

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITÉ A. MIRA-BEJAIA



جامعة بجاية
Tasdawit n Bgayet
Université de Béjaïa

FACULTÉ DES SCIENCES EXACTES

DÉPARTEMENT D'INFORMATIQUE

MÉMOIRE

EN VUE DE L'OBTENTION DU DIPLÔME DE MASTER RECHERCHE

Domaine : Mathématiques et Informatique Filière : Informatique

Spécialité : Réseaux et Systèmes Distribués

Présenté par

M^r SAIDANI Abderrezak.

Thème

Un nouvel algorithme de clustering pour
améliorer les performances des réseaux de
capteurs sans fil

Soutenu le 24 sept 2020

Devant le jury composé de :

Nom et Prénom

Grade

M. MOKTEFI Mohand MAA Univ. de Béjaïa Encadrant

M. ATMANI Mouloud MAB Univ. de Béjaïa Examineur

Mme. ZIDANI Faroudja MAB Univ. de bejaia Examineur

Année Universitaire : 2019/2020.

Remerciements

Je remercie vivement mon promoteur M^r Moktefi Mohand pour avoir accepté de m'encadrer, et me suivi durant toute l'année en assurant le suivi scientifique et technique du présent mémoire. Je le remercie pour sa grande contribution à l'aboutissement de ce travail, et pour s'être montré disponible.

Mes remerciements vont aussi aux membre du jury pour l'honneur qu'ils me fait en acceptant de juger ce modeste travail.

Je remercie tous qui ont participé de près ou de loin à l'élaboration de ce travail.



A MES PARENTS

Auxquels Je dois Ce Que Je Suis
Que Dieu Vous Protège
et Vous Prête Une Bonne Santé
et Une Longue Vie

A MA CHERE FEMME

Aida

Pour Les Encouragements Continus et L'aide Constant

A MES ENFANTS

Belsem, Ritel et le petit joli Anes

A TOUS CEUX QUI M'ONT AIMÉ

Et Qui Ne Méritent Pas D'être Oubliés

Je dédie ce travail.

Abderrezak.

TABLE DES MATIÈRES

Table des matières	i
Table des figures	iv
Liste des abréviations	vi
Introduction générale	1
1 Introduction aux réseaux des capteurs	3
1.1 Introduction	3
1.2 Définition d'un capteur	3
1.3 Architecture d'un nœud capteur	4
1.3.1 Architecture matérielle	4
1.3.2 Architecture logicielle	5
1.4 Réseaux de capteurs sans fil RCSFs	5
1.4.1 Présentation d'un RCSF	5
1.4.2 Caractéristiques des RCSFs	6
1.4.3 Architecture des réseaux de capteurs sans fils	8
1.5 Présentation des applications à base de capteurs	11
1.6 Contraintes	14
1.7 Défis des RCSFs	15

1.8	Conclusion	16
2	Protocoles de routage pour les RCSFs	18
2.1	Introduction	18
2.2	Protocoles de routage	18
2.2.1	Définition du routage	18
2.2.2	Présentation des différents types de routage	19
2.3	Travaux antérieurs	25
2.4	Clustering et partitionnement de données	26
2.4.1	Définition	26
2.4.2	Clustering basé sur une représentation non supervisée	26
2.4.3	Méthodes de clustering	29
2.4.4	Domaines d'applications du clustering	30
2.4.5	Types de transmission dans le routage	31
2.4.6	Greedy forwarding	31
2.5	Conclusion	32
3	Schéma de routage basé sur le clustering pour les RCSFs	33
3.1	Introduction	33
3.2	K-Means	33
3.3	Méthode silhouette [36]	34
3.4	Méthode de k-means distribuée dans les réseaux de capteurs	35
3.5	Principe de fonctionnement du schéma de routage proposé	36
3.5.1	Avantages	36
3.5.2	Inconvénients	36
3.5.3	Schéma de routage proposé	36
3.6	Description du déroulement de l'algorithme	38
3.7	Conclusion	42
	Conclusion générale	43
	Bibliographie	43

Résumé

48

TABLE DES FIGURES

1.1	Exemple d'un nœud capteur	4
1.2	Architecture d'un nœud capteur Source	4
1.3	Réseau de capteurs, liaison directe	9
1.4	Réseau de capteurs à routage Multi sauts	10
1.5	Architecture basée sur les clusters	11
1.6	Application des réseaux de capteurs dans le domaine militaire	12
1.7	Mise en œuvre d'un réseau de capteurs pour la surveillance d'activité sismique ou volcanique.	13
1.8	Capteur médical	13
1.9	Des noeuds capteurs dans une installation industrielle	14
2.1	Protocole SPIN	20
2.2	Protocole LEACH avec CH et BS.	23
2.3	Protocole LEACH avec CH et BS.	24
2.4	Fonctionnement du protocole PEGASIS	25
2.5	Greedy Forwarding basée sur la distance	32
3.1	Schéma de création des clusters	39
3.2	Processus d'élection des cluster-heads	41
3.3	Schéma de construction des chemins	41

3.4 Schéma de transmission des données 42

LISTE DES ABRÉVIATIONS

ADC	Analog to Digital Converter
CH	Cluster Head
MAC	Media Access Control
DD	Directed Diffusion
GPS	Global Positioning System
HEED	Hybrid, Energy-Efficient, Distributed approach
LEACH	Low Energy Adaptive Clustering Hierarchy
MEMS	Electro Mechanical Systems
OS	Open Systems Interconnection
PEGASIS	Power-Efficient GATHERing in Sensor Information Systems
QoS	Quality of Service
RCSF	Réseau de Capteurs Sans Fil
RSSI	Received Signal Strength Indication
SPIN	Sensor Protocol for Information via Negotiation
TCP/IP	Transfer Control Protocol/Internet Protocol

INTRODUCTION GÉNÉRALE

Ces dernières années, les progrès dans le domaine de la technologie des systèmes micro électromécaniques (MEMS) et le domaine des communications sans fil ont permis l'apparition de nouveaux équipements électroniques dotés de moyens de communication sans fil, peu coûteux et peuvent être configurés pour former un réseau autonome. Ces réseaux sont les Réseaux de Capteurs Sans Fil (RCSF).

Les réseaux de capteurs se composent d'un ensemble de capteurs capables de communiquer entre eux et de fournir des informations par la combinaison des mesures prises par les différents capteurs. Ces RCSFs sont utilisés dans plusieurs domaines tels que : le domaine militaire, médical, industriel, transport, et ils sont utilisés aussi dans les applications de surveillance de l'environnement (monitoring)... etc.

Parmi les défis essentiels des réseaux de capteurs sans fil nous pouvons citer : le maintien des mêmes performances lors du passage à l'échelle, la consommation d'énergie et surtout la durée de vie puisque les capteurs sont généralement dotés de batteries non rechargeables et leur remplacement est quasiment impossible quand ils sont déployés dans des zones hostiles. À cet effet, dans le cadre de ce projet de fin d'études, nous proposons une nouvelle approche de routage qui permet de relever ces défis. Cette approche proposée est basée sur le clustering pour organiser le réseau et un schéma de routage multi sauts (CH-to-CH) selon un algorithme glouton amélioré (Greedy Forwarding) pour acheminer les données à la station de base. Ceci est dans le but de minimiser la consommation d'énergie et par conséquent augmenter la durée de vie du réseau.

Ce manuscrit s'articule autour de trois chapitres :

- Dans le premier chapitre, une présentation générale des réseaux de capteurs sans fil.
- Dans le deuxième chapitre, une présentation de quelques protocoles de routage qui ont fourni de bonnes performances pour les RCSFs.
- Le troisième chapitre détaille ma solution de routage proposée.

Et enfin, je termine par une conclusion générale qui récapitule le travail réalisé dans ce projet de fin d'études et quelques perspectives

CHAPITRE 1

INTRODUCTION AUX RÉSEAUX DES CAPTEURS

1.1 Introduction

Les progrès dans le domaine de l'électronique miniaturisée et les communications sans fil ont donné naissance à des composants peu coûteux capables de prélever des grandeurs physiques environnementales. Ces composants sont appelés des nœuds capteurs qui ont la capacité de s'auto-organiser pour former un réseau de capteurs sans fil (RCSF).

Dans ce chapitre, je nous allons présenter les RCSFs, en commençant par définir les nœuds capteurs et leurs architectures ensuite définir les RCSFs ainsi que leurs caractéristiques, puis leurs domaines d'application et les contraintes liées aux réseaux de capteurs sans fils et les défis des RCSFs et nous terminons ce chapitre par une petite conclusion.

1.2 Définition d'un capteur

Un capteur ou un nœud capteur est un système qui permet de mesurer et transformer une grandeur physique en donnée exploitable. Il permet donc de convertir en valeur

numérique une donnée environnementale mesurée. Un capteur peut regrouper plusieurs sous-systèmes responsables chacun d'une ou plusieurs tâches. Il permet l'acquisition des données, leurs traitements et leurs transmissions. Un capteur sans fils est également équipé d'une source d'énergie.



FIGURE 1.1 – Exemple d'un nœud capteur

1.3 Architecture d'un nœud capteur

1.3.1 Architecture matérielle

Un nœud capteur est représenté par quatre unités (figure 1.2) :

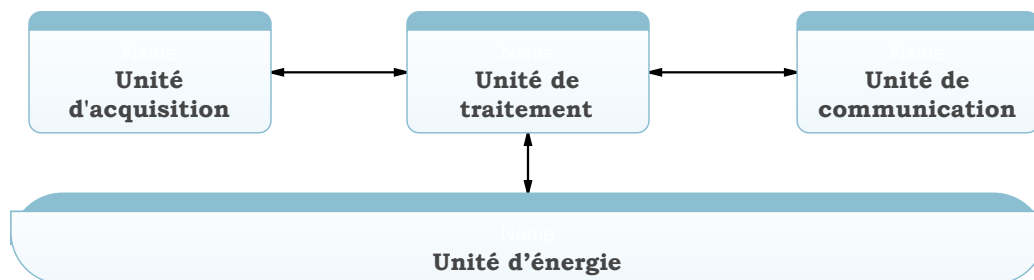


FIGURE 1.2 – Architecture d'un nœud capteur Source

- **Unité d'acquisition** : elle est responsable de la capture des événements du milieu physique et de la conversion du signal analogique vers un signal numérique ;

- **Unité de traitement** : les événements captés sont transformés en données numériques et sont traités par le processeur ;
- **Unité de communication** : elle est composée d'un module radio (émetteur/récepteur) permettant l'échange des données entre les différents nœuds ou entre les capteurs et la station de base ;
- **Unité d'énergie** : c'est la batterie qui dans la plupart des cas n'est ni remplaçable ni rechargeable.

1.3.2 Architecture logicielle

Comme pour les équipements informatiques, un système d'exploitation est mis en place afin de permettre à toutes les unités d'accomplir leurs tâches. Les systèmes les plus connus sont TynOS et Contiki.

1.4 Réseaux de capteurs sans fil RCSFs

1.4.1 Présentation d'un RCSF

Un Réseau de Capteurs Sans Fil (RCSF ou WSN : Wireless Sensor Network en anglais) est un système distribué mettant en communication un grand nombre d'entités autonomes appelées "capteurs sans fil", ou simplement "capteurs". C'est un ensemble de nœuds capteurs intelligents de petite taille, à faible coût et de puissance limitée, variant de quelques dizaines d'éléments à plusieurs milliers (appelés aussi nano ordinateurs). Les capteurs communiquent par radio sans fil afin d'acheminer l'information à une ou plusieurs stations de base.

