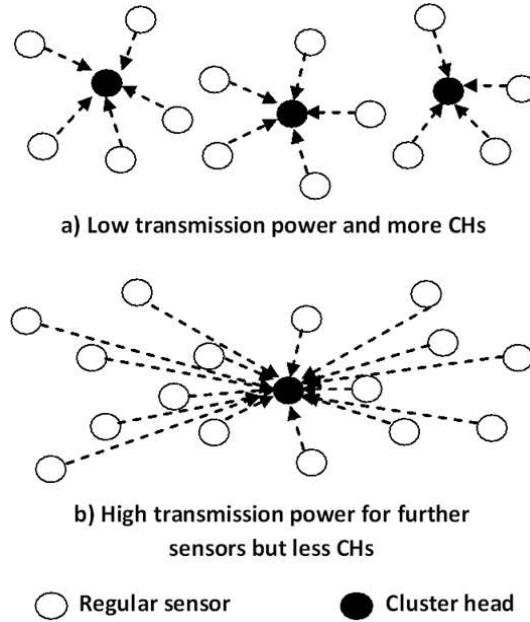


FIGURE 4.3 Two limit scenarios : low transmit power and many CHs required vs high transmit power and less CHs required

problem combines optimal coverage and clustering, which are strongly tied because optimal clustering depends on which sensors are switched on. Besides, performing optimal coverage and optimal clustering in two separate stages does not guarantee an optimal configuration because, for example, some of the cluster heads elected during the optimal clustering stage could belong to the set of switched-off sensors. Hence, coverage and clustering have to be handled within the same optimization process and any admissible network configuration (admissible solution) must contain :

- a full-covering set of active sensors ;
- a set of CHs so that every *Active* sensor is either a CH or connected to its closest CH, using one of the transmit powers it is equipped with.

Our objective is to find the network-lifetime-optimal allocation of states (*Sleep*, *Active*( $P_k$ ), *CH*) that meets these two conditions. Before modeling our problem, we make the following assumptions :

1. Each sensor has a unique ID, known to the PN and to the sensor itself.

2. The position of each sensor is fixed and known to the PN. The location information can be obtained either through a GPS (Global Positioning System), as assumed in Cardei *et al.* (2005) (but this technique is still expensive due to the high cost of placing a GPS on each sensor), or using one of the many GPS-free localization techniques proposed in the literature (Chu et Jan, 2007). However, in this paper, we do not specifically address any localization technique and we assume that, whatever the localization technique used, sensors' location information is available at the PN.
3. Active sensors capture events occurring in their sensing range and transmit data associated with these events straightaway, without any buffering, because sensors are usually not equipped with large (and costly) buffers.
4. All sensors have the same sensing range  $R_s$ . Every sensor in  $Active(P_k)$  state has a transmission range  $R_t(P_k)$  that depends on the transmit power  $P_k$  used by the sensor.
5. All CHs have the same transmission range  $R_t^{CH}$  and  $\forall k = 1..K, R_t^{CH} > R_t(P_k)$ ;
6. Each sensor has an initial energy  $E_0$ . The PN has no energy limitation. Besides, we assume that, when a sensor is  $Active(P_k)$ , it dissipates a constant amount of energy  $E_{Active}(P_k)$  during a unit of time, no matter how the events' distribution is.
7. The network is dense enough so that when all the sensors are  $Active$ , the monitored area is fully covered. Besides, we assume that, when every sensor uses its highest transmit power, the graph representing the sensor network is connected (two sensors being connected when they are within the transmission range of each other).
8. Network lifetime is defined as the time separating the instant the network starts operating and the instant at which the network cannot be covered anymore because of the expiration of some nodes.

9. We assume ideal MAC layer conditions, i.e., perfect transmission of data on a node-to-node wireless link.
10. We assume that sensors have ideal sensing capabilities, i.e., that inside the sensing range, the quality of sensing does not depend on the distance from the sensor.

#### 4.4 Problem modeling

Our problem consists in finding the optimal assignment of states to sensors (*Sleep*, *Active*( $P_k$ ), *CH*) to sensors, that maximizes network lifetime under the joint constraint of coverage and clustering. We call this problem *OPTimal Clustering with Power Assignment under Coverage Constraint (OPT-CPACC)*. To model OPT-CPACC, we first define the following sets and variables :

Let  $X_i$  be a binary variable defined by :  $\forall i = 1..N$ ,

$$X_i^k = \begin{cases} 1 & \text{if sensor } i \text{ is in } Active(P_k) \text{ state} \\ 0 & \text{if sensor } i \text{ is in } Sleep \text{ state} \end{cases}$$

Let  $Y_i$  be a binary variable defined by :  $\forall i = 1..N$ ,

$$Y_i = \begin{cases} 1 & \text{if sensor } i \text{ is CH} \\ 0 & \text{else} \end{cases}$$

Let  $D = \{d_{ij}^k\}$  be a three-dimensional matrix defined by :  $\forall i = 1..N, j = 1..N, k = 1..K$ ,

$$d_{ij}^k = \begin{cases} 1 & \text{if sensor } i \text{ can reach sensor } j \text{ using transmit power } P_k \\ 0 & \text{else} \end{cases}$$

To model the area coverage constraint, we subdivide our monitored area into elementary square cells and we define the binary matrix  $\rho = \{\rho_{ij}\}$  by :  $\forall i = 1..N, j = 1..M$ ,

$$\rho_{i,j} = \begin{cases} 1 & \text{if sensor } i \text{ covers cell } j \\ 0 & \text{else} \end{cases}$$

Finally, let  $\lambda$  be a positive integer.

The input matrices  $D$  and  $\rho$  are supposed to be known.

To find the optimal subset of sensors that will be activated as well as the optimal set of CHs and the optimal assignment of sensors to CHs, we have to find the solution  $(X_i^k, Y_i)$ ,  $\forall i = 1..N, \forall k = 1..K$ , of the following system :

**Minimize :**

$$\sum_{i=1}^N \sum_{k=1}^K \left[ X_i^k \cdot (1 - Y_i) \cdot \left( \frac{E_{Active}(P_k)}{Er_i} \right)^\lambda + X_i^k \cdot Y_i \cdot \left( \frac{E_{CH}}{Er_i} \right)^\lambda \right] \quad (4.1a)$$

**Subject to :**

$$\forall i = 1..N, \sum_{k=1}^K X_i^k \leq 1, \quad (4.1b)$$

$$\forall i = 1..N, \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{k=1}^K X_i^k \cdot d_{ij}^k \cdot Y_j + Y_i \geq \sum_{k=1}^K X_i^k, \quad (4.1c)$$

$$\forall j = 1..M, \sum_{i=1}^N \sum_{k=1}^K X_i^k \cdot \rho_{ij} \geq 1, \quad (4.1d)$$

$$\forall i = 1..N, X_i^K \geq Y_i, \quad (4.1e)$$

$$\forall i = 1..N, \forall k = 1..K, X_i^k \in \{0, 1\}, \quad (4.1f)$$

$$\forall i = 1..N, Y_i \in \{0, 1\}. \quad (4.1g)$$

The exponential character of the weighting factors  $\left(\frac{E_{Active}(P_k)}{Er_i}\right)^\lambda$  and  $\left(\frac{E_{CH}}{Er_i}\right)^\lambda$  in the objective function (4.1a) aims at balancing the dissipated energy and thus prolonging network lifetime. In fact, when  $\lambda \geq 1$ , these weighting terms will tends to penalize the choice of sensors having low residual energy when the latter is small, favoring the activation of more sensors with relatively high residual energy than fewer sensors with low residual energy. In other words, the lower their residual energy, the less likely sensors will be selected in the optimal solution. And, the greater  $\lambda$ , the more balanced the energy consumption. Constraint (4.1b) ensures every *Active* sensor uses one transmit power at a time. Constraint (4.1c) is the clustering constraint and constraint (4.1d) is the coverage one. Constraint (4.1e) ensures any CH is *Active*. Finally, (4.1f) and (4.1g) are the integrality constraints.

To resolve the above problem (4.1) using an integer-variable optimization tool, we need to linearize the non-linear terms in constraint (4.1c) and in the objective function (4.1a). For that, we first note that, because of constraint (4.1e), the integer expression  $X_i.(1 - Y_i)$  in (4.1a) is equivalent to  $(X_i - Y_i)$  and the objective function can be linearized accordingly. Then, to linearize the constraint (4.1c), we define the binary variable  $W_{ij}^k = X_i^k.Y_j$  and we add the following constraints :

$$\forall i = 1..N, \forall j = 1..N, \forall k = 1..K, W_{ij}^k \leq X_i^k, \quad (4.2a)$$

$$\forall i = 1..N, \forall j = 1..N, \forall k = 1..K, W_{ij}^k \leq Y_j, \quad (4.2b)$$

$$\forall i = 1..N, \forall j = 1..N, \forall k = 1..K, W_{ij}^k \geq X_i^k + Y_j - 1, \quad (4.2c)$$

$$\forall i = 1..N, \forall j = 1..N, \forall k = 1..K, W_{ij}^k \in \{0, 1\}. \quad (4.2d)$$

We end up with the following binary-variable linear system :



**Minimize :**

$$\sum_{i=1}^N \sum_{k=1}^K \left[ (X_i^k - Y_i) \cdot \left( \frac{E_{Active}(P_k)}{Er_i} \right)^\lambda + Y_i \cdot \left( \frac{E_{CH}}{Er_i} \right)^\lambda \right] \quad (4.3a)$$

**Subject to :**

$$\forall i = 1..N, \sum_{k=1}^K X_i^k \leq 1, \quad (4.3b)$$

$$\forall i = 1..N, \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{k=1}^K W_{ij}^k \cdot d_{ij}^k + Y_i \geq \sum_{k=1}^K X_i^k, \quad (4.3c)$$

$$\forall j = 1..M, \sum_{i=1}^N \sum_{k=1}^K X_i^k \cdot \rho_{ij} \geq 1, \quad (4.3d)$$

$$\forall i = 1..N, X_i^K \geq Y_i, \quad (4.3e)$$

$$\forall i = 1..N, \forall j = 1..N, \forall k = 1..K, W_{ij}^k \leq X_i^k, \quad (4.3f)$$

$$\forall i = 1..N, \forall j = 1..N, \forall k = 1..K, W_{ij}^k \leq Y_j, \quad (4.3g)$$

$$\forall i = 1..N, \forall j = 1..N, \forall k = 1..K, W_{ij}^k \geq X_i^k + Y_j - 1, \quad (4.3h)$$

$$\forall i = 1..N, \forall k = 1..K, X_i^k \in \{0, 1\}, \quad (4.3i)$$

$$\forall i = 1..N, Y_i \in \{0, 1\}, \quad (4.3j)$$

$$\forall i = 1..N, \forall j = 1..N, \forall k = 1..K, W_{ij}^k \in \{0, 1\}. \quad (4.3k)$$

**Proposition 4.4.1.** *OPT-CPACC is NP-Complete*

**Proof.** To prove the NP-Completeness of *OPT-CPACC*, we will derive a polynomial reduction to the set covering problem, known to be NP-Complete (Cormen *et al.*, 2001). For that, we propose to define an instance *SCP* of the set covering

problem, build an instance  $I$  of *OPT-CPACC* and then show that any algorithm that resolves  $I$  resolves *SCP* as well.

Any instance *SCP* of the set covering problem is defined by a set of nodes, a set of node subsets and a cost for each node subset. Let :

- $S = \{s_1..s_N\}$  be a set of  $N$  nodes (sensors) ;
- $J = \{S_1..S_M\}$  be a set of  $M$  node subsets defined by :  $\forall j = 1..M, S_j = \{s_j\} \cup \{s_m \in S - \{s_j\} \text{ such that } s_m \text{ can reach } s_j \text{ using its lowest transmit power } P_1\}$  ;
- $\forall j = 1..M$ , the cost of subset  $S_j$  is defined by :  $Cost(S_j) = \left(\frac{E_{CH}}{E_{r_j}}\right)^\lambda$ .

The instance *SCP* being defined, let us consider the instance  $I$  of *OPT-CPACC*, defined by the following input :

- Each cell of the monitored area is covered by exactly one sensor ;
- The residual energy of nodes  $s_i \in S, \forall i = 1..N$  is equal to  $E_0$  and  $E_0 > E_{CH} > E_{Active}(P_k), \forall k = 1..K$  ;
- Each sensor is equipped with exactly one transmit power  $P_1$ .

Since each sensor covers exactly one single cell, any admissible solution of the instance  $I$  will have all its sensors turned on, to satisfy the full coverage constraint. If we could resolve  $I$ , then the solution of *SCP* can be derived from the solution of  $I$  as follows : for every sensor  $s_i, (1 \leq i \leq N)$  designated as CH in the solution of  $I, S_i$  is a member of the minimum-cost covering set. In fact, since the solution of  $I$  is fully clustered and since sensors use their lowest power  $P_1$  to connect to their CH, the optimal solution of  $I$  exhibits the lowest-cost set of CHs. And, since the cost of subset  $S_j, \forall j = 1..M$ , is exactly the weight of cluster head  $s_j$  in the optimal solution of  $I$ , the set  $\{S_j, \text{ such that } s_j \text{ is designated CH in the optimal solution of } I\}$  is the lowest-cost set of node subsets that fully covers all network nodes. Hence, any algorithm that resolves the above instance  $I$  of *OPT-CPACC* can resolve any instance *SCP* of the set covering problem. Such reduction proves that *OPT-CPACC* is NP-Complete.

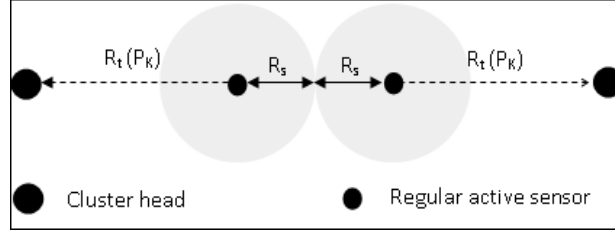


FIGURE 4.4 Necessary condition to have a connected overlay CH graph :  $R_t^{CH} \geq 2.(R_s + R_t(P_K))$

**Proposition 4.4.2.** *When the inter-cluster transmission range  $R_t^{CH}$  is greater or equal to  $2.(R_s + R_t(P_K))$ , any admissible solution exhibits a clustered configuration where the overlay graph composed of cluster heads is connected.*

**Proof.** Since the whole network is assumed to be covered and clustered, every node is either a CH or attached to a CH. Figure 4.4 depicts the admissible configuration exhibiting the longest distance between two adjacent CHs. In fact, when sensors are lined-up, each regular sensor has to be connected to a CH. Therefore, the maximum distance separating a regular *Active* sensor from an adjacent CH is  $R_t(P_K)$ . On the other hand, the monitored area is supposed to be fully covered, so the maximum distance separating two adjacent regular *Active* sensors is  $2.R_s$ . Hence, the maximum distance separating two CHs is  $2.R_s + 2.R_t(P_K)$ . In this worst-case configuration, it is clear that when  $R_t^{CH} \geq 2.(R_t(P_K) + R_s)$ , the overlay graph composed of CHs will be connected. This condition is very interesting because it will guarantee that the overlay graph formed by the CHs is connected and thus could be used as a routing topology, to convey data from any sensor toward the PN. However, this condition is a very conservative because, in practice, when the network is dense enough, the average distance between two adjacent CHs is likely to be less than  $2.(R_s + R_t(P_K))$ .

## 4.5 Proposed heuristic

As the considered problem is NP-Complete, we propose a Tabu search heuristic, called *Tabu search heuristic for Clustering and Power Assignment under Coverage Constraint* (T-CPACC), to tackle to exponentially-increasing processing time of the exact solution. T-CPACC will be run by the PN to find a near-optimal sensor state configuration. As shown in Algorithm 4.1, T-CPACC starts with an admissible solution and iteratively performs movements that consist in changing the state of one sensor at a time. The best solution found after the predefined number of iterations is transposed on sensors to form the new network configuration. The network will ope-

---

**Algorithm 4.1** T-CPACC : Tabu search algorithm for clustering and power assignment in WSN under coverage constraint

---

- **Initial solution** : the Tabu algorithm starts with a configuration where all sensors are activated as CHs. This configuration is obviously admissible ;
  - **Admissible configuration** : a configuration  $S$  is defined by the states of its sensors (*Sleep*,  $Active(P_k)$ , *CH*) where  $Active(P_k)$  is a state in which a sensor is *Active* and is using transmit power  $P_k$ ,  $k = 1..K$ . Only feasible configurations (i.e, satisfying model constraints (4.1b-4.1d) are considered ;
  - **Score function** : a configuration is evaluated using the score function of Equation (4.1a) ;
  - **Neighborhood investigation** : a search movement  $M < i, u, v >$  consists in changing the state of a single sensor  $i$  from state  $u$  to state  $v$ , such that the model constraints (4.1b-4.1d) are satisfied ;
  - **Aspiration criterion** : Tabu movements are allowed when the score of the resulting configuration is lower than the score of the best solution found so far over the whole search process ;
  - **Stop criterion** : The search algorithm stops after a predefined number of iterations.
-

rate with this configuration for a predefined period  $T$  during which residual energies of *Active* nodes and CHs will decrease, then T-CPACC is run again to find a new configuration based on the new values of residual energies. This new configuration will be kept for another period  $T$  and so fourth. The periodic execution of T-CPACC by the PN requires sensors-related information (e.g, residual energies) to be transmitted periodically to the PN (upstream communication) and the newly-computed sensor states to be transmitted to the sensors (downstream communication). In our architecture, sensor-related data will be conveyed exactly in the same manner as the sensed data, i.e, through the cluster-based hierarchical structure of the network. *Active* sensors that have data to report will send it to the PN via their respective CHs. They will append the value of their respective residual energies to the data packets they are sending. When they have no data to send, they will synchronize their energy information with the PN periodically. As for the switched-off sensors, their energy consumption is constant and very low and can therefore be estimated by the PN whenever needed, as long as they are in *Sleep* state. As far as state assignment to sensors is concerned, the PN will broadcast a notification message holding associations between sensors and their newly-computed states (*Sensor ID*, *Sensor State*). Only the sensors which state has changed will have an entry in the notification message while all other sensors will keep their ongoing state. The notification message will be routed to the sensors via their respective CHs.

#### 4.6 Implementation and simulation results

In this paragraph, we assess the efficiency of T-CPACC. We first evaluate the quality of T-CPACC's solutions with respect to a lower bound provided by the exact resolution of the integer-variable model (4.3). Then, we compare, in terms of network lifetime and energy consumption balancing, T-CPACC to HEED (Younis et Fahmy, 2004), one of the most popular algorithms in literature for energy-efficient

clustering with transmit powers assignment.

#### 4.6.1 Quality of the solution : T-CPACC vs its lower bound

To evaluate the quality of our heuristic, we compared the objective function provided by T-CPACC to the lower bound derived from the exact resolution of the mathematical model (4.3). Let's note though that the measure of the objective function has no practical usefulness but is only meant to assess the quality of the heuristic's solutions with respect to the lower bound (i.e, how far is the solution from optimality). To compute the optimal solution, we implemented our linear integer-variable model using ILOG CPLEX (ILOG, 2009). CPLEX is a mathematical programming and optimization software that solves linear problems with continuous, integer or mixed variables, using the branch-and-bound method. Figure 4.5 depicts the objective function of the solutions generated by CPLEX and by T-CPACC, for different network sizes. Let's note here that, because of the exponentially-increasing processing time of CPLEX, we were not able to solve the integer-variable system (4.3) in a reasonable

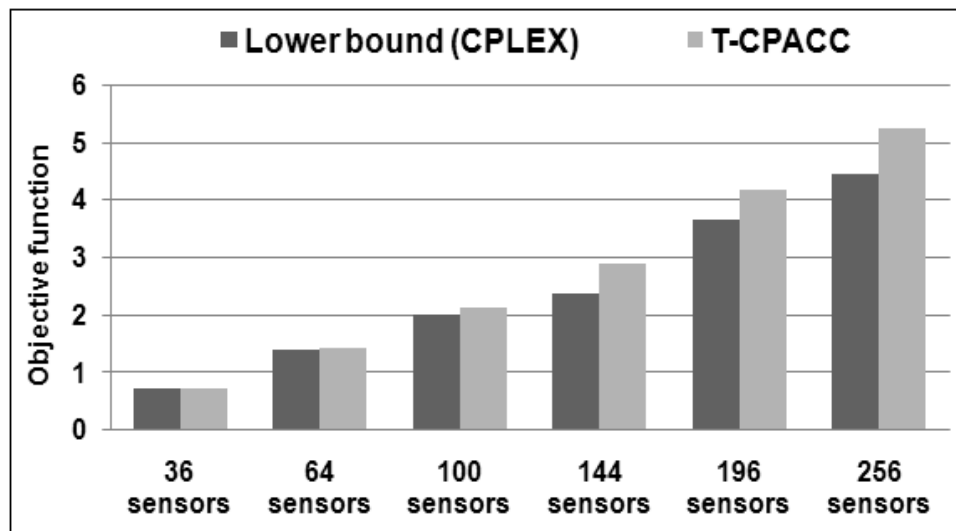


FIGURE 4.5 Objective function generated by T-CPACC vs Cplex

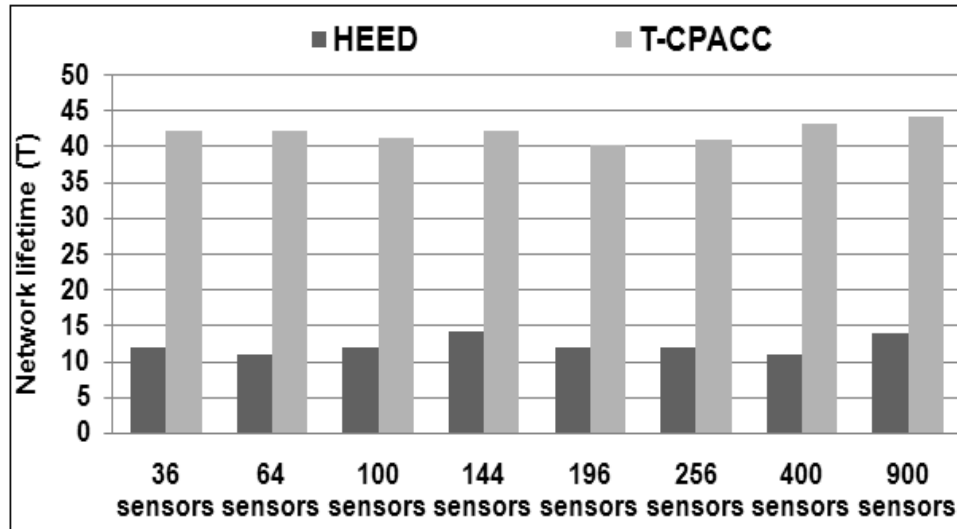


FIGURE 4.6 Network lifetime generated by T-CPACC vs HEED

time for configurations higher than 64 nodes, so we measured and reported the solution of the continuous relaxation of the system (4.3), since the relaxed system takes much less time to be resolved than the original integer-variable system and obviously provides a lower bound. Nevertheless, for configurations with more than 256 sensors, it was difficult to get the relaxed system resolved in reasonable processing times. In Figure 4.5, we can observe that the objective function provided by T-CPACC is very close to the optimal solution for configurations up to 100 sensors and is less than 20% above the lower bound for configurations of more than 100 sensors. The increase in gap is due to the fact that the objective function of the relaxed system's solution is lower than that of the original system. Still, the solution provided by T-CPACC is quite close to the optimal one, which shows that T-CPACC has a very acceptable performance.

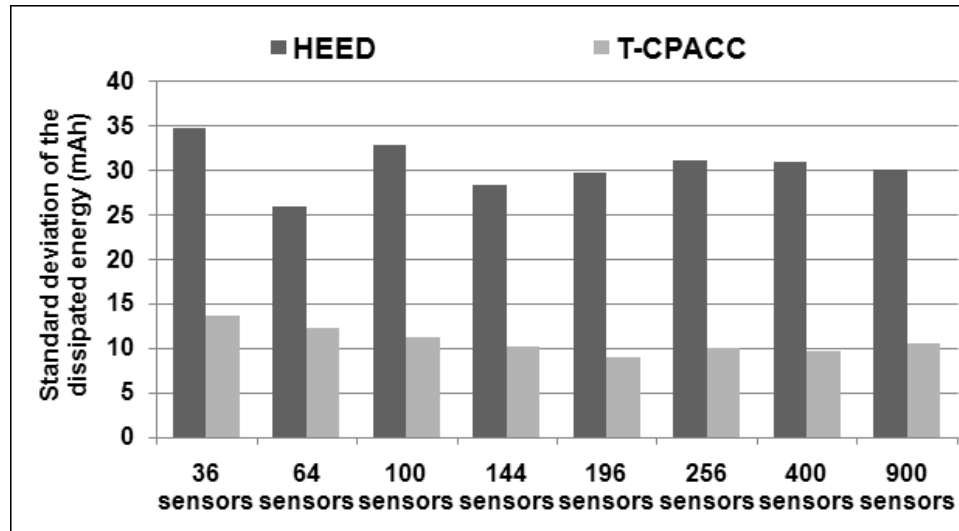


FIGURE 4.7 Standard deviation of the dissipated energy, generated by T-CPACC vs HEED

#### 4.6.2 Measure of the network lifetime without coverage constraint : T-CPACC vs HEED

To evaluate the performance of T-CPACC in terms of network lifetime, we compared it to a well-recognized clustering protocol : Hybrid Energy-Efficient Distributed Clustering (HEED) (Younis et Fahmy, 2004). HEED is a distributed clustering protocol that uses a hybrid combination of the residual energy and the intra-cluster communication cost as attribute for CH selection. In its initialization phase, HEED allows sensors to compute a probability of becoming CH, proportional to its residual energy and to a pre-determined percentage of CHs. Then, during a repetition phase, sensors seek the best CH to connect to. If no CH is found, the sensor doubles its probability to become CH and broadcasts it again to its neighbors, and so forth. This phase stops either when this probability equals 1 (i.e, the sensor elects itself as CH) or when it finds a CH to connect to. HEED performs clustering with one or many transmit powers, but it does not address optimal coverage (i.e, it performs clustering on a set of switched-on sensors). So, to be able to compare our heuristic



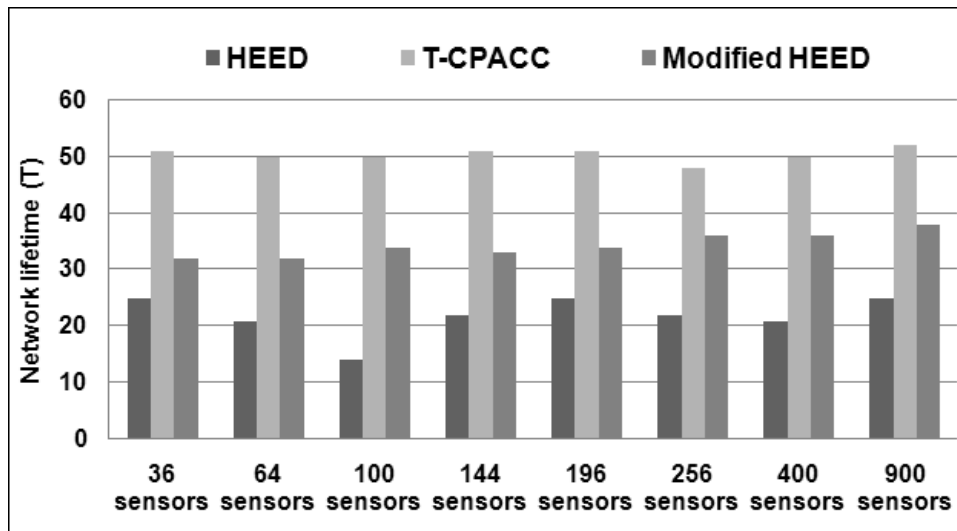


FIGURE 4.8 Network lifetime generated by T-CPACC vs HEED and the Modified HEED, for  $R_s=1$

to HEED without penalizing the latter with regards to optimal coverage, we considered, for T-CPACC, a configuration in which every sensor covers one and only one cell. With this configuration, any admissible solution provided by T-CPACC will have all its sensors switched on and only the clustering performance will be accounted for in comparison to HEED. Figure 4.6 depicts the network lifetime generated by T-CPACC versus HEED. We can see that T-CPACC outperforms HEED for all considered network sizes. This can be explained by a better balancing of the consumed energy over the sensors, so that all nodes' batteries are drained roughly at the same time and thus the network lifetime is longer. This argument is confirmed by the results in Figure 4.7, on which we can see that T-CPACC has a lower standard deviation of the dissipated energy than HEED, which means the dissipated energy is better balanced over sensors.