Dans ces réseaux, chaque nœud est capable de surveiller son environnement et de réagir en cas de besoin en envoyant l'information collectée. Les nœuds sont équipés d'une batterie qui est une source d'énergie non durable et que son remplacement est difficile, voire impossible, surtout dans des endroits hostiles [1, 6].

1.4.2 Caractéristiques des RCSFs

La conception des réseaux de capteurs est influencée par de nombreux facteurs qui représentent la base de la conception de protocoles ou d'algorithmes pour les réseaux de capteurs. Il existe plusieurs caractéristiques qui distinguent les RCSFs des autres réseaux sans fil dont on cite les plus importantes [7, 8] :

- **Durée de vie du réseau** : C'est l'intervalle de temps qui sépare l'instant de déploiement du réseau de l'instant où l'énergie du premier nœud s'épuise ou le dernier nœud ou un pourcentage de nœuds. Selon l'application, la durée de vie exigée pour un réseau peut varier entre quelques heures et plusieurs années.
- **Bande passante limitée** : Afin de minimiser l'énergie consommée lors de transfert de données entre les nœuds, les capteurs opèrent à bas débit. Typiquement, le débit utilisé est de quelques dizaines de kb/s. Un débit de transmission réduit n'est pas handicapant pour un réseau de capteurs où les fréquences de transmission ne sont pas importantes.
- **Média du transport** : Dans un réseau de capteurs, la communication multi-sauts entre les nœuds est réalisée avec des liens sans fil à l'aide de média optique, infrarouge ou radio. La plupart des réseaux de capteurs utilisent des circuits de communication à radio fréquence grâce à leur faible coût ainsi que leur facilité d'installation [6].
- **La topologie du réseau** : Elle est en constante évolution à cause du changement de l'état d'activité des capteurs (actif, mise en veille et passif). Il faut que les capteurs soient capables d'adapter leur fonctionnement afin de maintenir la topologie souhaitée et assurer la couverture totale de la zone de déploiement.
- **Déploiement** : Les capteurs sont soit déployés manuellement quand leur nombre est petit soit de manière aléatoire lancés en masse depuis un avion, par exemple.
- **Le passage à l'échelle** : La plupart des protocoles sont conçus pour des réseaux de capteurs de taille moyenne. Cependant, ces protocoles sont dits efficaces si les performances des réseaux ne doivent pas chuter d'une manière drastique quand le nombre de capteurs augmente dans le réseau.

- **La consommation d'énergie** : L'économie d'énergie est l'une des problématiques majeures dans les réseaux de capteurs. En effet, la recharge des sources d'énergie est souvent trop coûteuse et parfois impossible. Il faut donc que les capteurs économisent au maximum l'énergie afin de pouvoir fonctionner pour une durée maximale.
- **L'auto-configuration** : Les capteurs sont généralement déployés aléatoirement dans des zones d'intérêt hostiles et en grand nombre, par conséquent, aucune intervention humaine ne peut être requise pour assurer leur organisation. L'auto-configuration de ces réseaux s'avère nécessaire pour leur bon fonctionnement.
- **La qualité de service** : cette caractéristique est visée dans les réseaux de capteurs afin d'assurer la fiabilité de livraison des paquets entre les nœuds sources et réduire le délai de réception de ces paquets. Les protocoles doivent vérifier la stabilité du réseau ainsi que les données redondantes transmises dans le réseau selon la répartition du trafic [7].
- **Le mode de transmission** : Il joue un rôle important dans les réseaux de capteurs. Les nœuds peuvent transmettre des données vers d'autres nœuds dans le réseau en utilisant une seule fréquence ou bien plusieurs fréquences.
- **Les nœuds de détection** : Ils peuvent être mobiles ou statiques, selon l'application.
- **La scalabilité** : Contrairement aux réseaux sans fil traditionnels (personnel, local ou étendu), un RCSF peut contenir un très grand nombre de nœuds capteurs (des centaines, des milliers...) [9]. Un réseau de capteur est scalable parce qu'il a la faculté d'accepter un très grand nombre de nœuds qui collaborent ensemble afin d'atteindre un objectif commun.
- **La tolérance aux pannes** : Dans le cas de dysfonctionnement d'un nœud à cause de l'épuisement de son énergie par exemple, ou aussi en cas d'ajout de nouveaux nœuds capteurs dans le réseau, ce nœud doit continuer à fonctionner normalement sans interruption [9]. Ceci explique le fait qu'un RCSF n'adopte pas de topologie fixe, mais plutôt dynamique.

- **Une densité importante des nœuds :** Les RCSFs sont caractérisés par leur forte densité [10]. Cette densité peut atteindre, selon le type d'application $20\text{noeuds}/m^2$ [9] surtout lorsqu'il s'agit de capteurs associés à des petits objets connectés.
- **Une collaboration entre les nœuds :** Les contraintes strictes de consommation d'énergie mènent les nœuds capteurs à détecter et traiter les données d'une manière coopérative afin d'éviter le traitement redondant d'une même donnée observée, et qui aura un impact négatif sur la perte d'énergie [10].

1.4.3 Architecture des réseaux de capteurs sans fils

Un réseau de capteurs est composé de plusieurs capteurs mis en réseau avec une station de base dite « Puits » qui collecte les informations transmises par ces derniers. Ce genre de réseau a connu à travers le temps une évolution de point de vue architecture. Cette évolution concerne la façon avec laquelle les capteurs sont organisés au sein du réseau et la manière qu'ils utilisent pour communiquer entre eux ou avec la station de base. On distingue trois types d'architecture.

- **Les réseaux à liaison directe** [1] Dans ces premiers réseaux, les capteurs communiquaient directement avec le « puits ». Cette architecture est illustrée dans la figure 1.3.

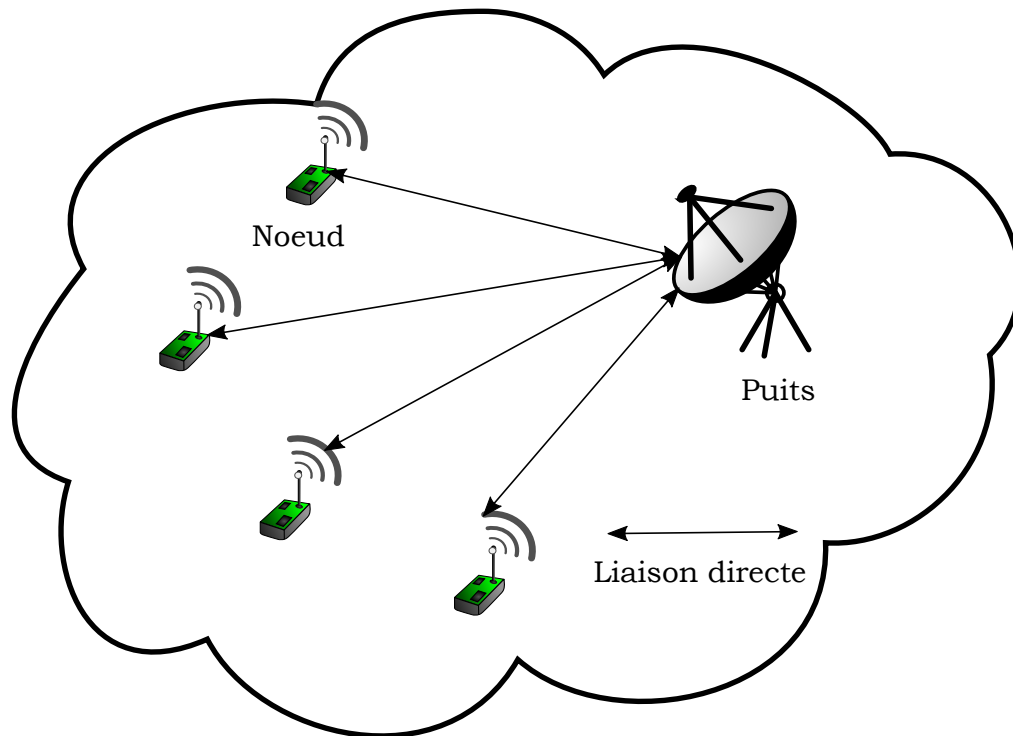


FIGURE 1.3 – Réseau de capteurs, liaison directe

Ce sont les premières architectures qui ont été élaborées, le routage dans ce type d'architecture ne présentait pas un problème puisque tous les nœuds du réseau savaient où envoyer l'information.

- **Réseaux multisauteurs** : Dans ce cas, les nœuds capteurs sont interconnectés entre eux et l'acheminement de l'information jusqu'à la station de base peut se faire de deux façons. Les données sont acheminées directement vers le nœud « puits » si le capteur est à portée de ce dernier. Dans le cas où les données proviennent d'un capteur loin du nœud puits, celles-ci vont passer par les autres capteurs intermédiaires jusqu'à arriver à la station de traitement (figure 1.4).

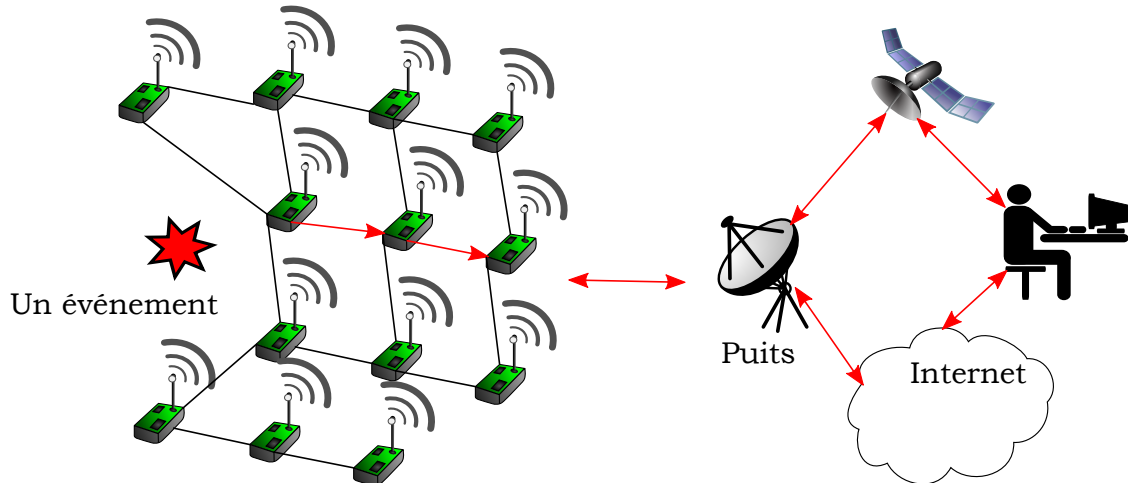


FIGURE 1.4 – Réseau de capteurs à routage Multi sauts

Les données collectées par les stations de bases sont ensuite acheminées vers des machines puissantes qui auront la tâche de les analyser et les exploiter [3].