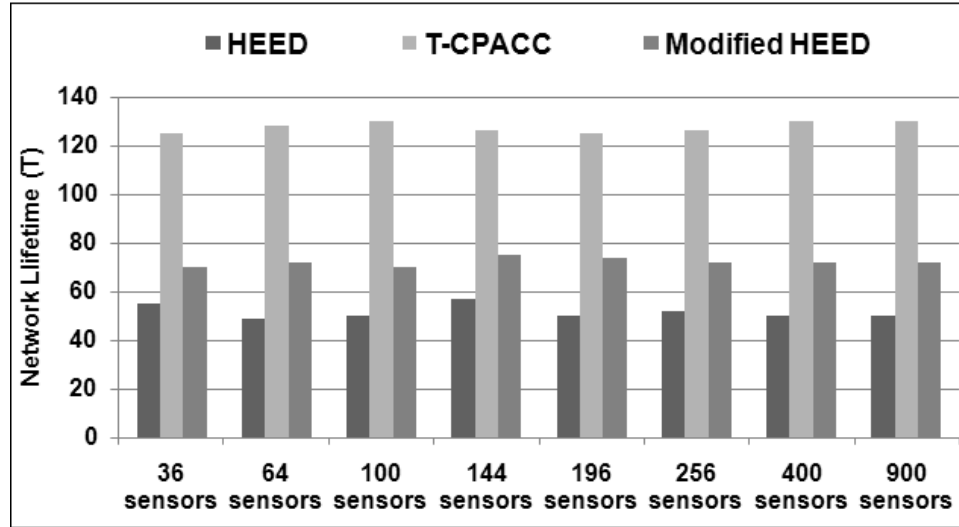


FIGURE 4.9 Network lifetime generated by T-CPACC vs HEED and the Modified HEED, for  $R_s=2$

#### 4.6.3 Measure of the network lifetime with coverage constraint : T-CPACC vs Modified HEED

As mentioned above, HEED does not address optimal area coverage. Hence, to compare T-CPACC to HEED in terms of the combined clustering and coverage constraint, we operated a small modification on HEED to include the optimal area coverage feature : this modification consists in computing, for each cluster generated by HEED, an optimal subset of non-CH sensors that will be activated while the other sensors are turned off, such that full area coverage is insured (Chamam et Pierre, 2007a). Figure 4.8 and Figure 4.9 depict the network lifetime generated by T-CPACC versus HEED and the modified version of HEED that handles optimal coverage, for two different values of the sensing range. We can see that, for  $R_s = 1$ , the solution generated by T-CPACC generates approximately twice the network lifetime generated by HEED's solution and 1.5 times the network lifetime of the enhanced HEED's solution. For  $R_s = 2$ , T-CPACC performs even better. Let's note that the value of the network lifetime, whatever the algorithm used, increases

when the sensing range increases. This is due to the fact that when  $R_s$  increases, less sensors need to be activated to insure full area coverage and thus, more energy is conserved, which allows for a higher network lifetime.

#### 4.7 Conclusion and future work

In this paper, we proposed a novel centralized mechanism for near-optimal state and transmit power assignment to sensors in large-scale, cluster-based, monitoring wireless sensor networks. We considered dense sensor networks with overlapping sensing ranges and where sensors were equipped with multiple transmit powers, allowing them to reach a bigger set of cluster heads at a higher energy cost. Our mechanism seeks energy-optimal configurations where every sensor can be switched off, activated or promoted cluster head and where every active non-cluster-head sensor must be connected to a cluster head, reached using one of the transmit powers a sensor is equipped with. Our mechanism is based on a tabu algorithm that computes a near-optimal network configuration that maximizes network lifetime while ensuring the full coverage of the monitored area and the connectivity of every active sensor to a cluster head. Moreover, our algorithm provides an optimal assignment of transmit powers to active sensors. Simulations show that our mechanism achieves good results with respect to the exact solution of the derived ILP model. Despite its centralized aspect, our mechanism exhibits low complexity and low computation time making its practical implementation adaptable to large-scale networks.

As future research directions, we intend to develop a more sophisticated heuristic to improve the network lifetime. Furthermore, we intend to consider distance-dependent probabilistic event detection, where the probability that a sensor detects an event is function of the distance of that sensor from the event.

## CHAPITRE 5

# A DISTRIBUTED ENERGY-EFFICIENT CLUSTERING PROTOCOL FOR WIRELESS SENSOR NETWORKS

Ali Chamam and Samuel Pierre

Department of Computer & Software Engineering,

École Polytechnique de Montréal,

C.P 6079, succ. Centre-Ville, Montreal (QC), Canada H3C 3A7.

*E-mail : {ali.chamam, samuel.pierre}@polymtl.ca*

### ABSTRACT

Minimizing energy dissipation and maximizing network lifetime are among the central concerns when designing applications and protocols for sensor networks. Clustering has been proven to be energy-efficient in sensor networks since data routing and relaying are only operated by cluster heads. Besides, cluster heads can process, filter and aggregate data sent by cluster members, thus reducing network load and alleviating the bandwidth. In this paper, we propose a novel distributed clustering algorithm where cluster heads are elected following a three-way message exchange between each sensor and its neighbors. Sensor's eligibility to be elected cluster head is based on its residual energy and its degree. Our protocol has a message exchange complexity of  $O(1)$  and a worst-case convergence time complexity of  $O(N)$ . Simulations show that our algorithm outperforms EESH, one of the most recently published distributed clustering algorithms, in terms of network lifetime and ratio of elected cluster heads.

**Key words :**

Wireless Sensor Networks, clustering, network lifetime, distributed, energy efficiency.

**5.1 Introduction**

Wireless sensor networks consist of a large number of low-power, limited-processing capability, MEMS (Micro-Electro-mechanical Systems) capable of measuring and reporting physical variables related to their environment. In surveillance applications, sensors are deployed in a certain field to detect and report events like presence, movement or intrusion in the monitored area (Chong et Kumar, 2003).

As depicted in Figure 5.1, data collected by sensors is transmitted to a special node equipped with higher energy and processing capabilities called "Processing Node" (PN) or "sink" (Chong et Kumar, 2003). The PN collects, filters and aggregates data sent by sensors in order to extract useful information. Due to their energy constraint, wireless sensors usually have a limited transmission range making multi-hop data routing toward the PN more energy-efficient than one-hop transmissions. Unlike in wired networks or in regular wireless networks (cellular, WLAN, etc.), energy conservation in wireless sensor networks is a critical issue that has been

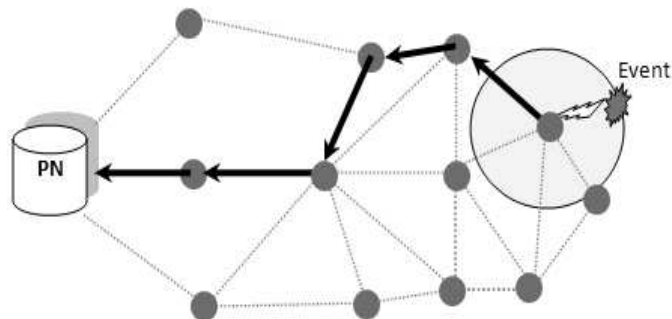


FIGURE 5.1 Multi-hop routing of collected data

addressed by substantial research works (Cardei *et al.*, 2005; Raghunathan *et al.*, 2002). Generally, energy conservation is dealt with on five different levels (Cardei *et al.*, 2005; Chong et Kumar, 2003) :

1. efficient scheduling of sensor states to alternate between sleep and active modes ;
2. efficient control of transmission power to ensure an optimal tradeoff between energy consumption and connectivity ;
3. data compression (source coding) to reduce the amount of uselessly transmitted data ;
4. efficient channel access and packet retransmission protocols on the Data Link Layer ;
5. energy-efficient routing, clustering and data aggregation.

This paper addresses the fifth issue and more specifically energy-efficient clustering. Clustering has been proven energy-efficient in sensor networks (Arboleda et Nasser, 2006; Chong et Kumar, 2003; Younis *et al.*, 2006). In fact, CHs can process, filter and aggregate data sent by sensors belonging to their cluster, thus reducing network load and alleviating the bandwidth. Besides, clustering provides for substantial energy saving (Arboleda et Nasser, 2006; Karl et Willig, 2005) since non-CH sensors are not involved in routing and relaying data and transmissions are only operated by CHs. Yet, since CHs consume more energy in aggregating and routing data, it is important to have an energy-efficient mechanism for CH election and rotation (Heinzelman *et al.*, 2000; Kwon et Gerla, 1999). In flat networks, data is routed from sensors to the PN through peer sensors using one of the many routing protocols proposed in the literature (Akkaya et Younis, 2005). In contrast, in hierarchical (clustered) sensor networks, CHs transmit aggregated data to the PN, either directly (one-hop) (Heinzelman *et al.*, 2000) or in a multi-hop fashion (Mhatre et Rosenberg, 2004b). For the sake of minimizing energy consumption, both optimal number of CHs and

their optimal placement have to be sought.

In this paper, we propose a distributed clustering protocol for sensor networks, called Energy-Efficient Cluster Formation protocol (EECF). We suppose that any sensor can be either a regular sensor or a CH and that every regular sensor has to be connected to a CH, i.e., it has to have at least one CH in its fixed transmission range. We also suppose that only CHs can route data (because they have enough energy to do so and because this limits broadcasting). In EECF, the set of CHs is elected following a three-way message exchange between every sensor and its one-hop neighbors. CH election is mainly based on sensors' respective residual energies and degrees. Run on a flat network, our protocol ends up with a clustered (admissible) sensor configuration that maximizes the network lifetime, while ensuring, under a certain condition, the connectivity of the overlay network composed of CHs.

This paper is organized as follows : in the next section, we present some related work. In Section 5.3, we outline our problem and enumerate our assumptions. In Section 5.4, we describe EECF in detail. In Section 5.5, we provide an analytical study of the complexity of EECF and the connectivity of the solutions. In Section 5.6, we discuss our simulation results. Finally, Section 5.7 concludes this paper and points out some future directions.

## 5.2 Related work

Energy-efficient clustering algorithms for wireless sensor networks have been widely addressed in literature. The main goal of clustering is to efficiently maintain the energy consumption of sensor nodes by involving them in multi-hop communication within a particular cluster and by performing data aggregation and fusion in order to decrease the number of transmitted messages to the sink. Cluster formation is typically based on the energy reserve of sensors and sensor's proximity to the CH (Younis *et al.*, 2006). For instance, Low-Energy Adaptive Clustering Hierarchy

(LEACH)(Heinzelman *et al.*, 2000), one of the first clustering algorithms proposed for sensor networks, is a distributed, proactive, dynamic algorithm that forms clusters of sensors based on the received signal strength and uses local CHs as routers to the sink. Each node makes its own decision whether to become CH based on how often and the last time it has been CH but also on the optimal percentage of CHs in the network (pre-determined value). Transmissions are operated only by CHs which saves energy. LEACH provides a balance of energy consumption through a random rotation of CHs. However, CHs transmit data directly to the sink, which can be energy-consuming in large-scale sensor networks. Power-Efficient GATHERing in Sensor Information Systems (PEGASIS) (Lindsey et Raghavendra, 2002) and its variation Hierarchical-PEGASIS are two improvements of LEACH. Rather than forming multiple clusters, PEGASIS forms chains of sensor nodes so that each sensor transmits and receives from a neighbor and only one node is selected from that chain to convey data to the PN. Still, communication between the elected CH and the PN is one-hop, which may waste energy and prove to be unsuitable for large-sized networks. Weighted Clustering Algorithm (WCA) (Chatterjee *et al.*, 2002) is a reactive clustering algorithm where cluster election is based on the evaluation, for every sensor, of a score function called *combined weight*. This function is a weighted linear combination of the degree, the mobility level, the transmission power and the residual energy of the sensor. Every sensor broadcasts its *combined weight* to its neighbors and the sensor having the lowest weight is elected CH. Hybrid Energy-Efficient Distributed Clustering (HEED) (Younis et Fahmy, 2004) is a distributed clustering protocol that uses a hybrid combination of the residual energy and the intra-cluster communication cost as attribute for cluster head selection. HEED ensures a uniform distribution of CHs across the network and adjusts the probability of CH-selection to ensure inter-CH connectivity. In its initialization phase, HEED allows sensors to compute a probability of becoming CH, proportional to its residual energy and to a pre-determined percentage of CHs. Then, during a repetition phase, sensors seek



the best CH to connect to. If no CH is found, the sensor doubles its probability to become CH and broadcasts it again to its neighbors, and so forth. This phase stops either when this probability equals 1 (i.e., the sensor elects itself as CH) or when it finds a CH to connect to. Distributed Weight-based Energy-Efficient Hierarchical Clustering (DWEHC)(Ding *et al.*, 2005) is a distributed clustering algorithm similar to HEED. In DWEHC, every sensor computes its score as a function of its residual energy and its proximity to its neighbors. In a neighborhood, the sensor with the largest score will be elected CH and all the neighborhood members become attached to that CH. Then, every non-CH member tries to find a "cheaper" way to reach the CH through its non-CH neighbors. These neighbors broadcast their cost to reach the CH and the sensor can then decide whether to use a multi-hop or a direct connection to the CH. Energy-Efficient Hierarchical Clustering (EEHC)(Bandyopadhyay et Coyle, 2003) is another distributed clustering algorithm that performs single and multi-level clustering. In the single-level process, every node announces itself to its  $k$ -hops neighbors (by forwarding, for  $k > 1$ ) as a "volunteer" CH, with probability  $p$ . Then, every sensor connects to the closest CH. Sensors that do not receive any announcement and are not CHs become "Forced CHs". The same process is extended to the multi-level clustering with the exception that only designated CHs participate in the process of election of the second-level cluster set. Energy-Efficient Strong Head clustering (EESH) (Zhou *et al.*, 2007) is a recently published clustering protocol. In EESH, nodes are promoted CHs according to their respective residual energies, their respective degrees and the distance to and the residual energy of their neighbors. For that, EESH evaluates a cost function for every sensor in the network and iteratively elects the node having the greatest cost as CH. This process terminates when all the sensors in the network are connected to at least one CH. EESH has been shown to outperform HEED and LEACH, so we used it as a comparative base in our performance evaluation.

### 5.3 Problem description and assumptions

As discussed in the Introduction, clustering has been proven energy-efficient in sensor networks (Arboleda et Nasser, 2006; Heinzelman *et al.*, 2000; Kwon et Gerla, 1999; Younis et Fahmy, 2004). During the clustering process, it is necessary to take into account aspects such as : cluster size and form, criteria for CH selection, how to control inter-cluster and intra-cluster collisions and energy saving issues. The design of the clustering process is one of the more important issues for the correct functioning of the network, due to the proved efficiency of using a hierarchical communications scheme. In this paper, we consider a sensor network randomly deployed in a certain area and we propose an energy-efficient cluster formation protocol (EECF) that maximizes network lifetime. Run on a flat topology of sensors, EECF ends up with a hierarchical topology in which sensors are organized into clusters having, each, one sensor promoted as CH and all the other regular sensors connected to the closest CH. CHs also act as data relays since they route data received from peer CHs toward the PN, either directly if the PN is within their range, or through a neighbor CH if not. To control the sensors that perform data relaying, we restrict the routing task to CHs, since they are more eligible than the other nodes in terms of residual energy. Therefore, a route must exist from any CH to the PN. On the other hand, because of the energy-intensive data processing and routing tasks, a CH sensor consumes much more energy than regular sensors. If CHs were chosen a priori and fixed throughout the system's lifetime, these nodes would quickly exhaust all their limited energy making them no longer operational and therefore, all the nodes that belong to the cluster would lose their communication ability. In this paper, we dynamically designate CHs so that the elected CH sets are adaptive to the change in the residual energies and the network configuration.

In this paper, we make the following assumptions :

1. The ID and the position of each sensor are fixed and known to both the PN and the sensor itself.
2. Active sensors capture events occurring in their range and transmit data associated with these events straightaway, without any buffering, because sensors are usually not equipped with large (and costly) buffers. However, for some channel access protocols, like TDMA, some little buffers are still needed to hold data waiting to be sent.
3. All sensors have the same sensing range  $R_s$ .
4. All regular (non-CH) sensors have the same transmission range  $R_t$  and consume the same power to transmit one bit of data. Likewise, all CHs have the same transmission range  $R_t^{CH}$  and consume the same power for transmitting one bit of data. Typically,  $R_t^{CH} \gg R_t$ .
5. CHs consume more energy during a time period than regular sensors, since they have a higher transmission range and since they perform additional data-related tasks like aggregation, filtering and routing.
6. Each sensor has an initial energy  $E_0$ . The PN is assumed to have no energy limitation. We assume that regular sensors dissipate the same energy during a unit of time, no matter how the events' distribution is.
7. Network lifetime is defined as the time separating the instant the network starts operating and the instant the network is no more connected (because of the failure of one or more sensor).
8. Sensors have idealized sensing capabilities. That is, inside the sensor's range, quality of sensing is not altered with distance from the sensor.
9. We assume ideal MAC layer conditions, that is, perfect transmission of data on a node-to-node wireless link.

## 5.4 Description of the EECF protocol

EECF is fully distributed. As shown in Figure 5.2, it consists of a number of actions and message exchanges executed by each sensor, which leads to the election of a set of CHs and ensures that each regular sensor (not CH) is connected to a CH.

We will show later that, under a certain condition, any configuration provided by EECF will exhibit a spanning tree connecting all CHs (i.e., a route exists between any pair of CHs). This makes the logical overlay network formed by the set of CHs sufficient to route data from any sensor toward the PN. Here is the description of EECF :

1. Firstly, each sensor broadcasts to all its neighbors a message called Score Advertisement Broadcast (SAB) containing a score function expressing its ability to be promoted CH. In EECF, the score function advertised by each sensor through the SAB is a linear combination of its degree and its residual energy. It is expressed by :  $Score(i) = \alpha \cdot \frac{degree(i)}{N-1} + (1 - \alpha) \cdot \frac{Er(i)}{E_0}$ . The coefficient  $\alpha$  will weigh the importance of the residual energy with respect to the number of neighbors. We will discuss the choice of the value of  $\alpha$  later.
2. After receiving a SAB from all its neighbors, each node  $i$  partitions the set of its neighbors into two subsets :  $N^+(i) = \{v \in \text{Neighbors}(i) \text{ such that } Score(v) \geq Score(i)\}$  and  $N^-(i) = \{v \in \text{Neighbors}(i) \text{ such that } Score(v) < Score(i)\}$ . The idea behind this partitioning is that a regular sensor tends to be connected to a CH that has a higher score (more eligible to be a CH) and thus belonging to  $N^+(i)$ . Subsequently, node  $i$  computes, for each node  $v \in N^+(i)$ , a function  $RCR_i(v)$  representing the relative preference that sensor  $i$  has for sensor  $v$  as its potential CH. In our implementation,  $RCR_i(v)$  takes the form of a probability function for sensor  $v$  to be CH for sensor  $i$ , that is proportional to the rank of

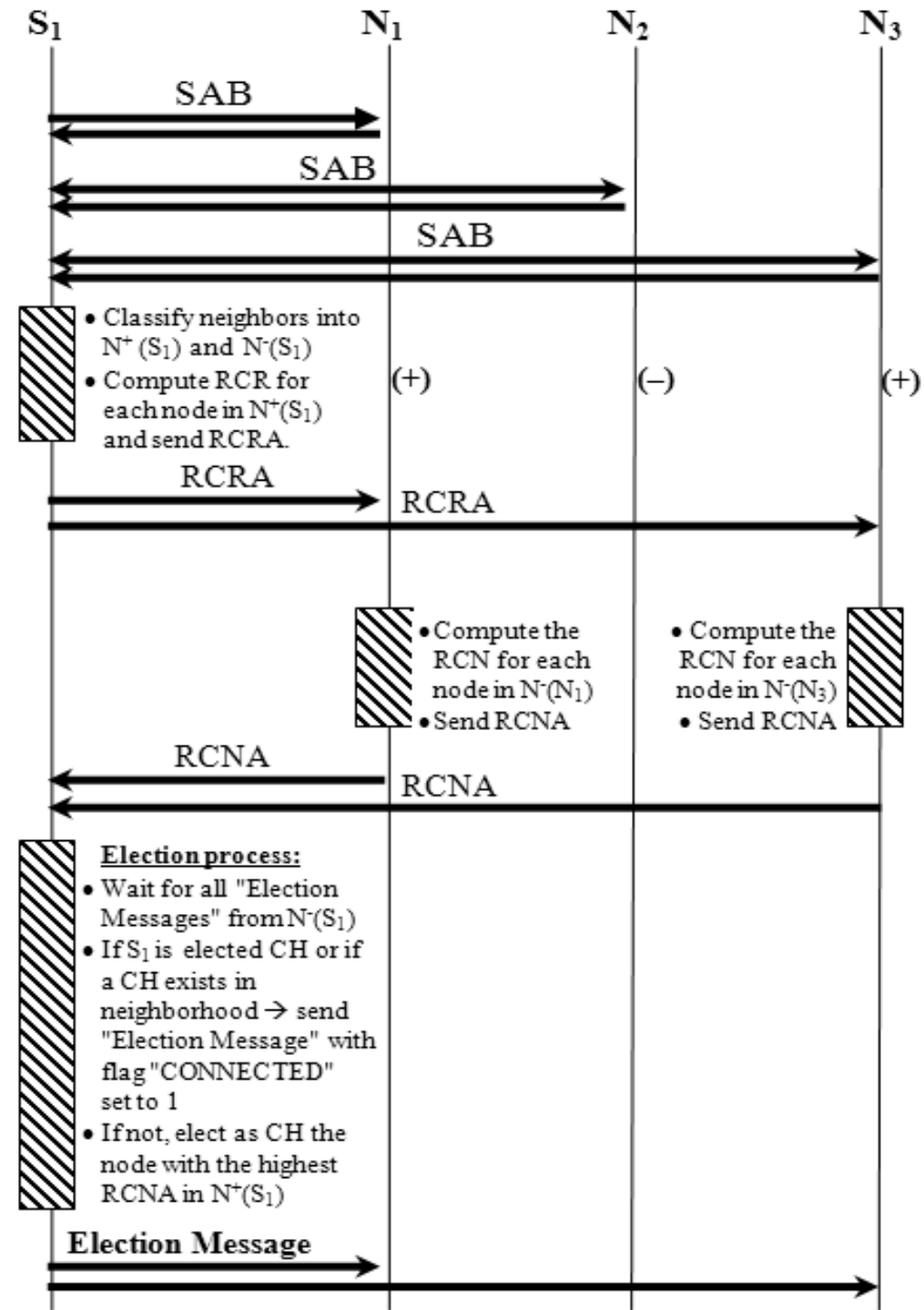


FIGURE 5.2 Description the EECF protocol, executed by every sensor

sensor  $v$  within  $N^+(i)$ . It is expressed as follows :

$$RCR_i(v) = \frac{|N^+(i)| + 1 - Rank(v)}{\frac{|N^+(i)| * (|N^+(i)| + 1)}{2}} \quad (5.1)$$

where  $Rank(v)$  is the rank of sensor  $v$  within  $N^+(i)$  and  $\frac{|N^+(i)| * (|N^+(i)| + 1)}{2}$  is the sum of all possible ranks.

3. Next, sensor  $i$  broadcasts a message called Relative Cluster head Rank Advertisement (RCRA) containing the values  $RCR_i(v)$  of all neighbors  $v \in N^+(i)$ . In the same manner, each sensor  $i$  receives a distinct  $RCRA$  from all its  $N^-(v)$  neighbors expressing its relative "importance" (as a potential CH) to each of them.
4. After receiving all  $RCRA$  messages from its  $N^-(i)$  neighbors, every sensor  $i$  computes a function  $RCN(i)$  representing its *Relative Contribution To the Network (RCN)*.  $RCN(i)$  represents the probability for a node  $i$  to be elected CH, based on the information received from its  $N^-(i)$  neighbors. Since every sensor  $i$  receives from each of its  $N^-(i)$  neighbors an  $RCRA$  representing sensor  $i$ 's probability to be elected CH and since all  $N^-(i)$  neighbors have the same weight, we express  $RCN(i)$  as follows :

$$RCN(i) = 1 - \prod_{v \in N^-(i)} (1 - RCR_v(i)) \quad (5.2)$$

Let us note that  $RCN(i)$  represents the probability for sensor  $i$  to be elected CH. Indeed, in our protocol, a sensor promotes itself CH if, at least, one of its  $N^-(i)$  neighbors elects it as CH. Therefore, if we assume that the CH election probabilities are independent, the probability for a sensor of not being elected CH equals :  $\prod_{v \in N^-(i)} (1 - RCR_v(i))$  and thus  $RCN(i)$  represents the probability for sensor  $i$  to be elected CH.

5. Afterwards, each sensor  $i$  broadcasts a message, called Contribution to the

Network Advertisement (RCNA). The *RCNA* sent by sensor  $i$  will be received by all sensors in  $N^-(i)$ . In the same manner, each node  $i$  receives a distinct *RCNA* from all its  $N^+(i)$  sensors.

6. Finally, after receiving all *RCNAs* from its  $N^+(i)$  neighbors, each node  $i$  takes its decision about the CH it will connect to, unless it has been elected, itself, CH by one of its  $N^-(i)$  neighbors. If  $N^-(i) \neq \emptyset$ , sensor  $i$  has to wait until all *RCNAs* are received from its  $N^+(i)$  neighbors or until a predefined timeout has elapsed to make its decision. This decision is then broadcast to all neighbors through an *ElectionMessage* holding the ID of the sensor designated by sensor  $i$  as its CH (either an existing CH or a newly elected one). The decision process for every node  $i$  is as follows :
  - (a) If sensor  $i$  has no neighbor with lower score ( $N^-(i) = \emptyset$ ), it proceeds with step 6b. Otherwise, it waits until all the decisions of its  $N^-(i)$  neighbors are received or until a predefined timeout has elapsed and then proceeds with step 6b.
  - (b) If, at least, one sensor in  $i$ 's one-hop neighborhood has already been elected CH,  $i$  can connect to it and thus does not need to designate a CH among its  $N^+(i)$ . In this case, sensor  $i$  will broadcast its *ElectionMessage* with the ID of its CH and a flag *CONNECTED* set to 1, which tells the sensors in  $N^+(i)$  that sensor  $i$  does not need to elect a new CH.
  - (c) If sensor  $i$  does not have any already-elected CH in its range, it has to elect one among its  $N^+(i)$  neighbors. It will then select the node  $v_{max} = \arg \max_{N^+(i)} \{RCN(i)\}$  as its CH and broadcasts its *Election Message* with the ID of  $v_{max}$  so that  $v_{max}$  knows that it has been designated CH.
  - (d) Finally, if a sensor  $i$  has no neighbors with higher score ( $N^+(i) = \emptyset$ ) and has no elected CH within its range, it promotes itself CH.
7. When a sensor is elected CH, it broadcasts a message holding its ID and its new state to its neighbors.

Let us note here that a sensor is elected CH if, at least, one of its  $N^-$  neighbors elects it as CH. Furthermore, to avoid having multiple CHs within the same cluster, every sensor waits for the decisions of its  $N^-$  neighbors before making its own decision. In this way, CH election will be made in chain : sensors with the lowest *Score* (having an empty  $N^-$  set) will be the first ones to elect their CH and broadcast their decision. When a sensor is elected CH, it will inform its neighbors so that they can connect to it instead of electing, uselessly, another CH within their respective  $N^+$  neighbors. For the election chain to function properly, every sensor  $i$  has to send its *Election Message* to its  $N^+(i)$  neighbors even if it chooses to connect to an already-elected CH neighbor, because sensors in  $N^+(i)$  will wait for  $i$ 's decision to send their own decisions to their respective  $N^+$  neighbors.