- **Les réseaux à base de clusters** [6] : Dans ce type de réseaux, les deux premières techniques sont combinées. En effet, il s'agit de former des petits réseaux (cluster) au sein d'un seul réseau. Pour chaque cluster, on trouve un Clusterhead (CH) et les autres capteurs. Ces derniers ne vont échanger les données qu'avec le CH, cette partie constitue le réseau à liaison directe. Ensuite, les CH vont utiliser la technique multisauts pour acheminer l'information jusqu'à la station de base (voir figure 1.5)

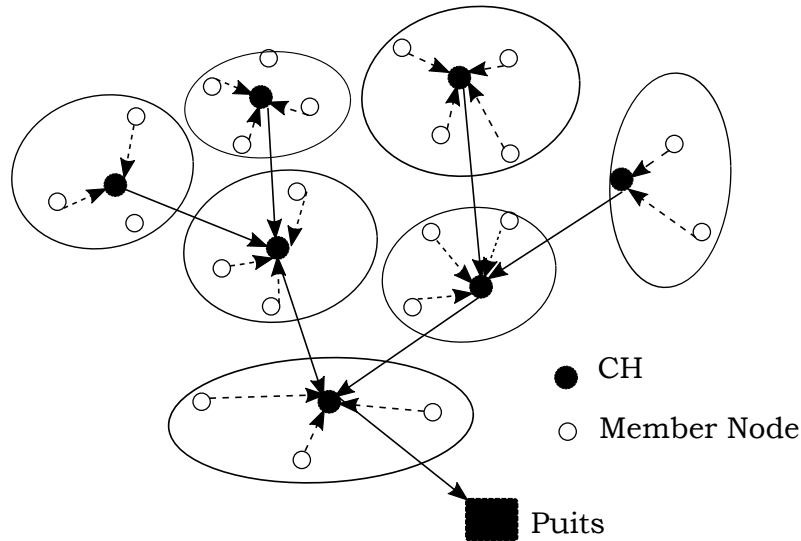


FIGURE 1.5 – Architecture basée sur les clusters

Dans un seul réseau, on peut rencontrer différents types de capteurs qui ont la charge de récolter des informations différentes selon le domaine d'intérêt et les événements environnementaux que l'on souhaite surveiller. On parle d'hétérogénéité de capteurs.

1.5 Présentation des applications à base de capteurs

On distingue quatre grandes familles d'applications [6] :

- **Les applications militaires** : leurs objectifs incluent la surveillance des forces amies, des équipements, des munitions et du champ de bataille, la détection des attaques (biologiques, chimiques, etc.), la reconnaissance des forces ennemies, l'évaluation des dommages, etc. Un exemple d'application militaire est VigilNet [7]. C'est un réseau de surveillance à grande échelle conçu pour le pistage de cibles dans des environnements difficiles.

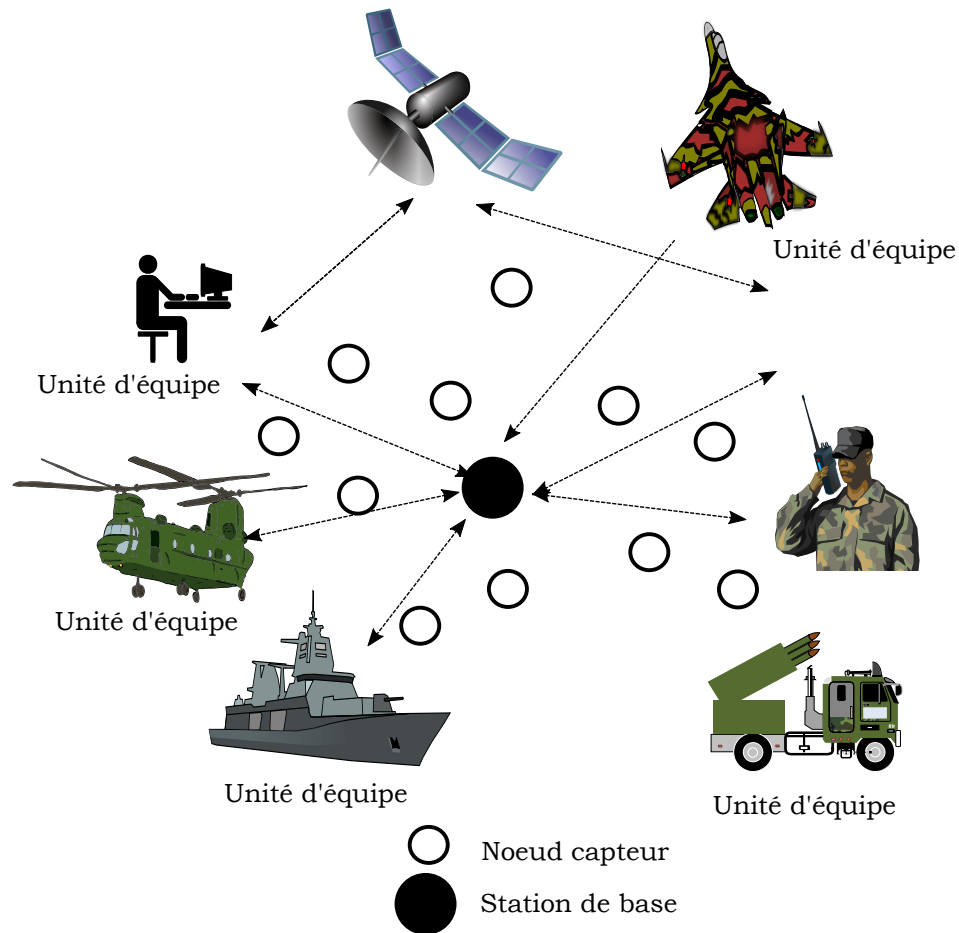


FIGURE 1.6 – Application des réseaux de capteurs dans le domaine militaire

- **Les applications pour l'environnement** : comme par exemple la détection de feux dans les forêts, la détection d'inondation, l'agriculture de précision, le recensement et la surveillance d'animaux, etc. Le projet "Great Duck Island" [8] est un exemple d'application pour l'environnement. C'est le résultat d'une collaboration entre "College of Atlantic" et "Intel Research Laboratory" à Berkeley pour étudier la distribution et l'affluence des oiseaux de mer à Great Duck Island, Maine.

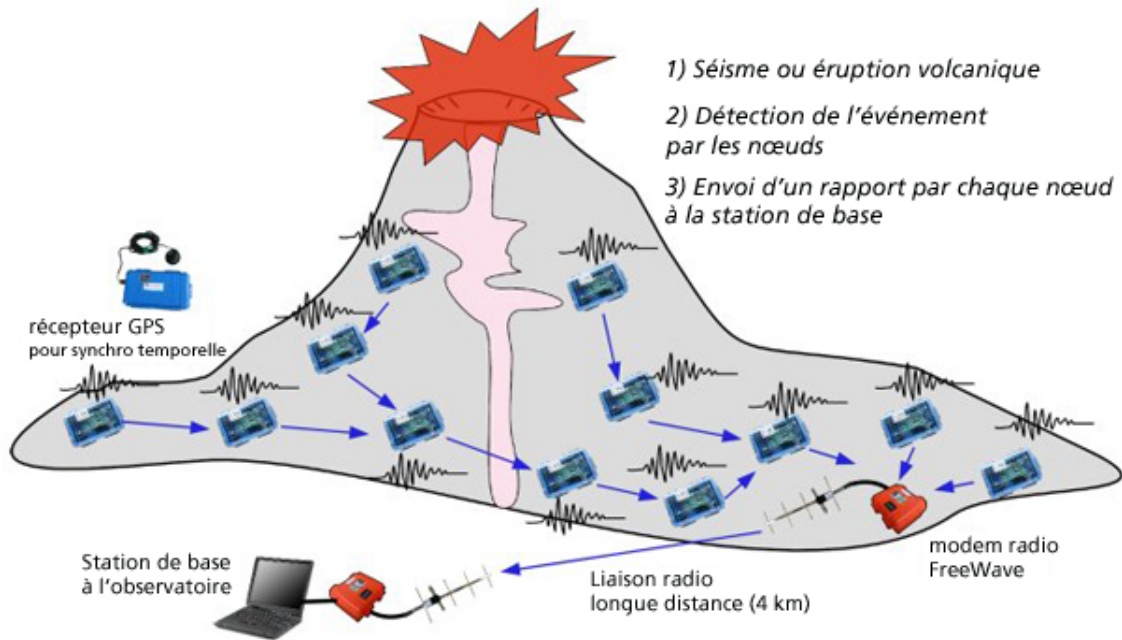


FIGURE 1.7 – Mise en œuvre d'un réseau de capteurs pour la surveillance d'activité sismique ou volcanique.

- **Les applications médicales** : comme celles permettant la télésurveillance des patients et de leurs fonctions vitales, le repérage des patients, des médecins et du matériel dans un hôpital, etc.



FIGURE 1.8 – Capteur médical

- **Les applications industrielles** : la supervision, la localisation et l'inventaire des équipements industriels sont des exemples d'applications industrielles. Par

exemple, la société fireflies-rtls propose une solution de localisation de biens en temps réel destinée aux industriels.



FIGURE 1.9 – Des noeuds capteurs dans une installation industrielle

Selon le type d'applications et son environnement de déploiement, nous pouvons être confrontés à des contraintes qui auront une influence directe sur la qualité de service.

1.6 Contraintes

Plusieurs contraintes sont liées aux réseaux de capteurs sans fils. Les principales sont :

- **Le support utilisé** : c'est le média sans fils qui par nature n'est pas fiable et son état varie dans le temps et est fortement lié à l'environnement de déploiement ;
- **La densité du réseau** : comme nous l'avons déjà mentionné dans le paragraphe « caractéristiques des réseaux de capteurs », ces derniers sont déployés dans certains cas par milliers. Ceci doit être pris en compte dans le développement des algorithmes liés à la gestion des données et ceux liés à la configuration du réseau ;
- **Le matériel** : cette contrainte est liée en partie à la densité du réseau. En d'autres mots, le coût de production des capteurs ne doit pas être trop élevé puisque ces derniers peuvent être déployés par milliers. Donc, optimiser le coût

de production revient à optimiser le coût du matériel et par conséquent nous aurons des capteurs à faible capacité de mémoire, de puissance de calcul et de traitement ;

- **L'énergie** : c'est la contrainte la plus forte dans les réseaux de capteurs sans fils. En effet, dans la plupart des cas, la source d'énergie utilisée est des batteries dont la durée de vie est limitée. Cette contrainte impacte directement la qualité de service.

Toutes ces contraintes rendent difficile la garantie d'une qualité de service optimale surtout quand il s'agit d'applications critiques où la tolérance aux fautes et l'imprécision de mesures par exemple ne sont pas acceptées. À ces contraintes, s'ajoute l'aspect gestion de données provenant de ces capteurs.

1.7 Défis des RCSFs

L'évolution du nombre d'applications dans les RCSFs a donné naissance à de nouvelles problématiques qui nécessitent d'être résolues sur le plan recherche et pratique. Ces problématiques concernent plusieurs aspects. En effet, il n'y a aucun standard spécifique aux RCSFs, les problèmes de communications, de traitement des données et de la gestion du capteur lui-même, sont partiellement identifiés et les solutions apportées sont souvent spécifiques à un cas précis. Les principaux défis lancés par la communauté des chercheurs se rapportent essentiellement aux problèmes de : la découverte du réseau, le routage, la couverture, la durée de vie, la sécurité,... Par ailleurs, les caractéristiques intrinsèques des RCSFs, telles que la capacité limitée des batteries, les moyens limités de calcul et de traitement des données, diffèrent de celles d'un réseau TCP/IP traditionnel et des réseaux ad hoc en général[17].