## 5.5 Protocol analysis

As described in the previous section, CH election is performed after a three-way message exchange between each sensor and its neighbors. After receiving the score of each of its neighbors through the SAB, every sensor processes and broadcasts the *RCR* function of its neighbors that have higher SAB. Some of these neighbors can be two-hops away from each other, so the *RCR* received by a sensor may encompass information about some neighbors of its neighbors that are two-hops away, which provides that sensor with aggregated information about its extended (two-hops) neighborhood. Then, every sensor assesses its capacity to become CH by computing its *RCN* function, based on the *RCRs* received from its neighbors having a lower score. Therefore, the *RCN* sent by a node to its one-hop neighbors to express its "eligibility" to become CH takes indirectly into account information about its two-hops neighbors, through the received *RCRs*. In this three-way message exchange, sensors play the role of "information relay" for their potential CH (neighbors having higher scores) by informing them through the *RCRA* about their relative "rank" among



the other potential CH in their range. Similarly, every sensor  $i$  plays the same role of "information relay" for sensors in  $N^-(i)$ , by informing them (through the *RCNA*) of its "importance" (or eligibility) as a potential CH, so that they can decide which one of their neighbors will be chosen as their CH. This "importance", expressed by the *RCN*, cannot be computed by  $i$ 's neighbors since they do not necessarily reach each other and thus cannot necessarily hear each other's *RCRA*. Hence, the final CH election will be based not only on local information (one-hop) but also on aggregated information from the two-hops neighbors and this makes the election protocol more efficient than traditional localized clustering algorithms (Heinzelman *et al.*, 2000; Younis et Fahmy, 2004) where each node's decision is based solely on its one-hop neighbors information. This is the main originality of EECF and we will show later that this novel clustering technique performs very well in terms of energy consumption and network lifetime. But let us discuss first the complexity of EECF and the connectivity of the set of elected CHs.

**Proposition 5.5.1.** *EECF has an algorithmic complexity of  $O(N)$  at each node.*

**Proof.** The execution of EECF at each node  $i$  consists of a sequence of four steps : (1) calculating the score and broadcasting it through the *SAB* message ; (2) calculating the  $RCR(v)$  function for each neighbor  $v$  in  $N^+(i)$  (after all *SAB* messages are received from neighbors) and broadcasting them through the *RCRA* ; (3) calculating  $RCN(i)$  (using the *RCRs* of  $N^-(i)$  neighbors) and broadcasting it through the *RCNA* message ; (4) choosing the neighbor having the highest *RCN* among  $N^+(i)$  as CH and broadcasting the decision. Processing time at Steps (2) and (3) is proportional to the number of  $i$ 's neighbors which, in the worst case, is equal to  $N$ . Hence, the overall complexity is  $O(N)$ .

**Proposition 5.5.2.** *EECF has a worst-case convergence time complexity of  $O(N)$*

and an average convergence time complexity of  $O(\log(N))$ .

**Proof.** We denote by "convergence time" the time it takes for a flat network to become fully clustered using EECF. The worst-case sensor configuration for the convergence time is the one in which sensors are sorted according to their *Scores* and placed in a chain of increasing scores, from right to left, for instance. In this case, the rightmost sensor will start the election process since it has the lowest *Score*. It will elect its left neighbor as CH. Then, every sensor will wait for the *Election Message* from its right neighbor ( $N^-$ ) and send its own decision to its left neighbor ( $N^+$ ). The process terminates at the leftmost sensor which will either be CH or connect to its right neighbor if this latter has been elected CH. With this worst-case configuration, the convergence time is proportional to the number of sensors and thus the complexity is  $O(N)$ . However, in a dense network where sensors are randomly scattered, we have, in average,  $|N^+(i)| = |N^-(i)| = \frac{\delta_i}{2}$ , where  $\delta_i$  denotes the average degree of node  $i$ . In such a case, every sensor has to wait for the *Election Messages* of its  $N^-(i)$  neighbors, which average number is  $(\delta_i/2)$  and then sends its own decision to the set of its  $N^+(i)$  neighbors, having the same average cardinality. The set of sensors having the lowest score (i.e., an empty  $N^-$  set) start the process. Each of these sensors will send its decision to its  $(\delta_i/2)$  neighbors. Then, each of these latter neighbors will send, in its turn, its own decision to its  $(\delta_i/2)$  neighbors, and so forth. Hence, in average, the whole set of sensors can be seen as a layered structure of sensors in which sensors on a certain layer have their  $N^-$  neighbors on the lower layer and their  $N^+$  neighbors on the upper one. Therefore, if we start with  $p$  sensors having the lowest score (and thus forming the first layer), the number of sensors on layer  $k$  ( $k > 1$ ) is  $p \cdot (\frac{\delta_i}{2})^k$ . And since the average convergence time is proportional to the number of layers, it has thus a complexity of  $O(\log(n))$ .

**Proposition 5.5.3.** *When the inter-cluster transmission range  $R_t^{CH}$  is greater or*

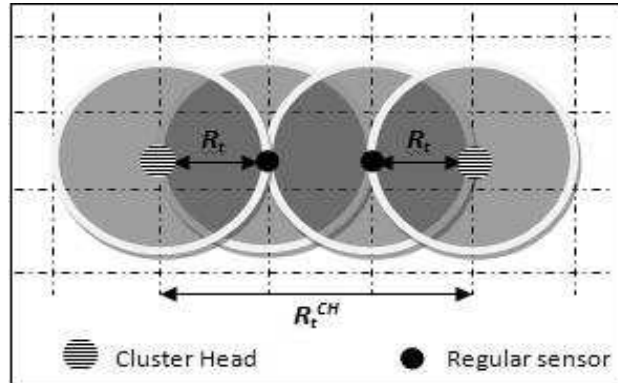


FIGURE 5.3 Necessary condition to have a connected overlay CH graph :  $R_t^{CH} \geq 3.R_t$

*equal to three times the intra-cluster transmission range  $R_t$ , the EECF protocol ends up with a clustered configuration where the overlay graph composed of cluster heads is connected.*

**Proof.** Since the whole network is assumed to be connected and clustered, every node is either a CH or attached to a CH.

Figure 5.3 depicts a configuration exhibiting the longest distance between two adjacent CHs. In fact, when sensors are lined-up, each regular sensor has to be connected to a CH and the two regular sensors have to be directly connected since the network is assumed to be connected. In this worst-case configuration, it is clear that when the transmission range of a CH is equal to or greater than three times  $R_t$ , the overlay graph composed of CHs will be connected.

## 5.6 Protocol implementation and Performance evaluation

### 5.6.1 Experiment setup

In our EECF protocol, CHs are elected based on their respective degrees (which are static) and their residual energies at the instant the clustering protocol is executed. Hence, different residual-energy maps can lead to different sets of CHs. When run once, the EECF protocol elects the set of CHs that tends to maximize the network lifetime as if this configuration will be kept for the whole network lifetime. However, since sensors' residual energies are gradually depleted during network operation, our protocol will be run periodically to adapt the dynamic residual energies : every pre-determined period of time  $T$ , a new set of CHs will be elected, based on the actual residual energies at the instant EECF is run. This adaptivity will balance the CH role among sensors and thus balance energy dissipation. The period  $T$  could be chosen arbitrarily but it has to provide a tradeoff between network dynamic and stability. In fact, when  $T$  is too small, the set of CHs will be re-computed too frequently leading to an instability in the network because of the frequent changes of sensor states. On the other hand, when  $T$  is too big, the adaptivity aspect of the protocol will fade out due to the relatively quick change in residual energies. Without loss of generality, we chose  $T = 1$  *unit*. To measure the sensors' energy dissipation during a time period  $T$ , we assumed a dense network and we used a simplified form of the log-distance path loss channel model described by Heinzelman *et al.* (2002) : we used the free-space fading model for both regular sensors and CHs. Thus, the energy dissipated by a sensor to transmit  $l$  bits over a range  $R$  is :  $E_{Tx} = l.E_e + l.\epsilon.R^2$  and the energy spend to receive an  $l$ -bits message is :  $E_{Rx} = l.E_e$  where  $E_e$  is the energy dissipated by the radio hardware to transmit or receive the signal and  $\epsilon.R^2$  is the amplifier's energy necessary to transmit 1 bit of data over a relatively short distance (Actually,  $\epsilon = \frac{1}{G_e.G_r} \cdot (\frac{4.\pi}{\lambda})^2$  where  $G_t$  is the gain of the transmitter's antenna,  $G_r$  is the gain of

the receiver's antenna and  $\lambda$  is the wavelength). Thus, the energy dissipated by a CH to transmit  $l$  bits of data over its range  $R^{CH}$  and to receive and aggregate an  $l$ -bit length message from each of the regular sensors connected to it, is :

$$E_{CH} = l.E_e + l.\epsilon.(R_t^{CH})^2 + N_{CH}.l.E_e + N_{CH}.l.E_A$$

where  $E_A$  is the energy dissipated in data aggregation and  $N_{CH}$  is the number of regular sensors connected to the CH.

For our simulations, we considered the same parameter values as those used by Heinzelman *et al.* (2002), that are :  $\epsilon = 10pJ/bit/m^2$ ,  $E_e = 50nJ/bit$  and  $E_A = 50nJ/bit/signal$ . We also considered a unique data length value of 100 bytes for all messages.

To evaluate the performance of EECF, we implemented it in C++ and we first studied the effect of the score coefficients  $\alpha$  and  $\beta$  on the network lifetime and the cluster head ratio (number of elected CHs with respect to the total number of sensors). Then, we compared EECF to a recently published clustering protocol : EESH (Zhou *et al.*, 2007). As performance metrics, we chose the network lifetime and the ratio of CHs with respect to the total number of sensors. The second metric assesses the efficiency of our protocol in terms of the minimum number of CHs required to fulfill a fully clustered network. For performance evaluation, we considered 4 grid networks of increasing sizes, composed, respectively, of 100, 400, 900 and 1600 sensors.

### 5.6.2 Effect of the score coefficient $\alpha$ on performance

As we can see from the protocol description, the weighting coefficient  $\alpha$  is not analytically linked to the network lifetime or the cluster head ratio. In fact,  $\alpha$  will only "push" nodes having a good tradeoff between the degree and the residual energy to be elected CHs but we cannot derive any exact mathematical relationship between

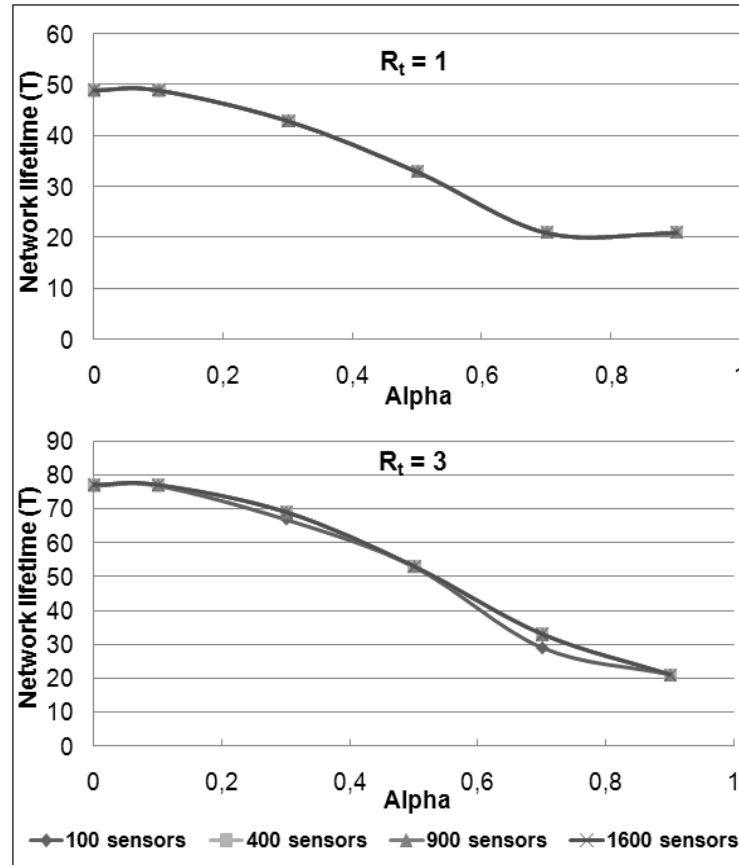


FIGURE 5.4 Network lifetime generated by EECF function of  $\alpha$ , for different values of  $R_t$

$\alpha$  and the network lifetime, or the cluster head ratio. In fact, these two metrics are average values that also depend on the energy map (residual energy of each sensor with respect to its neighbors' residual energies). Hence, we empirically studied the effect of the weighting coefficient  $\alpha$  on the network lifetime and the cluster head ratio.