Les axes de recherches dans les RCSFs sont passés des problèmes liés au routage des données et leur sécurité aux problèmes beaucoup plus pointus concernant la durée de vie des RCSFs, améliorer leur qualité de service, augmenter leur tolérance aux pannes. D'où l'apparition des travaux sur les réseaux de capteurs dans des domaines très spécialisés et qui traitent des problématiques qui n'ont pas été abordés auparavant comme l'utilisation

de la redondance pour augmenter la durée de vie de ces réseaux en mettant en veille les capteurs à tour de rôle suivant des politiques et des algorithmes d'ordonnancement d'activités. On peut classer les travaux de recherche dans le domaine des réseaux de capteurs comme suit :

- **Conception des capteurs** : Trouver de nouvelles conceptions adéquates et spécifiques pour des environnements donnés (détecteur de chute, vêtement intelligent, capteurs multimédias ...).
- **Protocoles de communication** : Trouver de nouveaux paradigmes et protocoles de communication qui prennent en compte l'évitement de collisions, les vides, les flux multimédias, le routage des données
- **Passage à l'échelle** Afin de garantir le bon fonctionnement du réseau, les nouveaux schémas de déploiement doivent être capables de travailler avec un grand nombre de nœuds. Par ailleurs, ils doivent utiliser la propriété de haute densité dans les réseaux de capteurs, et donc pouvoir déployer un grand nombre de nœuds dans une petite surface surtout avec l'avènement de l'internet des objets où le nombre de capteurs pourra atteindre plus de 10 capteurs par mètre carré.
- **Préservation de l'énergie et optimisation** : Trouver de nouvelles visions d'optimisations.
- **Sécurité** : Dans les applications sensibles, la sécurisation des données est nécessaire. D'où, il faudrait tenir compte des ressources limitées des capteurs pour proposer des solutions légères en termes de calcul et de stockage.
- **Conception de middleware** : Nouvelles plateformes dédiées aux RCSFs.

1.8 Conclusion

À travers ce chapitre, nous avons essayé de présenter les réseaux de capteurs sans fil, nous avons posé les briques de base sur les concepts généraux qui aident à la compréhension des réseaux de capteurs.

Dans le chapitre suivant, nous allons présenter les différents protocoles de routage

conçus pour les RCSFs, définir le clustering et ses méthodes ainsi que les outils matériels et logiciels qui permettent la mise en place de certains protocoles pour les RCSFs.

CHAPITRE 2

PROTOCOLES DE ROUTAGE POUR LES RCSFS

2.1 Introduction

Pour concevoir et réaliser un réseau de capteur sans fil, il nous faut un ensemble d'outils logiciels et matériels. Le matériel c'est l'ensemble de capteurs constituant le réseau et le l'outil logiciel c'est l'ensemble de protocoles interagissent entre ces capteurs.

Dans ce chapitre, nous allons présenter les différents protocoles de routage et quelques méthodes de clustering qui sont utilisées pour organiser les nœuds du réseau en clusters. L'approche de clustering permet le passage à l'échelle et d'améliorer la durée de vie des réseaux.

2.2 Protocoles de routage

2.2.1 Définition du routage

Le routage consiste à trouver le meilleur chemin pour envoyer le message de la source à la destination. Dans le cadre des RCSFs, il doit être efficace en termes d'énergie puisque les nœuds dans ce type de réseaux sont généralement dotés de batteries dont

la durée de vie est limitée. Pour cela, il faut bien sûr être capable de trouver un chemin qui ne consomme pas trop d'énergie (économe en énergie), et qui ne soit pas trop long également[18]. Toute conception de protocole de routage doit prendre en considération les contraintes suivantes :

1. La minimisation de la charge du réseau.
2. Offrir un support pour pouvoir effectuer des communications multipoints fiables.
3. Assurer un routage presque optimal.
4. Offrir une bonne qualité de service en termes de latence et permettre la tolérance aux pannes pour garantir l'acheminement des données quand certains nœuds cessent de fonctionner.

2.2.2 Présentation des différents types de routage

Les protocoles de routage sont en fait découpés en quatre familles : les protocoles de routage à plat, hiérarchiques, non hiérarchiques, et géographiques.

Le routage à plat

Le routage à plat et données centrales est le modèle le plus simple où chaque nœud dans le réseau transmet les données à la station de base. Chaque nœud joue typiquement le même rôle et les nœuds capteurs collaborent entre eux pour accomplir la même tâche. La station de base envoie des requêtes à certaines régions et se met en attente des données des capteurs situés dans les régions choisies. Dans ce qui suit quelques protocoles qui sont fondés sur cette architecture :

- a) **SPIN (Sensor Protocol for Information via Negotiation)** SPIN utilise trois types de messages, ADV, REQ et DATA. Le message ADV est diffusé par un nœud qui possède des données et compte les envoyer. Ce message indique le type de données à envoyer par le nœud source. Les nœuds concernés qui ont reçu le message ADV envoient un message REQ demandant les données. Le nœud ayant les données envoie les données aux nœuds concernés. Après la réception des données, les nœuds envoient un message ADV et le processus se poursuit. Ceci

peut être vu dans la figure ci-dessous. Le nœud 1 envoie un message ADV à tous

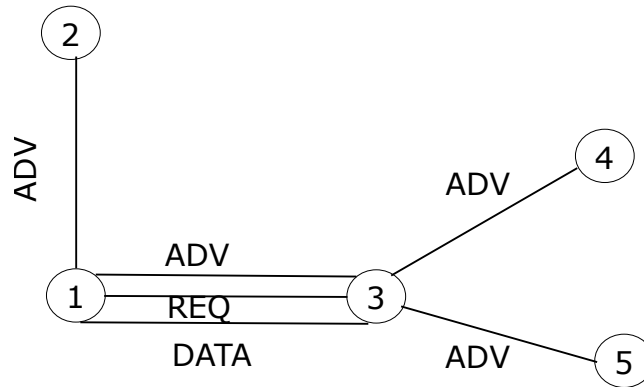


FIGURE 2.1 – Protocole SPIN

ses voisins, 2 et 3. Le nœud 3 demande les données à l'aide du message REQ, pour lequel le nœud 1 envoie les données à l'aide du message DATA au nœud 3. Après avoir reçu les données, le nœud 3 envoie le message ADV à ses voisins 4 et 5 et le processus continue. Il n'envoie pas à 1, car 3 sait qu'il a reçu des données de 1.

Les données sont décrites dans le paquet ADV à l'aide de descripteurs de données de haut niveau, qui permettent d'identifier les données. Ces descripteurs de données de haut niveau sont appelés métadonnées. Les métadonnées de deux données différentes doivent être différentes et les métadonnées de deux données similaires doivent être similaires. L'utilisation de métadonnées évite que les données réelles ne soient inondées via le réseau. Les données réelles ne peuvent être transmises qu'aux nœuds qui en ont besoin. Ce protocole rend également les nœuds plus intelligents, chaque nœud aura un gestionnaire de ressources, qui informera chaque nœud de la quantité de ressources restantes dans le nœud. En conséquence, le nœud peut décider s'il peut ou non être un nœud de transfert.

- b) **Protocole DD (Directed Diffusion)** Dans la diffusion dirigée, les nœuds sont identifiés par leurs extrémités, et la communication inter-nœud se superpose à un service de livraison de bout en bout fourni dans le réseau.

En diffusion dirigée, les nœuds du réseau sont sensibles aux applications, car nous permettons à un code spécifique à une application de s'exécuter sur le réseau et ai-

dans la diffusion dans le traitement des messages. Cela permet à la diffusion dirigée de mettre en cache et de traiter les données sur le réseau (agrégation), en diminuant la quantité de trafic de bout en bout et d'entraîner des économies d'énergie plus importantes. La diffusion dirigée [19] est un protocole basé sur une requête dans lequel une requête est inondée dans le réseau par le récepteur où plusieurs chemins sont établis entre le nœud destinataire et la source. Le nœud destinataire renforce l'un des chemins et reçoit les données dans un intervalle plus court via ce chemin renforcé. Les auteurs de [19] ont modifié le modèle d'énergie radioélectrique du simulateur NS-2, conçu à l'origine pour la radio 802.11, afin d'analyser la consommation d'énergie du réseau de capteurs à l'aide du protocole de diffusion dirigée. Il est montré que la diffusion dirigée pourrait économiser de l'énergie en choisissant le bon chemin c'est-à-dire le chemin économe en énergie. Néanmoins, la diffusion dirigée ne peut pas être appliquée aux applications nécessitant une transmission constante des données telles que les applications de surveillance. Le routage sensible à l'énergie [20] soutient que l'utilisation du même chemin d'énergie minimale épuise les nœuds de ce chemin d'énergie de ce fait il est judicieux de choisir de manière probabiliste entre différents chemins existants entre la source et la destination.

Le routage basé sur la localisation

Dans ce type de routage, les nœuds capteurs sont adressés en fonction de leurs localisations. La distance entre les nœuds voisins peut être estimée en fonction de la puissance du signal reçu (RSSI)¹. les coordonnées relatives des nœuds voisins peuvent être obtenues en échangeant une telle information entre les voisins. Alternativement, la localisation des nœuds peut être disponible directement en communiquant avec un satellite en utilisant GPS (système de positionnement global) ou par le biais d'un nœud mobile diffusant à plusieurs endroits sa localisation qui pourra être captée par les nœuds capteurs se trouvant dans rayon de transmission.

Dans la plupart des protocoles de routage, l'information sur la localisation des nœuds

1. Received Signal Strength Indication

est nécessaire afin de calculer la distance entre deux nœuds particuliers de sorte que la consommation d'énergie puisse être estimée[21].

Le routage hiérarchique

Le routage hiérarchique est considéré comme étant l'approche la plus favorable en termes d'efficacité énergétique surtout dans les réseaux de capteurs à grande échelle. Il se base sur le concept (nœud simple - nœud maître) où les nœuds simples acheminent leurs messages à leur maître, lequel les achemine ensuite dans le réseau tout entier via d'autres nœuds maîtres jusqu'à l'aboutissement à la station de base.

L'avantage principal ce type de routage est l'agrégation et la fusion des données afin de diminuer le nombre de messages redondants transmis à la station de base, ce qui permet une meilleure conservation de l'énergie. En fait, deux grandes approches sont dérivées de ce type de routage à savoir : chaîne-based approach (approche chaînée) comme PEGASIS et cluster-based approach (approche à grappe) comme LEACH[22]

- a) **HEED (Hybrid, Energy-Efficient, Distributed approach)** Le protocole hybride HEED (Hybrid Energy Efficiency Distributed) est l'un des protocoles qui permettent d'augmenter la durée de vie dans les RCSFs. Ce protocole prend en compte deux paramètres pour déterminer les cluster-heads (chefs de groupes), à savoir l'énergie résiduelle et la densité des nœuds, et a été conçu pour les RCSFs homogènes.

Dans HEED, les cluster-heads communiquent directement avec la station de base ce qui permet de minimiser la consommation d'énergie. Cependant, dans HEED, la topologie en clusters ne réalise pas la consommation minimale d'énergie dans les communications intra-cluster et les clusters générés ne sont pas tellement équilibrés en taille[23].

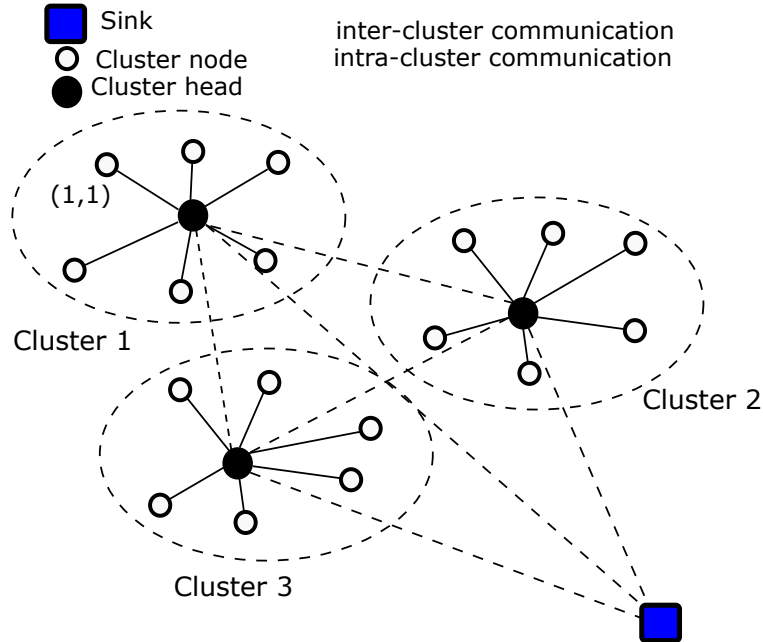


FIGURE 2.2 – Protocole LEACH avec CH et BS.

b) **LEACH (Low Energy Adaptive Clustering Hierarchy)** LEACH est un protocole conçu pour les réseaux de capteurs où un utilisateur final souhaite surveiller l'environnement à distance. Dans une telle situation, les données des nœuds individuels doivent être envoyées à une station de base, souvent située loin du réseau de capteurs, par laquelle l'utilisateur final peut accéder aux données[24]. Il existe plusieurs propriétés souhaitables pour les protocoles conçus aux RCSFs :

- Supportent des réseaux de taille moyenne et à grande échelle de 100 à 1000 nœuds,
- Maximisent la durée de vie des réseaux,
- Maximisent la couverture de zone,
- Utilisent des nœuds uniformes de bon marché

LEACH permet deux types de communications : intra-cluster et inter-cluster pour réduire la consommation d'énergie et la rotation aléatoire des cluster-heads pour ne pas épuiser certains nœuds de leurs batteries rapidement. Ces fonctionnalités permettent à LEACH d'atteindre les propriétés souhaitées. Les études sur ce

protocole montrent que LEACH est un protocole économe en énergie qui prolonge la durée de vie des RCSFs.

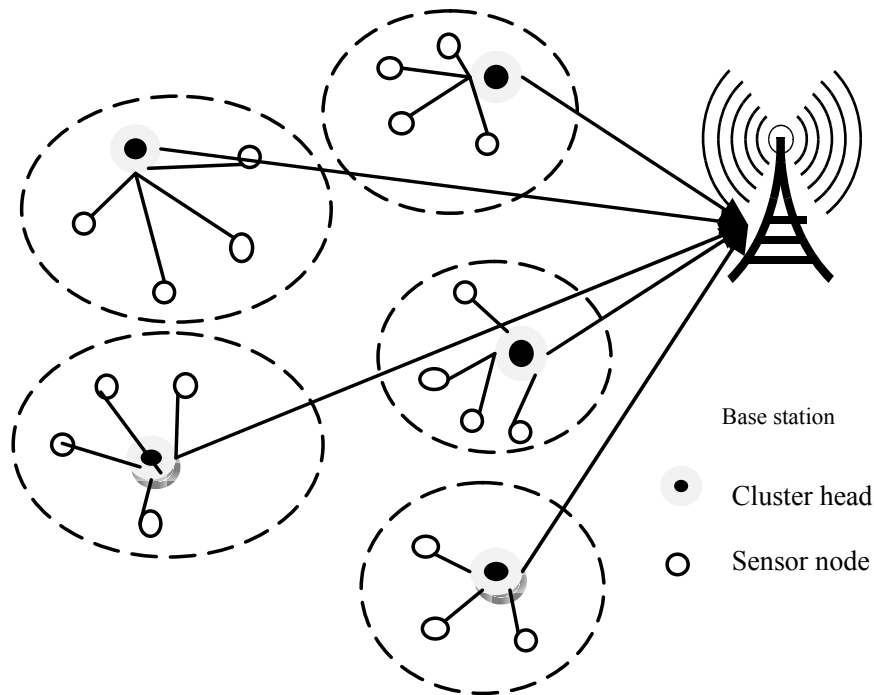


FIGURE 2.3 – Protocole LEACH avec CH et BS.

c) **PEGASIS (Power-Efficient Gathering in Sensor Information Systems)**

PEGASIS est un protocole hiérarchique basé sur la chaîne la plus favorisée. Dans ce protocole, les nœuds sont organisés sous la forme d'une chaîne pour le transport et l'agrégation des données. La création de chaînes peut être centralisée en fonction de l'application. PEGASIS est basé sur l'hypothèse qu'une connaissance globale du réseau est fournie à tous les nœuds. La création de la chaîne commence à partir du dernier nœud du récepteur et son voisin le plus proche est sélectionné en tant que nœud suivant de la chaîne, etc. Le dernier nœud doit être le nœud destinataire et le nœud avant que celui-ci agisse en tant que leader du nœud. Le traitement des données et l'agrégation sont effectués par le nœud principal. PEGASIS n'est pas aussi pertinent pour les réseaux à topologie dynamique ou variant dans le temps. Comme la taille du réseau sera plus grande, le délai de transmission sera d'autant plus long que PEGASIS est soumis à un problème d'évolutivité[25].

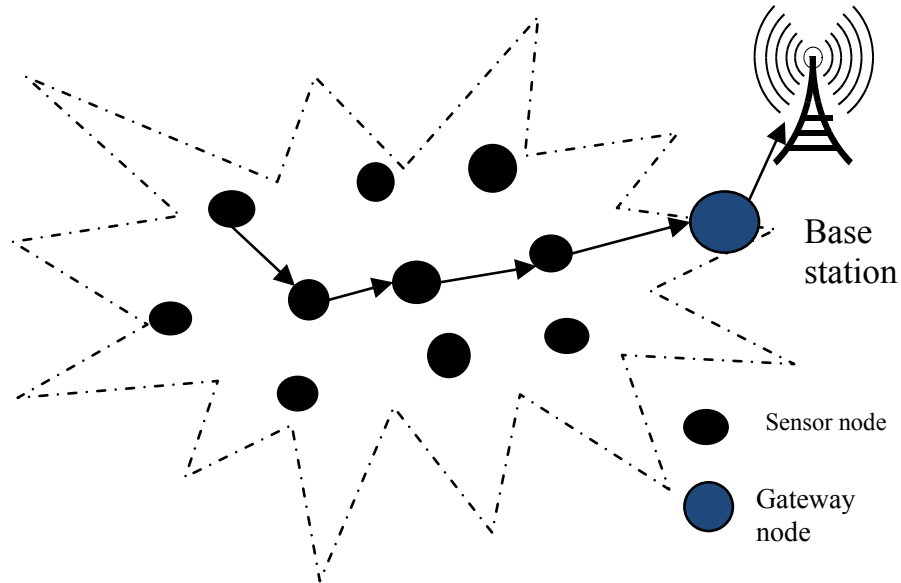


FIGURE 2.4 – Fonctionnement du protocole PEGASIS

2.3 Travaux antérieurs

Il existe plusieurs protocoles qui ont pour but de maximiser la durée de vie d'un réseau de capteurs, dans cette section nous nous sommes intéressés sur les protocoles de routage dans l'opération de formation de clusters est basée sur l'approche K-Means.

Dans[26], les auteurs ont proposé une approche K-Means distribuée dans laquelle le clustering est effectué par chaque nœud du réseau. Cette contribution est plus efficace que le clustering centralisé en termes de la vitesse du traitement, cependant la consommation d'énergie est presque la même dans les deux approches.

Dans[27], les auteurs ont proposé une approche de clustering basée sur K-Means pour minimiser la consommation d'énergie dans le réseau. Cette approche est adaptée pour les réseaux de stockage intégrés appelés KEAC. Dans KEAC, les cluster-heads sont élus en fonction de l'énergie restante et de la distance qui les sépare de leurs centres de gravité. Les CHs générés sont uniformément distribués dans le réseau en fonction de leurs poids et permettent une communication directe avec la station de base (un seul saut). À cet effet, la minimisation de la consommation d'énergie mène à prolonger la durée de vie du réseau, cependant l'ajustement de la radio pour communiquer avec la

station de base peut augmenter les frais de ce protocole

Dans[28], les auteurs ont proposé un schéma de routage hybride qui combine entre l'approche K- Means et le protocole LEACH. K-Means est utilisée pour regrouper les nœuds capteurs en clusters et LEACH est utilisé pour sélectionner les cluster-heads. Les résultats de simulation ont montré que ce protocole surpasse LEACH en termes de la consommation d'énergie et de la durée de vie du réseau, mais la principale faiblesse de ce protocole est la grande énergie consommée durant le processus d'élection des cluster-heads.

2.4 Clustering et partitionnement de données

2.4.1 Définition

Le clustering est un processus qui regroupe un ensemble d'objets (physiques ou abstraits) en clusters similaires de telle sorte que les données du même cluster aient des caractéristiques similaires, et celles appartenant à des clusters distincts soient dissimilaires.

Le clustering est une des méthodes d'analyse des données. Elle vise à diviser un ensemble de données en différents « paquets » homogènes, en ce sens que les données de chaque sous-ensemble partagent des caractéristiques communes, qui correspondent le plus souvent à des critères de proximité (similarité informatique) que l'on définit en introduisant des mesures et classes de distance entre objets[29].

2.4.2 Clustering basé sur une représentation non supervisée

L'approche K-Means

L'approche K-Means[30]est utilisée pour regrouper les éléments d'un ensemble de données en k clusters autour d'un centre de gravité (centroid). En général on ne connaît pas le nombre de classes que contient l'ensemble de données. Étant donné un ensemble

de points $\{x_1, x_2, \dots, x_n\}$ on cherche à partitionner les n points en k ensembles :

$$S = \{S_1, S_2, \dots, S_k\} (k \leq n)$$

en minimisant la distance entre les points à l'intérieur de chaque partition :

$$\arg \min_S \sum_{i=1}^n \sum_{x_j \in S_j} \|x_j - u_i\|^2 \quad (2.1)$$

où u_i est le barycentre des points dans S_i la méthode des k-means se déroule comme suit :

1. On tire au hasard k centres de gravité. Ces centres peuvent être tirés parmi les exemples de la base d'apprentissage.
2. On associe chaque exemple de l'ensemble de données au centre de gravité le plus proche. Après cette étape tous les exemples ont été affectés à un centre.
3. Chaque centre est mis à jour à l'aide de la moyenne des exemples qui lui sont associés.
4. Puis on recommence les étapes 2 et 3 jusqu'à ce que les exemples affectés à un centre ne changent plus.

Il existe une variante de cette méthode, qui s'appelle le k-means++ [31]. Cette variante est une amélioration de la version originale de K-means et consiste à changer la manière d'initialiser les centres de gravité.