Figure 5.4 depicts the variation of the network lifetime with respect to  $\alpha$  for two values of the transmission range :  $R_t = 1$  and  $R_t = 3$ . We can see that for  $\alpha = 0.1$ , an optimal value of the network lifetime can be obtained. This small value of  $\alpha$  means that the residual energy is a more important factor in CH election than the

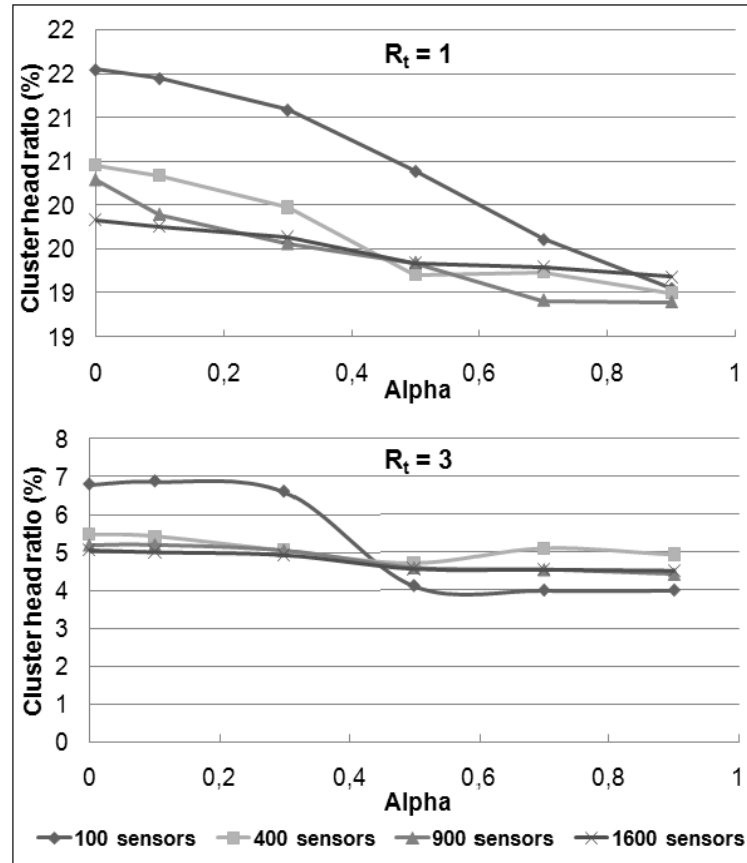


FIGURE 5.5 Ratio "Number of CHs/Total number of sensors" generated by EECF function of  $\alpha$ , for different values of  $R_t$

degree. However, since the optimal value of  $\alpha$  is not zero, the degree still has a small significance. This result was pretty expectable in our case because we considered a grid network where all nodes have almost the same degree (except edge nodes). The optimal value of  $\alpha$  is roughly the same for  $R_t = 1$  and  $R_t = 3$  because, even if the sensors have a greater degree for  $R_t = 3$ , sensors still have the same degree as their neighbors. As far as the CH ratio is concerned, Figure 5.5 shows that the bigger the value of  $\alpha$ , the lower (the better) this ratio. This is also expectable because higher values of  $\alpha$  favor the election of nodes having the highest degree as CHs, regardless of their residual energies, thus reducing the overall number of CHs. In the next paragraph, we will use the average value of  $\alpha=0.5$  to compare our protocol to EESH.

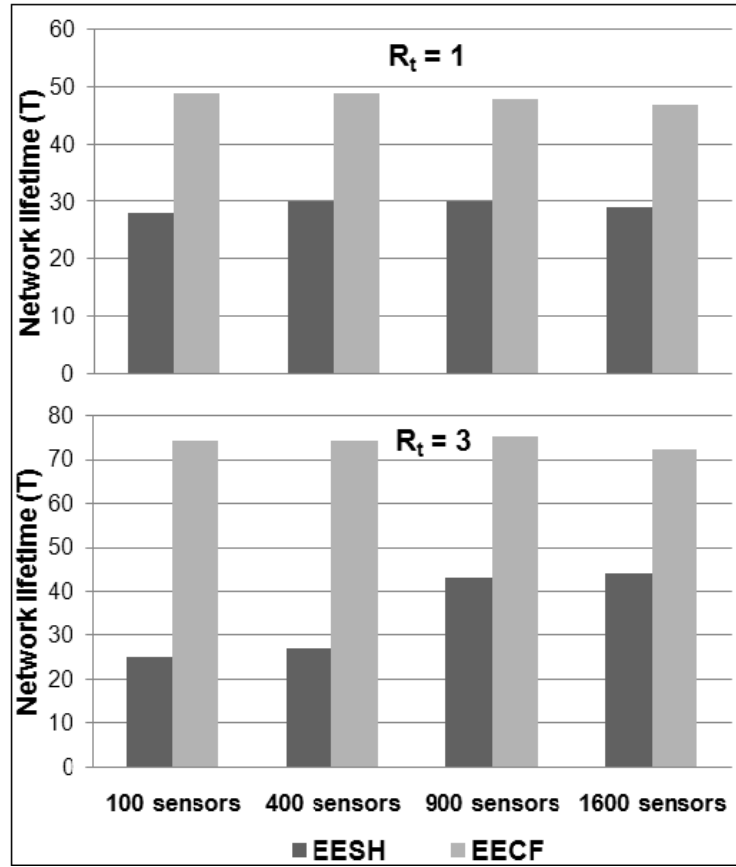


FIGURE 5.6 Network lifetime generated by EECF vs. EESH, for different values of  $R_t$

### 5.6.3 Comparison of EECF and EESH in terms of network lifetime

To measure the network lifetime generated by EECF and EESH, we run these two protocols periodically, at instants  $k.T$  ( $k \in \mathbb{N}$ ) and we used  $T$  as a measure unit of the network lifetime.

Figure 5.6 depicts the network lifetime generated by EECF and EESH, for 2 different values of the transmission range,  $R_t = 1$  and  $R_t = 3$ . We can see that EECF outperforms EESH by approximately 60% for  $R_t = 1$  and more than 70% for  $R_t = 3$ . This could be explained by the two-hops information (resulting from the 3-way message exchange) every sensor uses to make its decision, and which generates



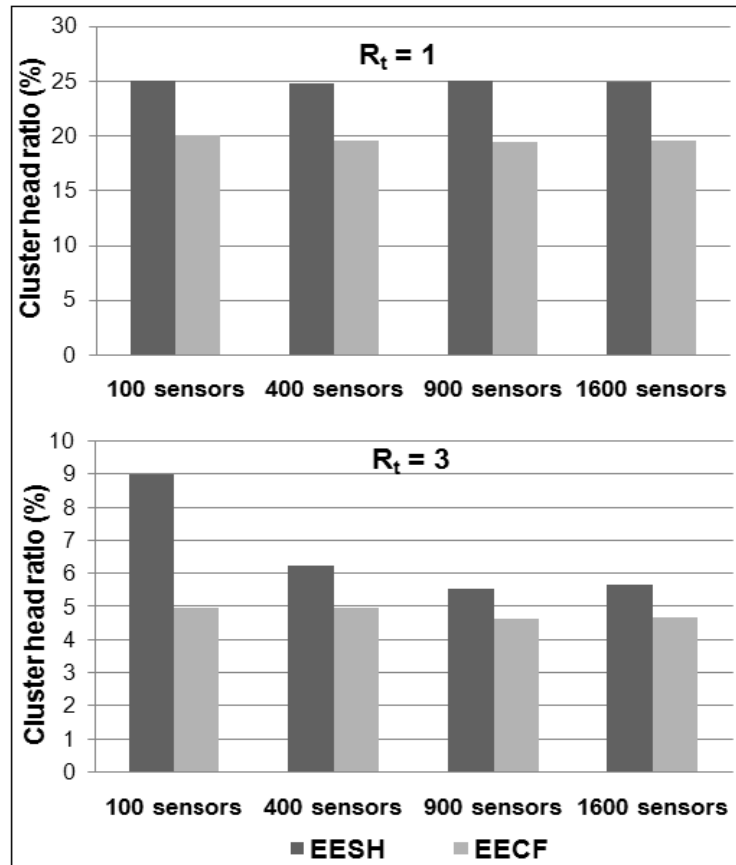


FIGURE 5.7 Ratio "Number of CHs/Total number of sensors" generated by EECF vs. EESH, for different values of  $R_t$

better-cost clusters. We also notice that the better performance of EECF compared to EESH is more emphasized for  $R_t = 3$  than for  $R_t = 1$ . This is because, the bigger the transmission range, the more potential CHS are reached by each sensor and thus EECF has more clustering possibilities and is more likely to find a better configuration compared to EESH.

#### 5.6.4 Comparison of EECF and EESH in terms of the CH ratio (Number of CHs/Total number of sensors)

The ratio "Number of CHs/Total number of sensors" expresses the efficiency of the protocol in terms of the number of CHs required to cluster the network. Figure 5.7 depicts the ratio of CHs for EECF vs. EESH, for  $R_t = 1$  and for  $R_t = 3$ . We can see that, for large-scale configurations (400 sensors and more), EECF outperforms EESH by approximately 20% for  $R_t = 1$  and  $R_t = 3$ . This means that our protocol is able to find admissible configurations with less CHs, which is very consistent with the results regarding the network lifetime, because using less CHs will, in average, increase the network lifetime. Here again, the better CH ratio of EECF compared to EESH can be explained by the two-hops information that every sensor uses to make its clustering-related decision.

### 5.7 Conclusion

In this paper, we proposed EECF, a novel dynamic distributed cluster formation protocol for sensor networks. Our protocol is based on a three-way message exchange between every sensor and its one-hop neighbors and ends up with a clustered network where the overlay graph composed of CHs is connected when the CH transmission range is equal or greater than three times the transmission range of non-CH sensors. We showed that EECF has a message exchange complexity of  $O(1)$  and a worst-case convergence time complexity of  $O(N)$ . Performance evaluation showed that EECF provides a better network lifetime and a better ratio "Number of CHs/Total number of sensors" than EESH, a recently published clustering protocol for wireless sensor networks.

As for future research, we intend to extend our protocol to sensor networks with

multiple transmission ranges, where sensors can reach more neighbors using higher transmission power (at higher cost). Furthermore, we plan on extending our protocol to handle multi-level clustering.

## CHAPITRE 6

### DISCUSSION GÉNÉRALE

Dans ce chapitre, nous faisons une synthèse de nos travaux et contributions en regard de nos objectifs de recherche. Nous discutons également l'approche méthodologique employée et nous analysons les résultats obtenus.

#### 6.1 Synthèse des travaux

Comme dans tout travail de recherche, nous avons commencé cette thèse par une analyse critique de l'état de l'art dans le domaine. Ainsi, nous avons consacré le Chapitre 2 à une revue de littérature exhaustive des mécanismes d'optimisation de l'énergie dissipée et de maximisation de la durée de vie dans les réseaux de capteurs sans fil, et ce, à tous les niveaux du modèle OSI, ainsi que les mécanismes inter-couches (*cross-layer*). Cet état de l'art nous a permis de déceler les faiblesses et les lacunes des mécanismes existants afin de dégager quelques problèmes ouverts pertinents qui ont servi comme orientations de recherche pour cette thèse. Ensuite, nous avons traité dans nos chapitres 3, 4 et 5, trois problématiques différentes, présentées sous forme de trois articles scientifiques de revue.

Dans notre premier article, nous avons considéré un RCSF formé de nœuds déployés aléatoirement pour surveiller une zone donnée, et où chaque capteur pouvait être soit éteint, soit TG, soit allumé mais non TG, chacun de ces états ayant une consommation énergétique spécifique. Nous avons abordé la problématique de la maximisation de la durée de vie du RCSF par deux techniques : l'extinction de certains capteurs non nécessaires à la couverture de la zone surveillée et la structuration du réseau en

grappes, les TG étant les seuls nœuds pouvant relayer les données. Nous avons alors cherché la configuration optimale des capteurs, c'est à dire l'allocation optimale des états (éteint, TG, allumé mais non TG) aux capteurs, qui permet de maximiser la durée de vie du réseau tout en respectant les contraintes de couverture totale de la zone surveillée, de structuration en grappes et de routage entre les TG, qui sont en fait des contraintes très réalistes dans la pratique. En effet, l'une des applications phare des RCSF étant la surveillance des espaces, il est important d'assurer qu'un nombre suffisant de capteurs soit allumé pour assurer la couverture totale de la zone surveillée, le reste des capteurs "redondants" pouvant être éteints pour économiser de l'énergie. Par ailleurs, la formation de grappes permet, elle aussi, de conserver l'énergie, donc chaque capteur, quand il est allumé mais pas TG, doit avoir au moins une TG dans son voisinage. Enfin, les TG désignées doivent pouvoir véhiculer les données au centre de traitement en les faisant relayer par d'autres TG voisines. Ces TG doivent donc appartenir à un arbre de recouvrement. Trouver la configuration optimale qui satisfait à toutes ces contraintes fut l'objectif central de notre premier article. Pour y parvenir, nous avons modélisé notre problème sous forme de problème en nombres entiers, que nous avons prouvé NP-complet. Ensuite, nous avons proposé un mécanisme à base d'heuristique taboue pour résoudre ce modèle en des temps de calcul raisonnables.

Dans notre deuxième article présenté au Chapitre 4, nous avons traité le problème de la maximisation de la durée de vie d'un RCSF dont les nœuds sont munis de plusieurs puissances de transmission. Un capteur peut donc atteindre un plus grand nombre de TG potentielles en utilisant une plus grande puissance de transmission, mais à un coût énergétique plus élevé. Un niveau de complexité est ajouté par rapport au premier article : il s'agit de trouver la configuration optimale des états (éteint, TG, allumé mais pas TG) des capteurs mais aussi des puissances de transmission qui maximise la durée de vie du réseau, toujours sous la contrainte de couverture totale

de la zone surveillée. Ici, chaque capteur doit être soit une TG, soit connecté à une TG voisine qu'il devra pouvoir atteindre moyennant la puissance de transmission qui lui aura été allouée. Nous modélisons notre problème sous forme d'un problème d'optimisation à variables entières, que nous démontrons NP-complet. Nous proposons ensuite un algorithme à base d'heuristique taboue pour trouver une solution proche de l'optimalité en un temps de calcul raisonnable.

Dans notre troisième article présenté au chapitre 5, nous avons proposé un nouvel algorithme réparti de formation de grappes dans les RCSF. Notre algorithme diffère de la majorité des algorithmes proposés dans la littérature par la pertinence des informations récoltées par chaque capteur de la part de ses voisins. Grâce à un mécanisme d'échange de fonctions score, chaque capteur collecte des informations compilées reflétant l'aptitude de ses voisins directs (*1-hop neighbors*) et ceux séparés par deux sauts (*2-hops neighbors*), à devenir TG. Ces informations seront utilisées, par chaque capteur, pour affiner sa prise de décision relativement à son état (devenir TG ou non). Comme chaque capteur a plus de "visibilité" sur son voisinage, sa prise de décision est meilleure, ce qui fait qu'à l'échelle du réseau, notre algorithme donne de meilleurs résultats, en termes de durée de vie du réseau, que d'autres algorithmes bien connus dans la littérature.