K-médianes

L'algorithme des k-médianes est similaire à l'algorithme des k-means en utilisant la médiane au lieu de la moyenne. Cet algorithme est souvent confondu avec celui des k-médoïdes (aussi appelé k-médianes discrètes). On définit la médiane (géométrique) d'un ensemble de points comme le point artificiel qui minimise :

$$\sum_{x \in X} \|c - x\|^2 \quad (2.2)$$

Avec x qui décrit les points de l'ensemble

Il n'y a aucune formule analytique pour calculer la médiane géométrique exacte à partir d'un ensemble de points. Par conséquent en pratique on cherche à approximer cette médiane. L'algorithme des k-médianes semble être le plus performant pour trouver une excellente solution au problème de partitionnement en k clusters, en revanche sa complexité est vraiment très élevée et il ne peut pas être appliqué sur de grandes bases de données.

La méthode des k-médianes s'exécute ainsi :

1. On initialise k médianes (par exemple aléatoirement)
2. On associe chaque élément de l'ensemble de données à la médiane la plus proche de lui.
3. Puis pour chaque ensemble associé à une médiane, on recalcule une approximation de sa médiane.
4. On recommence les étapes 2 et 3 à nouveau jusqu'à ce que les points associés à une médiane ne changent plus.

K-modes

La méthode des k-modes est similaire à celle des k-moyennes, mais adaptée pour les objets catégorisés qui sont des objets qui ne contiennent pas de valeurs numériques, mais des chaînes par exemple[32] Le principe de la méthode des k-modes se déroule ainsi :

1. On initialise k-modes (aléatoirement par exemple).
2. On associe chaque objet de l'ensemble de données au mode le plus proche ou le plus similaire à lui.
3. Puis on recalcule les modes de chaque ensemble, à partir de la fréquence des champs des objets.
4. Et on recommence les étapes 2 et 3, jusqu'à ce que les modes ne changent plus.

2.4.3 Méthodes de clustering

Les méthodes de clustering sont généralement classifiées en quatre catégories majeures :

Les méthodes hiérarchiques

Dans un clustering hiérarchique, un cluster peut être divisé en sous clusters. L'ensemble des clusters étant généralement représenté par un arbre. Un objet appartient à une et une seule feuille dans la hiérarchie, mais également à son nœud père, et ainsi de suite jusqu'à la racine. Les méthodes de clustering hiérarchique permettent d'obtenir ce type de résultats.

Il existe deux types d'approches de clustering hiérarchique :

- Les approches par agglomération (ou ascendantes).
- Les approches par division (ou descendantes).

Les méthodes de partitionnement

Les méthodes de partitionnement ont généralement comme résultat un ensemble de M clusters, chaque objet appartenant à un seul cluster. Chaque cluster peut être représenté par un centroïde (représentant du cluster) qui peut être considéré comme une description récapitulative de tous les objets contenus dans le cluster. La forme précise de cette description dépendra du type des objets qui sont groupés. Au cas où les données à valeurs réelles sont disponibles, la moyenne arithmétique des vecteurs d'attribut pour tous les objets dans un cluster fournit un représentant approprié ; des types alternatifs de centroïdes peuvent être requis dans d'autres cas.

Si le nombre de clusters est élevé, les centroïdes peuvent encore être groupés de manière hiérarchique.

Les méthodes basées sur la densité

Les algorithmes basés sur la densité sont capables de découvrir des clusters de formes arbitraires, ce qui assure l'isolement des bruits et la prévention contre la formation de

clusters non pertinents. Ces algorithmes regroupent des objets selon des fonctions de densité spécifiques. La densité est habituellement définie comme nombre d'objets dans un voisinage particulier des éléments de données. Dans cette approche, un cluster donné continue à augmenter de taille tant que le nombre d'objets dans le voisinage dépasse un certain seuil.

Les méthodes basées sur la grille

Un algorithme de clustering basé sur les grilles utilise des structures de données multi résolution, où l'espace d'objets est quantifié en un ensemble de cellules, puis identifie l'ensemble de cellules denses connectées pour former des clusters.

2.4.4 Domaines d'applications du clustering

Le clustering possède des domaines d'applications extrêmement variés, parmi lesquels :

- **Le Marketing** : segmentation du marché en découvrant des groupes de clients distincts à partir de bases de données d'achats.
- **Le Marketing** : segmentation du marché en découvrant des groupes de clients distincts à partir de bases de données d'achats.
- **Les assurances** : identification de groupes d'assurés distincts associés à un nombre important de déclarations.
- **La planification des villes** : identification de groupes d'habitations suivant leurs types, valeur, localisation géographique.
- **La médecine** : Localisation de tumeurs dans le corps humain. Par exemple, dans un nuage de points fournis par le scan du cerveau, on identifie les points définissant une tumeur.
- **La segmentation d'images** : Détection des zones homogènes dans une image.
- **Text mining** : Classification des textes selon leur similitude dans des dossiers automatiques.

2.4.5 Types de transmission dans le routage

La maximisation de la durée de vie d'un réseau de capteur dépend fortement de certaines solutions de routage. Dans cette partie nous présentons quelques schémas de routage utilisant le multi saut. Dans[33], les auteurs ont proposé deux versions améliorées de LEACH : Energy-LEACH et Multihop-LEACH. Le protocole Energy-LEACH permet de choisir comme cluster-heads les nœuds qui ont plus d'énergie résiduelle, cependant le protocole multihop-LEACH permet d'améliorer la communication entre les CHs et entre les CHs et la station de base.

Dans[34], les auteurs ont proposé deux protocoles LEACH-D et LEACH-L. Ces deux protocoles travaillent en boucle, et chaque tour est divisé en phase d'établissement et en phase de stabilisation, dans LEACH-L, les CHs fusionnent les données provenant des nœuds membres. Si les CHs sont proches de la station de base, ils envoient les données directement à la station de base, sinon ils optent pour une stratégie multi sauts dans laquelle le prochain saut est choisi en fonction de son énergie et de la distance qui le sépare de la station de base.

2.4.6 Greedy forwarding

La méthode "Greedy Forwarding" est une stratégie qui utilise la distance ou la direction des nœuds comme paramètres pour choisir le prochain saut du chemin de la source vers la destination. Greedy Forwarding basée sur la distance consiste à minimiser le nombre de sauts, car chaque cluster-head envoie la donnée au voisin le plus proche à la station de base jusqu'à ce que la donnée atteigne envoyée à la station de base.

La figure 2.5 représente un exemple de la stratégie Greedy Forwarding basée sur la distance où RC est la plage de transmission du capteur S. Chaque fois que le nœud S a un paquet de données à envoyer à la station de base, il choisit le voisin le plus proche à la station de base comme prochain saut (le nœud B dans cet exemple). Ce processus de routage sera répété de la même manière jusqu'à ce que le paquet de données atteigne la station de base.

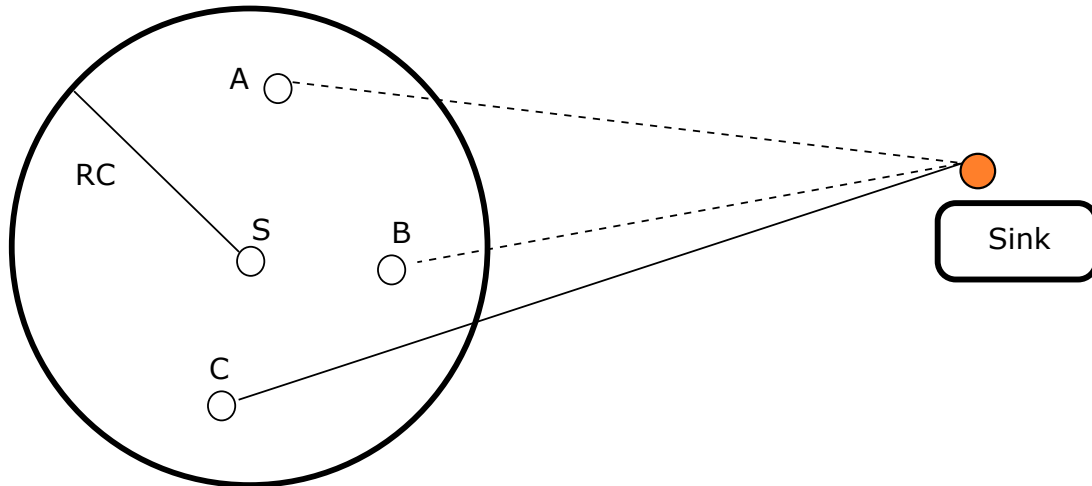


FIGURE 2.5 – Greedy Forwarding basée sur la distance

2.5 Conclusion

Dans ce chapitre, on a présenté les types de routage dans les RCSFs et les protocoles de routage les plus répandus associés à ces types. On a détaillé certaines méthodes de partitionnement telles que K-means et enfin les stratégies de routage dans les RCSFs

CHAPITRE 3

SCHÉMA DE ROUTAGE BASÉ SUR LE CLUSTERING POUR LES RCSFS

3.1 Introduction

Après avoir étudié les techniques de routage dans les réseaux de capteur sans fil , on va proposer une nouvelle technique basée sur le clustering, pour faire face au passage à l'échelle et pour réduire la consommation de l'énergie. Cette technique concerne une adaptation d'une nouvelle méthode de clustering aux réseaux de capteurs sans fil permettant d'éviter de définir le nombre de clusters comme paramètre initial contrairement à K-means.

3.2 K-Means

K-Means [35] est l'un des algorithmes d'apprentissage non supervisé les plus simples permettant de résoudre le problème de clustering. La procédure utilisée par K-Means suit un moyen simple pour classer un ensemble de données avec un nombre de clusters donné. L'idée principale est de définir un centroïde pour chaque cluster. L'étape suivante consiste à prendre chaque point appartenant à un ensemble de données et l'associer au centroïde le plus proche. Lorsqu'aucun point n'est en attente, la première étape est

terminée et les clusters sont formés. Ensuite, les k nouveaux centroïdes des clusters qui ont été déjà formés précédemment sont recalculés. Une fois que les k nouveaux centroïdes sont recalculés, une nouvelle liaison doit être effectuée entre les mêmes points et le nouveau centroïde le plus proche. Une boucle est générée jusqu'à ce que les k centroïdes ne changent plus d'emplacement, c'est-à-dire que les centroïdes ne bougent plus. Enfin, cette technique vise à minimiser la fonction objective suivante :

$$J = \sum_{j=0}^K \sum_{i=0}^n \|X_i^j - C_j\|^2 \quad (3.1)$$

Où $\|X_i^j - C_j\|^2$ est une mesure de distance choisie entre un point de données x et le centre du groupe C_j . Ceci est un indicateur de la distance entre les n points de données et leurs centres de cluster respectifs. K-means est composé des étapes suivantes :

1. Placer K points dans l'espace représenté par des objets. Ces points représentent les centroïdes du groupe initial.
2. Affecter chaque objet au groupe qui a le centroïde le plus proche.
3. Lorsque tous les objets ont été affectés, recalculer les positions des centroïdes K .
4. Répéter les étapes 2 et 3 jusqu'à ce que les centroïdes ne disparaissent plus. Cela produit une séparation des objets en groupes à partir de laquelle la métrique à minimiser peut être calculée.

3.3 Méthode silhouette [36]

La méthode Silhouette fournit un moyen quantitatif pour mesurer la position de chaque élément dans son cluster par rapport aux autres. La moyenne de la méthode silhouette de chaque point est une mesure de la similitude d'un point avec les points de son propre cluster par rapport aux points de départ des autres clusters. La valeur Silhouette d'un point de données est définie comme suit :

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (3.2)$$

Où, $a(i)$ est la moyenne de la distance entre le i^e point et les autres points du même cluster. Soit $b(i, k)$ la moyenne de la distance entre le i^e point et les autres points du

cluster. Alors $b(i)$ est le minimum de $b(i, k)$ sur tous les clusters auxquels le point de repère n'est pas attribué. La valeur de $S(i)$ n'est pas supérieure à 1, et si $S(i)$ est proche de 1, cela indique que le point concerné est bien dans son propre cluster. La méthode Silhouette est composée des étapes suivantes :

- Exécuter K-means en faisant varier le nombre de clusters de 1 à max ,
- Pour chaque k , calculer la moyenne de la méthode silhouette des observations,
- Tracer la courbe de avg selon le nombre de clusters k ,
- L'emplacement du maximum est considéré comme le nombre approprié de clusters.

3.4 Méthode de k-means distribuée dans les réseaux de capteurs

Les auteurs dans [35] ont proposé un nouveau schéma appelé K-means distribués dans lequel le processus de clustering est exécuté de manière distribuée dont le processus d'élection des cluster-heads est basé sur la distance par rapport aux centroïdes et l'énergie résiduelle comme illustrée par l'équation (3.3). Le schéma de routage proposé se déroule en deux phases : la phase Set-up et la phase steady-state. Dans la phase set-up la méthode K-means est utilisée pour la formation des clusters durant la première période et l'élection des cluster-heads se fait de manière distribuée. Chaque nœud envoie son coût à tous les membres de son cluster. Le nœud dont le coût est le plus élevé devient CH en tenant compte de l'énergie résiduelle et de la distance par rapport aux centroïdes selon l'équation suivante :

$$Cost(Noeud_i) = \alpha * Energie + \frac{\beta}{dist(Noeud_i, Centroid)} \quad (3.3)$$

Dans la deuxième phase, chaque membre du cluster envoie les données collectées à son CH respectif en fonction de son slot TDMA. Une fois que tous les paquets ont été reçus par le CH, ce dernier les regroupe dans un seul paquet et le transmet vers la station de base. Ce schéma de regroupement basé sur l'approche K-means offre de meilleures performances que le protocole LEACH.

3.5 Principe de fonctionnement du schéma de routage proposé

La solution de routage proposée permet de former des clusters en se basant sur la distance entre les nœuds en tenant compte du nombre des nœuds par cluster, tel que ce nombre dépasse un certain seuil. Ensuite, le nœud qui a l'énergie maximale et la meilleure position pour chaque cluster est élu comme cluster-head. En outre, chaque cluster-head identifie le prochain saut vers la station de base pour construire son chemin basé sur Greedy Forwarding et commence la transmission des données.

3.5.1 Avantages

- La transmission multi-saut inter-cluster a un avantage par rapport à la transmission directe à la station de base (en un seul saut).
- La méthode de clustering utilisée permet de garantir le passage à l'échelle.
- La sélection des clusters-heads n'est pas fixée et elle est déclenchée par une condition.
- L'élection du cluster-head est indépendante de la station de base.
- L'élection des cluster-heads a une efficacité d'élargir la durée de vie du réseau.

3.5.2 Inconvénients

Dans cette approche de routage proposée, les nœuds sont toujours affiliés à un même cluster et le processus de clustering est statique.

3.5.3 Schéma de routage proposé

Le schéma de routage hiérarchique proposé se déroule en rounds et dans chaque round deux étapes : une étape pour le clustering et une autre pour la construction des chemins

1. La première étape vise à organiser le réseau en clusters et se divise en deux phases : la formation des clusters et l'élection des cluster-heads.

Formation des clusters

Au début, la station de base lance le premier round. Chaque nœud calcule la distance moyenne par rapport aux autres nœuds. Ensuite les clusters sont formés en fonction la distance moyenne et chaque cluster va contenir les nœuds les plus rapprochés. Pour éviter le problème de clusters non homogènes en termes de nombre de nœuds par cluster à d'autres clusters, nous allons enlever les nœuds appartenant à des clusters qui ont un nombre de nœuds inférieurs à un certain seuil. Puis, nous associons chaque nœud restant au cluster le plus proche en se basant sur la distance entre le nœud et les centroïdes des clusters.

L'élection des cluster-heads

Chaque nœud vivant va calculer son score en fonction de la distance moyenne et de l'énergie restante du nœud selon l'équation suivante :

$$Cost(Noeud_i) = \alpha * Energie_{Res}(Noeud_i) + \frac{\beta}{dist_{Moy}(Noeud_i)} \quad (3.4)$$

Avec :

$Energie_{Res}$ l'énergie restante de chaque noeud.

La distance moyenne $dist_{Moy}$:c'est la moyenne des distances de chaque nœud par rapport aux autres nœuds dans le même cluster selon l'équation suivante :

$$dist_{Moy}(Noeud_i) = \frac{\sum_{j=1}^n dist(Noeud_i, Moeud_j)}{n - 1} \quad (3.5)$$

Avec :

n :le nombre des nœuds par chaque cluster.

Après, chaque nœud envoie son coût exprimé par l'équation (3.5) aux autres nœuds du même cluster et le nœud ayant le coût le plus élevé va être élu comme cluster-head.

2. La deuxième étape contient deux phases la construction des chemins et l'envoi des données

Construction des chemins

La transmission inter-cluster utilise le mode multi-saut et le choix du prochain saut est fait par une amélioration de la méthode Greedy Forwarding en se basant sur deux facteurs : l'énergie résiduelle du nœud (prochain saut) et la distance qui le sépare à la station de base. En plus des fonctionnalités du cluster-head, il va aussi jouer le rôle de nœud relais. Chaque cluster-head envoie son score défini dans l'équation(3.6) à ses voisins cluster-heads.

$$Cost(Noeud_i) = \alpha * Energie_{Res} + \frac{\beta}{dist(Noeud_i, BS)} \quad (3.6)$$

Chaque cluster-head choisit le cluster-head qui a le plus grand score comme prochain saut. On a favorisé le coefficient de la distance par rapport à l'énergie pour préserver le principe de l'approche "Greedy Forwarding".

L'envoi des données

Une fois que la création des clusters associés et la construction des chemins sont achevées, la transmission des données peut commencer. Chaque nœud envoie ces données à son clusterhead associé en mode mono-saut et chaque cluster-head agrège les données reçues en un seul paquet et l'envoie au cluster-head suivant jusqu'à ce qu'il atteigne la station de base.

3.6 Description du déroulement de l'algorithme

L'algorithme proposé se déroule en périodes et chaque période se compose de deux étapes : a) La première étape pour le clustering qui consiste à la formation des clusters et à l'élection des cluster-heads.

Formation des clusters

La première phase de l'agorithme consiste au premier lieu à construire des clusters selon l'organigramme suivant :

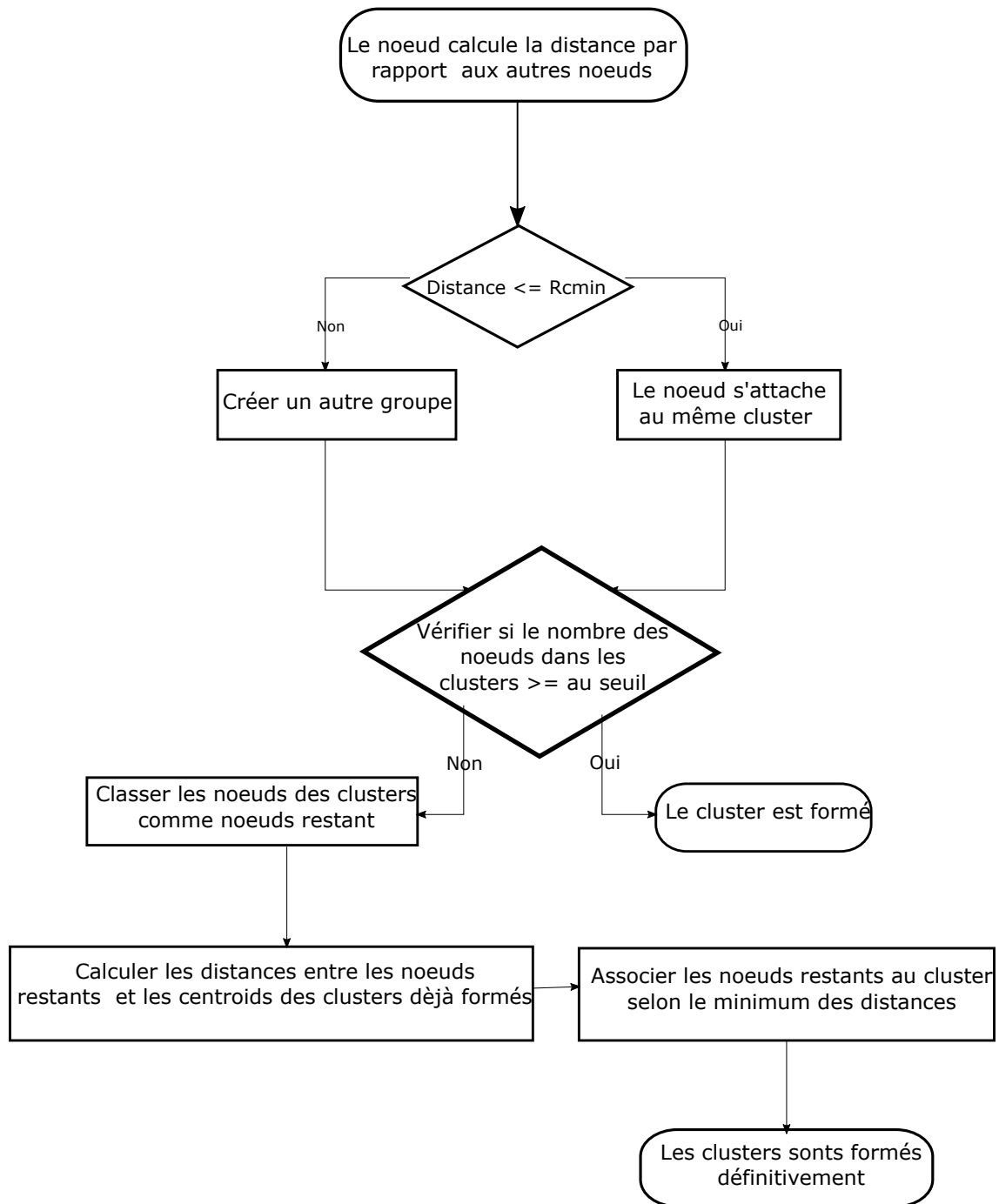


FIGURE 3.1 – Schéma de création des clusters

Algorithme de création des clusters

Début :

Pour tout nœud faire

 Calculer la distance par rapport aux autres nœuds ;

Si distance \leq RCmin alors

 Le nœud s'attache au même cluster ;

Si non

 Créer un autre groupe ;

Fin si

Si nbre_nœud \geq le seuil alors

 Le cluster est formé ;

Si non

 Classer les nœuds des clusters comme nœud restant ;

Fin si

 Calculer les distances entre les nœuds restant et les centroid des cluster déjà formé ;

 Associer le nœud restant au cluster selon le minimum des distances des cluster ;

 Les clusters sont formés définitivement ;

Fin pour

Fin

L'élection du cluster-head

Une fois la construction des clusters est achevée, la création des CHs commence au niveau de chaque cluster selon l'organigramme suivant :