## 6.2 Méthodologie

Dans cette thèse, nous avons proposé de nouveaux algorithmes et mécanismes à base d'heuristiques. Ces mécanismes et algorithmes ont été validés par des preuves de concept basées, tantôt sur des preuves théoriques, tantôt sur des simulations expérimentales. Afin d'évaluer la performance des mécanismes proposés, nous avons utilisé des outils de simulation reconnus dans la communauté scientifique tels que ILOG CPLEX (ILOG, 2009), le simulateur de réseaux OPNET ainsi que l'environnement

de développement Visual C++. Par ailleurs, pour générer nos résultats, nous avons suivi un mode opératoire formel et précis basé sur plusieurs batteries de tests qui ont permis de générer les valeurs moyennes ayant servi à la validation des résultats générés par nos heuristiques et algorithmes.

### **6.3 Analyse des résultats**

Dans notre premier article, les résultats de simulation ont montré que l'heuristique que nous avons proposée génère des solutions proches de la solution optimale trouvée, moyennant l'outil de résolution mathématique CPLEX, et qu'elle donne de meilleures performances que d'autres algorithmes répartis de formation de grappes proposés dans la littérature scientifique récente. Dans notre second article, les simulations démontrent que notre heuristique génère de bons résultats, comparée à la solution optimale générée par CPLEX. Et comparé à HEED, un protocole réparti de formation de grappes dans les RCSF, notre algorithme génère de meilleurs résultats en termes de durée de vie du réseau et de répartition de l'énergie dissipée sur les capteurs. Dans notre troisième article, Les simulations montrent que notre nouvel algorithme réparti de formation de grappes donne de meilleurs résultats, en termes de durée de vie du réseau, que d'autres algorithmes répartis proposés dans la littérature scientifique récente.

## CHAPITRE 7

### CONCLUSION

#### 7.1 Sommaire des contributions

Cette thèse a donné lieu à trois principales contributions :

- La modélisation et la résolution du problème de planification optimale des états des capteurs dans un RCSF, dans le but de maximiser la durée de vie du réseau, sous la triple contrainte de couverture, formation de grappes et routage. En effet, plusieurs travaux ont traité séparément certaines de ces contraintes comme, par exemple, la formation optimale de grappes (sans contrainte de couverture), ou encore la maximisation de la durée de vie du réseau par la sélection d'un sous-ensemble optimal de couverture, devant former un graphe connecté, pour qu'une route existe entre chaque paire de capteurs. Cependant, la maximisation de la durée de vie du réseau par une planification optimale des rôles des capteurs, tout en garantissant à la fois les trois contraintes fortement reliées de couverture, de formation de grappes et d'existence d'une arbre de recouvrement reliant les TG sélectionnées, constitue une contribution scientifique intéressante au domaine des RCSF. De plus, les solutions du modèle proposé constituent une borne supérieure, en termes de durée de vie du réseau, pour tout protocole réparti abordant le même problème. Notre modèle peut donc servir comme référence de comparaison pour évaluer la performance d'autres protocoles répartis traitant le même problème.
- La modélisation et la résolution du problème de maximisation de la durée de vie du réseau par formation de grappes dans un contexte de puissances de transmission variables, tout en assurant la couverture de la zone surveillée par le RCSF. Les solutions à ce problème offrent une borne supérieure, en termes de durée de vie du



réseau, aux algorithmes répartis traitant le même problème. En effet, la majorité des protocoles répartis proposés dans la littérature, bien que plus adaptés aux réseaux de grande taille que les algorithmes centralisés, sont généralement basés sur la connaissance, très limitée, des paramètres des nœuds voisins. Ces informations étant locales à chaque nœud, elles ne permettent pas d'aboutir à une configuration optimale. Par contre, un algorithme centralisé, ayant une connaissance globale des paramètres du réseau, génère des solutions optimales qui offrent une borne supérieure sur la durée de vie du réseau, qui peut servir comme base d'évaluation des performances des algorithmes répartis.

- La proposition d'un nouvel algorithme réparti de formation de grappes dans les réseaux de capteurs, qui génère de meilleures performances, en termes de durée de vie du réseau, que d'autres algorithmes répartis populaires dans la littérature scientifique relative au sujet. Notre algorithme constitue donc une amélioration de l'existant en ce qui a trait aux algorithmes répartis de formation de grappes dans les RCSF.

## 7.2 Limitation des travaux

Malgré les contributions énoncées ci-dessus, notre travail présente certaines limitations dont nous pouvons citer :

- Dans les chapitres 3 et 4, nous avons fait l'hypothèse d'un rayon de détection uniforme (i.e, tous les événements qui se produisent à l'intérieur du rayon de détection d'un capteur ont la même probabilité d'être détectés par le capteur en question quelle que soient leurs distances de celui-ci). Cette hypothèse, bien que simplificatrice et souvent considérée dans la littérature, n'est pas toujours valide en pratique car la probabilité de détection d'un événement par un capteur peut

diminuer lorsque la distance entre le capteur et l'événement augmente. Toutefois, nos modèles mathématiques peuvent facilement être adaptés à une détection probabiliste.

- Le caractère centralisé de nos modèles mathématiques 4.3 et 4.1, bien qu'ayant l'avantage de produire des solutions optimales (résolution exacte) ou proches de l'optimalité (heuristiques), ne s'adapte pas au passage à l'échelle (scalability). Dans la pratique, les protocoles répartis sont souvent préférés quand il s'agit de réseaux de grande taille. Toutefois, nos algorithmes centralisés ont le grand avantage d'offrir une borne supérieure, en termes de durée de vie du réseau, aux algorithmes répartis qui donnent généralement des résultats sous-optimaux. De plus, si la taille du réseau est grande mais stable, le centre de traitement peut être dimensionné en termes de processeur et de mémoire, en fonction de la taille du réseau pour résoudre le problème en des temps de calcul raisonnables.
- Dans nos travaux, nous avons abordé la minimisation de l'énergie dépensée par l'application d'une technique intégrée qui associe la désactivation de certains nœuds non nécessaires pour la couverture de la zone surveillée et la formation optimale de grappes. Nous nous sommes alors basés sur l'hypothèse que les liens sans fil au niveau de la couche MAC sont idéaux et nous avons fait abstraction des collisions et des retransmissions, gourmandes en énergie, qui peuvent avoir lieu au niveau de la couche 2. Les pertes et retransmissions de paquets sont souvent dues aux collisions et aux interférences autour des capteurs. Considérer de tels phénomènes dans l'équation de l'énergie va certainement améliorer la pertinence du problème étudié. Par exemple, une planification temporelle optimisée des transmissions et la prise en compte de l'interférence, surtout dans le cas des puissances de transmission variables, sont des sujets qui ne sont pas encore bien maîtrisés dans la communauté scientifique et constituent donc des avenues très intéressantes de recherche future.

### 7.3 Indication des travaux futurs

Même si certaines implantations de réseaux de capteurs ont vu le jour les dernières années, elles sont encore en phase de recherche et beaucoup de défis qui leur sont associés demeurent encore posés. Dans ce qui suit, nous allons citer quelques pistes de recherche future en relation avec les travaux présentés dans cette thèse.

Une première piste de recherche serait d'implémenter nos heuristiques à l'aide d'un simulateur de réseaux tel que OPNET ou NS-2, afin de mesurer la charge réseau générée par la transmission des données inhérentes au mécanisme d'allocation des états, comme les valeurs des énergies résiduelles, l'état alloué à un capteur ou encore la puissance de transmission allouée à celui-ci. La simulation des mécanismes proposés permettra également de mesurer, empiriquement, les retransmissions dues aux collisions et aux interférences et d'en déduire des valeurs moyennes plus réalistes de l'énergie dissipée par chaque capteur.

Une seconde piste consisterait à étendre notre modèle pour qu'il prenne en charge les modèles de détection probabilistes. Ceci consiste à considérer que le rayon de détection d'un capteur est divisé en plusieurs anneaux concentriques et que des probabilités de détection d'événements différentes leur sont associées.

Enfin, une autre piste serait d'aborder le problème d'allocation optimale des états dans les RCSF en tenant compte de l'énergie de transition entre les états "éteint" et "allumé". Sur le long terme, cette énergie de transition n'est pas négligeable (Ragunathan *et al.*, 2002) et le séquençage des états devient donc important. Il serait intéressant de bâtir un planning optimisé des états des capteurs sur  $N$  périodes de temps, tout en garantissant à chaque période les contraintes exigées par l'application, telles que la couverture, la structuration en grappes et le routage entre les TG. Un tel planning sur le temps peut ne pas être fait en temps réel et donc le temps de

calcul relativement important de la solution centralisée deviendrait moins problématique. Nous avons publié une première ébauche de solution à une version simplifiée de ce problème (Chamam et Pierre, 2007b), qui ne prend pas en considération ni la structuration en grappes du réseau, ni le routage entre les têtes de grappes. Il serait donc intéressant d'aborder le problème général avec toutes ses contraintes.

## CONTRIBUTIONS À LA RECHERCHE

Cette thèse a donné lieu à plusieurs publications dans des journaux scientifiques et à des articles de conférences. En voici la liste :

### Articles publiés dans des revues scientifiques :

1. CHAMAM A. et PIERRE, S. (2009). "On the Planning of Wireless Sensor Networks : Energy-Efficient Clustering under the Joint Routing and Coverage Constraint". *IEEE Transactions on Mobile Computing*, 8, 1077-1086. DOI : 10.1109/TMC.2009.16.
2. CHAMAM A. et PIERRE, S. (2009). "A distributed energy-efficient clustering protocol for wireless sensor networks", Accepté pour publication dans *Computers & Electrical Engineering (Elsevier)* (publié en ligne le 05 Mai 2009). ISSN 0045-7906.  
DOI : 10.1016/j.compeleceng.2009.03.008.

### Articles soumis à des revues scientifiques :

3. CHAMAM A. et PIERRE, S. "Energy-Efficient Clustering with Transmit Power Assignment under Coverage Constraint in Wireless Sensor Networks". Soumis pour publication à la revue *Wireless Communications and Mobile Computing (Wiley)*.

**Articles de conférences :**

1. CHAMAM A. et PIERRE S. (2009). "A Distributed Energy-Efficient Cluster Formation Protocol for Wireless Sensor Networks". *IEEE Consumer Communications and Networking Conference (CCNC)*. 1-5.  
DOI : 10.1109/CCNC.2009.4784896
2. CHAMAM A. et PIERRE S. (2008). "Power-Efficient Clustering in Wireless Sensor Networks under Coverage Constraint Networking and Communications". *IEEE International Conference on Wireless and Mobile Computing (WIMOB)*. 460-465. DOI : 10.1109/WiMob.2008.78
3. CHAMAM A. et PIERRE S. (2007). "Energy-Efficient State Scheduling for Maximizing Sensor Network Lifetime under Coverage Constraint". *IEEE International Conference on Wireless and Mobile Computing (WIMOB)*. 63-63.  
DOI : 10.1109/WIMOB.2007.4390857
4. Chamam A. et Pierre, S. (2007). "Optimal Scheduling of Sensors' States to Maximize Network Lifetime in Wireless Sensor Networks". *IEEE International Conference on Mobile Adhoc and Sensor Systems (MASS)*. 1-6. DOI : 10.1109/MOBHOC.2007.4428707
5. Chamam A. et Pierre, S. (2007). "Dynamic sensor activation for maximizing network lifetime under coverage constraint". *International Symposium on Communications and Information Technologies (ISCIT)*. 971-976. DOI : 10.1109/ISCIT.2007.4392156

## BIBLIOGRAPHIE

- AHMED, N., KANHERE, S. et JHA, S. (2005). "Probabilistic coverage in wireless sensor networks". *Proc. IEEE Conference on Local Computer Networks*. 678–681.
- AKAN, O. B. et AKYILDIZ, I. F. (2005). "Event-to-sink reliable transport in wireless sensor networks". *IEEE/ACM Transactions on Networking*, 13, 1003–1016.
- AKKAYA, K. et YOUNIS, M. (2005). "A survey on routing protocols for wireless sensor networks". *Elsevier Ad Hoc Networks*, 3, 325–349.
- AKYILDIZ, I., SU, W., SANKARASUBRAMANIAM, Y. et CAYIRCI, E. (2002). "A survey on sensor networks". *IEEE Communications Magazine*, 40, 102–114.
- AL-KARAKI, J. et KAMAL, A. (2004). "Routing techniques in wireless sensor networks : a survey". *IEEE Wireless Communications*, 11, 6–28.
- ANASTASI, G., CONTI, M., FRANCESCO, M. D. et PASSARELLA, A. (2009). "Energy conservation in wireless sensor networks : A survey". *Elsevier Ad Hoc Networks*, 7, 537–568.
- ARBOLEDA, L. et NASSER, N. (2006). "Comparison of clustering algorithms and protocols for wireless sensor networks". *Proc. Canadian Conference on Electrical and Computer Engineering (CCECE)*. 1787–1792.
- ARISHA, K., YOUSSEF, M. et YOUNIS, M. (2002). "Energy-aware TDMA-based MAC for sensor networks". *Proc. IEEE Workshop on Integrated Management of Power Aware Communications, Computing and Networking (IMPACCT)*.
- ASLAM, J., LI, Q. et RUS, D. (2002). "Three power-aware routing algorithms for sensor networks". *Wireless Communications and Mobile Computing*, 3, 187–208.

- BANDYOPADHYAY, S. et COYLE, E. (2003). "An energy efficient hierarchical clustering algorithm for wireless sensor networks". *Proc. Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*. vol. 3, 1713–1723.
- CARDEI, M., THAI, M., LI, Y. et WU, W. (2005). "Energy-efficient target coverage in wireless sensor networks". *Proc. Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*. vol. 3, 1976–1984.
- CARLE, J. et SIMPLOT-RYL, D. (2004). "Energy-efficient area monitoring for sensor networks". *Computer*, 37, 40–46.
- CAYIRCI, E. (2004). "Wireless sensor networks". In : D. Katsaros et al. (éd), *Wireless information highways* (pp. 273-301). Hershey : Idea group inc.
- CERPA, A. et ESTRIN, D. (2004). "ASCENT : adaptive self-configuring sensor networks topologies". *IEEE Transactions on Mobile Computing*, 3, 272–285.
- CHAMAM, A. et PIERRE, S. (2007a). "Dynamic sensor activation for maximizing network lifetime under coverage constraint". *Proc. International Symposium on Communications and Information Technologies (ISCIT)*. 971–976.
- CHAMAM, A. et PIERRE, S. (2007b). "Optimal scheduling of sensors' states to maximize network lifetime in wireless sensor networks". *Proc. IEEE International Conference on Mobile Adhoc and Sensor Systems (MASS)*. 1–6.
- CHAMAM, A. et PIERRE, S. (2009a). "A distributed energy-efficient clustering protocol for wireless sensor networks". *Elsevier Computers and Electrical Engineering (To appear)*.
- CHAMAM, A. et PIERRE, S. (2009b). "On the planning of wireless sensor networks : Energy-efficient clustering under the joint routing and coverage constraint". *IEEE Transactions on Mobile Computing*, 8, 1077–1086.



- CHATTERJEE, M., DAS, S. et TURGUT, D. (2002). "WCA : A weighted clustering algorithm for mobile ad hoc networks". *Journal of Cluster Computing (Special Issue on Mobile Ad hoc Networks)*, 5, 193–204.
- CHONG, C.-Y. et KUMAR, S. (2003). "Sensor networks : evolution, opportunities, and challenges". 91, 1247–1256.
- CHOU, J., PETROVIC, D. et RAMCHANDRAN, K. (2004). "A distributed and adaptive signal processing approach to exploiting correlation in sensor networks". *Elsevier Ad Hoc Networks*, 2, 387–403.
- CHU, H.-C. et JAN, R.-H. (2007). "A GPS-less, outdoor, self-positioning method for wireless sensor networks". *Elsevier Ad Hoc Networks*, 5, 547–557.
- CI, S., SHARIF, H. et NULI, K. (2005). "Study of an adaptive frame size predictor to enhance energy conservation in wireless sensor networks". *IEEE Journal on Selected Areas in Communications*, 23, 283–292.
- CMOP (2009). "Projet CORIE", consulté le 02 juin 2009, tiré de : <http://www.ccalmr.ogi.edu/>.
- CORMEN, T., LEISERSON, C., RIVEST, R. et STEIN, C. (2001). *Introduction to Algorithms*. MIT Press and McGraw-Hill, seconde édition.
- CORNELL, U. (2009). "Projet COUGAR", consulté le 06 juin 2009, tiré de : <http://www.cs.cornell.edu/bigreddata/cougar/index.php>.
- DING, P., HOLLIDAY, J. et CELIK, A. (2005). "Distributed energy-efficient hierarchical clustering for wireless sensor networks". *Distributed Computing in Sensor Systems*, 32–339.
- ELSON, J. et ESTRIN, D. (2001). "Random, ephemeral transaction identifiers in dynamic sensor networks". *Proc. 21st International Conference on Distributed Computing Systems*. 459–468.

- HEDETNIEMI, S. et LIESTMAN, A. (1988). "A survey of gossiping and broadcasting in communication networks". *IEEE Networks*, 18, 319–349.
- HEFEEDA, M. et AHMADI, H. (2007). "A probabilistic coverage protocol for wireless sensor networks". *Proc. IEEE International Conference on Network Protocols (ICNP)*. 41–50.
- HEINZELMAN, W., CHANDRAKASAN, A. et BALAKRISHNAN, H. (2000). "Energy-efficient communication protocol for wireless microsensor networks". *Proc. 33rd Annual Hawaii International Conference on System Sciences*. vol. 2, 1–10.
- HEINZELMAN, W., CHANDRAKASAN, A. et BALAKRISHNAN, H. (2002). "An application-specific protocol architecture for wireless microsensor networks". *IEEE Transactions on Wireless Communications*, 1, 660–670.
- HEINZELMAN, W., KULIK, J. et BALAKRISHNAN, H. (1999). "Adaptive protocols for information dissemination in wireless sensor networks". *Proc. 5th annual ACM/IEEE international conference on Mobile computing and networking (MobiCom)*. 174–185.
- HWANG, S.-F., SU, Y.-Y., LIN, Y.-Y. et DOW, C.-R. (2006). "A cluster-based coverage-preserved node scheduling scheme in wireless sensor networks". *Proc. 3rd Annual International Conference on Mobile and Ubiquitous Systems*. 1–7.
- ILOG (2009). "ILOG CPLEX", consulté le 02 juin 2009, tiré de : <http://www.ilog.com/products/cplex/>.
- INANC, M., MAGDON-ISMAIL, M. et YENER, B. (2003). "Power optimal connectivity and coverage in wireless sensor networks", vol. TR-03-06. Department of Computer Science, Rensselaer Polytechnic Institute.

- INTANAGONWIWAT, C., GOVINDAN, R. et ESTRIN, D. (2000). "Directed diffusion : A scalable and robust communication paradigm for sensor networks". *ACM International Conference on Mobile Computing and Networking (MobiCom)*. 56–67.
- KARL, H. et WILLIG, A. (2005). *"Protocols and Architectures for Wireless Sensor Networks"*. John Wiley & Sons.
- KOHVAKKA, M., KUORILEHTO, M., HÄNNIKÄINEN, M. et HÄMÄLÄINEN, T. D. (2006). "Performance analysis of ieee 802.15.4 and zigbee for large-scale wireless sensor network applications". *Proc. 3rd ACM international workshop on Performance evaluation of wireless ad hoc, sensor and ubiquitous networks (PEWASUN)*. ACM, 48–57.
- KUBISCH, M., KARL, H., WOLISZ, A., ZHONG, L. et RABAEY, J. (2003). "Distributed algorithms for transmission power control in wireless sensor networks". *Proc. IEEE Wireless Communications and Networking (WCNC)*. vol. 1, 558–563.
- KWON, T. et GERLA, M. (1999). "Clustering with power control". *Proc. IEEE Military Communications Conference (MILCOM)*. vol. 2, 1424–1428.
- LI, X.-Y., WAN, P.-J. et FRIEDER, O. (2002). "Coverage in wireless ad-hoc sensor networks". *Proc. IEEE International Conference on Communications (ICC)*. vol. 5, 3174–3178.
- LINDSEY, S. et RAGHAVENDRA, C. (2002). "PEGASIS : Power-efficient gathering in sensor information systems". *Proc. IEEE Aerospace Conference*. vol. 3, 1125–1130.
- LIU, J.-S. et LIN, C.-H. R. (2005). "Energy-efficiency clustering protocol in wireless sensor networks". *Elsevier Ad Hoc Networks*, 3, 371–388.

- LU, J., WANG, J. et SUDA, T. (2006). "Scalable coverage maintenance for dense wireless sensor networks". *Proc. 3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks (SECON)*. vol. 2, 651–660.
- MADDEN, S. R., FRANKLIN, M. J., HELLERSTEIN, J. M. et HONG, W. (2005). "TinyDB : an acquisitional query processing system for sensor networks". *ACM Transactions on Database Systems*, 30, 122–173.
- MAINWARING, A., CULLER, D., POLASTRE, J., SZEWCZYK, R. et ANDERSON, J. (2002). "Wireless sensor networks for habitat monitoring". *Proc. 1st ACM international workshop on Wireless sensor networks and applications (WSNA)*. 88–97.
- MHATRE, V. et ROSENBERG, C. (2004a). "Design guidelines for wireless sensor networks : communication, clustering and aggregation". *Elsevier Ad Hoc Networks*, 2, 45–63.
- MHATRE, V. et ROSENBERG, C. (2004b). "Homogeneous vs heterogeneous clustered sensor networks : a comparative study". *Proc. IEEE International Conference on Communications*. vol. 6, 3646–3651.
- MHATRE, V., ROSENBERG, C., KOFMAN, D., MAZUMDAR, R. et SHROFF, N. (2005). "A minimum cost heterogeneous sensor network with a lifetime constraint". *IEEE Transactions on Mobile Computing*, 4, 4–15.
- NGAN, H., ZHU, Y., NI, L. et RENYI, X. (2005). "Stimulus-based adaptive sleeping for wireless sensor networks". *Proc. International Conference on Parallel Processing (ICPP)*. 381–388.
- PANICHPAPIBOON, S., FERRARI, G. et TONGUZ, O. (2006). "Optimal transmit power in wireless sensor networks". *IEEE Transactions on Mobile Computing*, 5, 1432–1447.

- PANTAZIS, N. et VERGADOS, D. (2007). "A survey on power control issues in wireless sensor networks". *IEEE Communications Surveys & Tutorials*, 9, 86–107.
- PATWARI, N., ASH, J., KYPEROUNTAS, S., HERO, A.O., I., MOSES, R. et CORREAL, N. (2005). "locating the nodes : cooperative localization in wireless sensor networks". *IEEE Signal Processing Magazine*, 22, 54–69.
- POLASTRE, J., HILL, J. et CULLER, D. (2004). "Versatile low power media access for wireless sensor networks". *Proc. 2nd international conference on Embedded networked sensor systems (SenSys)*. ACM, New York, NY, USA, 95–107.
- POTTIE, G. J. et KAISER, W. J. (2000). "Wireless integrated network sensors". *ACM Communications*, 43, 51–58.
- RAGHUNATHAN, V., SCHURGERS, C., P., S. et SRIVASTAVA, M. (2002). "Energy-aware wireless microsensor networks". *IEEE Signal Processing Magazine*, 19, 40–50.
- RAJENDRAN, V., OBRACZKA, K. et GARCIA-LUNA-ACEVES, J. J. (2003). "Energy-efficient collision-free medium access control for wireless sensor networks". *Proc. 1st international conference on Embedded networked sensor systems (SenSys)*. ACM, New York, NY, USA, 181–192.
- RUDAFSHANI, M. et DATTA, S. (2007). "Localization in wireless sensor networks". *Proc. 6th international conference on information processing in sensor networks (IPSN)*. ACM, New York, NY, USA, 51–60.
- SEDDIK-GHALEB, A., GHAMRI-DOUDANE, Y. et SENOUCI, S.-M. (2006). "A performance study of TCP variants in terms of energy consumption and average goodput within a static ad hoc environment". *Proc. International Conference on Wireless Communications and Mobile Computing (IWCMC)*. ACM, New York, NY, USA, 503–508.

- SHAH, R. et RABAEY, J. (2002). "Energy aware routing for low energy ad hoc sensor networks". *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*. vol. 1, 350–355.
- SHEN, C.-C., SRISATHAPORNPHAT, C. et JAIKAEAO, C. (2001). "Sensor information networking architecture and applications". *IEEE Personal Communications*, 8, 52–59.
- SLEPIAN, D. et WOLF, J. (1973). "Noiseless coding of correlated information sources". *IEEE Transactions on Information Theory*, 19, 471–480.
- STANN, F. et HEIDEMANN, J. (2003). "RMST : reliable data transport in sensor networks". *Proc. 1st IEEE International Workshop on Sensor Network Protocols and Applications*. 102–112.
- SUN, B., GAO, S.-X., CHI, R. et HUANG, F. (2008). "Algorithms for balancing energy consumption in wireless sensor networks". *Proc. 1st ACM international workshop on Foundations of wireless ad hoc and sensor networking and computing (FOWANC)*. ACM, New York, NY, USA, 53–60.
- TIAN, D. et GEORGANAS, N. (2002). "A coverage-preserving node scheduling scheme for large wireless sensor networks". *Proc. 1st ACM international workshop on Wireless sensor networks and applications (WSNA)*. ACM, New York, NY, USA, 32–41.
- TINYOS (2009). "Projet tinyOS", consulté le 06 juin 2009, tiré de : <http://www.tinyos.net>.
- WAN, C.-Y., CAMPBELL, A. T. et KRISHNAMURTHY, L. (2002). "PSFQ : a reliable transport protocol for wireless sensor networks". *Proc. 1st ACM international workshop on Wireless sensor networks and applications (WSNA)*. ACM Press, 1–11.

- WYNER, A. et ZIV, J. (1976). "The rate-distortion function for source coding with side information at the decoder". *IEEE Transactions on Information Theory*, 22, 1–10.
- XING, G., WANG, X., ZHANG, Y., LU, C., PLESS, R. et GILL, C. (2005). "Integrated coverage and connectivity configuration for energy conservation in sensor networks". *ACM Transactions on Sensor Networks*, 1, 36–72.
- XU, Y., HEIDEMANN, J. et ESTRIN, D. (2001). "Geography-informed energy conservation for ad hoc routing". *Proc. 7th annual international conference on Mobile computing and networking (MobiCom)*. 70–84.
- YAN, T., HE, T. et STANKOVIC, J. (2003). "Differentiated surveillance for sensor networks". *Proc. 1st international conference on Embedded networked sensor systems (SenSys)*. ACM, New York, NY, USA, 51–62.
- YAO, Y. et GIANNAKIS, G. (2005). "Energy-efficient scheduling for wireless sensor networks". *IEEE Transactions on Communications*, 53, 1333–1342.
- YE, W., HEIDEMANN, J. et ESTRIN, D. (2004). "Medium access control with coordinated adaptive sleeping for wireless sensor networks". *IEEE/ACM Transactions on Networking*, 12, 493–506.
- YEN, Y.-S., HONG, S., CHANG, R.-S. et CHAO, H.-C. (2007). "An energy efficient and coverage guaranteed wireless sensor network". *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*. 2923–2928.
- YOUNIS, M. et ABBASI, A. (2007). "A survey on clustering algorithms for wireless sensor networks". *Elsevier Computer Communications*, 30, 2826–2841.
- YOUNIS, M., MUNSHI, P., GUPTA, G. et ELSHARKAWY, S. (2006). "On efficient clustering of wireless sensor networks". *Proc. IEEE Workshop on Dependability and Security in Sensor Networks and Systems (DSSNS)*. 10–91.

- YOUNIS, O. et FAHMY, S. (2004). "HEED : a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks". *IEEE Transactions on Mobile Computing*, 3, 366–379.
- ZHOU, W., CHEN, H.-M. et ZHANG, X.-F. (2007). "An energy efficient strong head clustering algorithm for wireless sensor networks". *Proc. International Conference on Wireless Communications, Networking and Mobile Computing (WiCom)*. 2584–2587.
- ZIXIANG, X., LIVERIS, A. et CHENG, S. (2004). "Distributed source coding for sensor networks". *IEEE Signal Processing Magazine*, 21, 80–94.