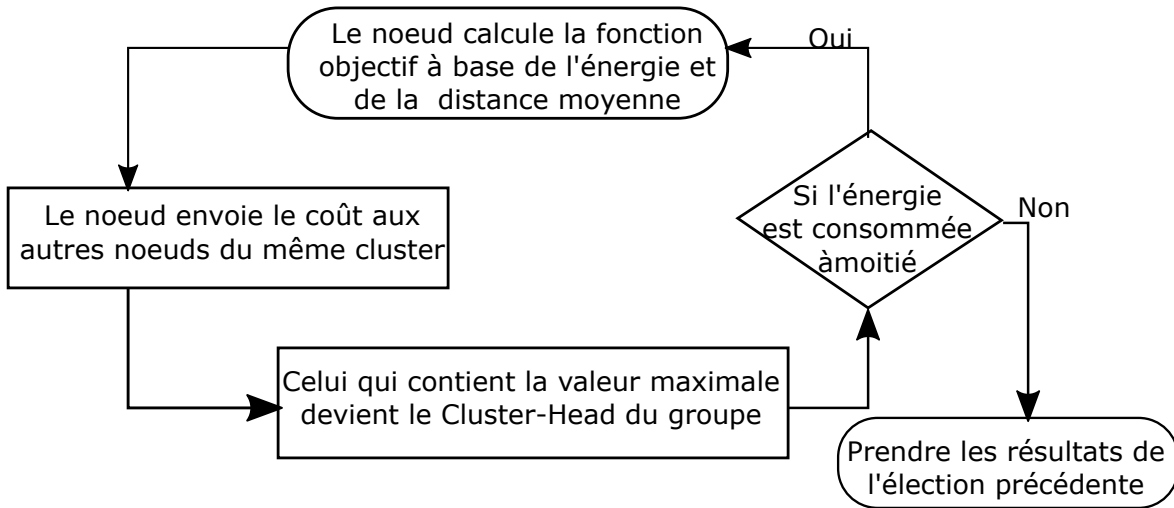


FIGURE 3.2 – Processus d'élection des cluster-heads

b) La deuxième étape concerne la transmission de données qui se compose en deux phases construction des chemins et l'envoi des données

Construction des chemins

Quand le processus d'élection des CHs est terminé, l'étape de la construction des chemins commence, selon l'organigramme suivant :

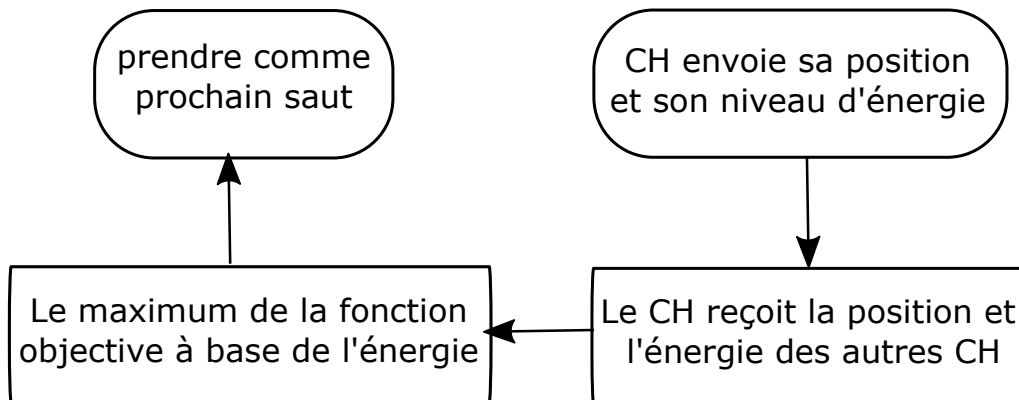


FIGURE 3.3 – Schéma de construction des chemins

L'envoi des données

Cette phase consiste de transmettre les données collectées par les nœuds à leurs cluster-head, ensuite chaque CH transfère ces données au CH voisin ayant le plus grand score (prochain saut) jusqu'à la station de base selon l'organigramme suivant :

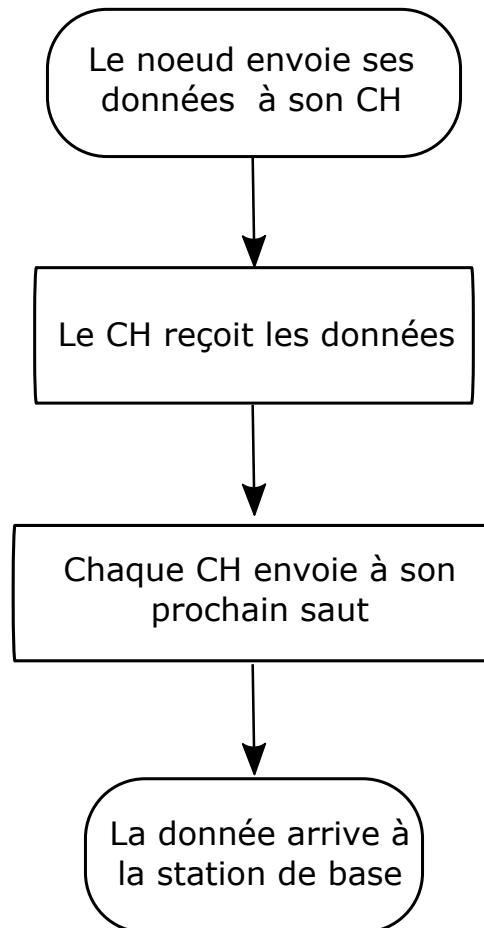


FIGURE 3.4 – Schéma de transmission des données

3.7 Conclusion

Dans ce chapitre, un algorithme de routage hiérarchique pour les réseaux de capteurs sans fil a été proposé.

Au cours des travaux à venir, nous aurons tendance à optimiser les solutions de routage proposées afin d'offrir de meilleures performances.

CONCLUSION GÉNÉRALE

Les réseaux de capteurs sans fil sont des réseaux ad-hoc composés de dispositifs minuscules qui ont des capacités de stockage, de calcul, ils doivent leur essor fulgurant à la panoplie d'applications offertes par cette technologie née de l'évolution vers la miniaturisation de plus en plus poussée des composants. Les perspectives applicatives offertes par les réseaux de capteurs sans fil sont nombreuses et variées. Cependant, la durée de vie des réseaux est encore actuellement un facteur limitant le développement de ces applications. Il est donc nécessaire de concevoir des protocoles de communication et d'auto-organisation permettant de minimiser l'énergie consommée dans le réseau, ainsi que des plateformes matérielles à faible consommation.

Dans le cadre de ce projet, on a proposé un algorithme de clustering ; cet algorithme consiste à réduire la distance entre les nœuds capteurs pour consommer moins d'énergie lors de la transmission des données et par conséquent pour rendre meilleure la durée de vie du réseau. En perspectives, on propose comme amélioration de ce travail l'utilisation du clustering dynamique au lieu du clustering statique pour que les nœuds ne soient pas toujours affectés au même cluster.

BIBLIOGRAPHIE

- [1] Y.Challal, "Réseaux de capteurs sans fils," Cours, Systèmes Intelligents pour le Transport, Université de Technologie de Compiègne, France, vol. 17, 2008
- [2] J.Brignell, "The future of intelligent sensors : A problem of technology or ethics ?," Sensors and Actuators A : Physical, vol. 56, no. 1-2, pp. 11-15, 1996
- [3] H. Schödel, "Utilization of fuzzy techniques in intelligent sensors," Fuzzy Sets and Systems, vol. 63, no. 3, pp. 271-292, 1994
- [4] C.-T. Kone, "Conception de l'architecture d'un réseau de capteurs sans fil de grande dimension," Université Henri Poincaré-Nancy I, 2011
- [5] M. Lehsaini, "Diffusion et couverture basées sur le clustering dans les réseaux de capteurs : application à la domotique," Besançon, 2009
- [6] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks : a survey," Computer networks, vol. 38, no. 4, pp. 393-422, 2002
- [7] S. MOAD, "Optimisation de la consommation d'énergie dans les réseaux de capteurs sans fil " Master recherche en 2^{ème} année informatique, vol. Université : FSICRennes 1, Laboratoire de recherche : DYONISOS-IRISA, 2008
- [8] R. Kacimi, "Techniques de conservation d'énergie pour les réseaux de capteurs sans fil," 2009.
- [9] O. Boudaa, "Conception et réalisation d'un protocole de routage pour les réseaux de capteurs sans fil," Université A/Mira de Bejaia, 2015.

- [10] A. Mehiaoui, "Etude comparative entre les deux protocoles de routage LEACH et PEGASIS dans les réseaux de capteurs sans fil."
- [11] M. Ilyas and I. Mahgoub, Handbook of sensor networks : compact wireless and wired sensing systems. CRC press, 2004
- [12] A. Manjeshwar and D. P. Agrawal, "APTEEN : A hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks," in ipdps, 2002, p. 0195b : Citeseer.
- [13] K. M. Modieginyane, B. B. Letswamotse, R. Malekian, and A. M. Abu-Mahfouz, "Software defined wireless sensor networks application opportunities for efficient network management : A survey," vol. 66, pp.274-287, 2018.
- [14] H. Wu and M. Shahidehpour, "Applications of wireless sensor networks for area coverage in microgrids," IEEE Transactions on Smart Grid, vol. 9, no. 3, pp. 1590-1598, 2018
- [15] L. Iannone, F. Benbadis, M. D. de Amorim, and S. Fdida, "Some applications of wireless sensor networks," in Journées Scientifiques Techniques CetMef (Centre d'Etudes Techniques Maritimes et Fluviales), 2004, pp. 239-247.
- [16] . Karl and A. Willig, Protocols and architectures for wireless sensor networks. John Wiley and Sons, 2007.
- [17] T. Allik, H. Makhlof, and N. Koulalene, "Mise en place d'un réseau de capteurs sans fil pour la surveillance des paramètres de production de l'unité de conditionnement d'huile au sein de Cévital," Université abderrahmane mira béjaia, 2017.
- [18] Y. Yousef, "Routage pour la gestion de l'énergie dans les réseaux de capteurs sans fil, Mulhouse, 2010.
- [19] L. J. Zhong J, "Cognitive radio cognitive network simulator," Michigan Tech University, 2009
- [20] J. Wang, M. Ghosh, and K. Challapali, "Emerging cognitive radio applications : A survey," IEEE Communications Magazine, vol. 49, no. 3, pp. 74-81, 2011
- [21] C. PRODHON, "LE PROBLÈME DE LOCALISATION-ROUTAGE," 2006.

- [22] A. Awasthi, "Développement d'un système de routage hiérarchique pour les réseaux urbains," Université de Metz, 2004
- [23] O. Younis and S. Fahmy, "HEED : a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on mobile computing*, no. 4, pp. 366-379, 2004.
- [24] H. Balakrishnan, A. P. Chandrakasan, and W. B. Heinzelman, "Method for low-energy adaptive clustering hierarchy," ed : Google Patents, 2006
- [25] S. Lindsey and C. S. Raghavendra, "PEGASIS : Power-efficient gathering in sensor information systems," in *Proceedings, IEEE aerospace conference*, 2002, vol. 3, pp. 3-3 : IEEE
- [26] S. K. P. Sasikumar, "K-means clustering in wireless sensor networks," in *Proceedings of 4th International Conference on Computational Intelligence and Communication Networks*, pp. 140-144, 2012
- [27] A. Gachhadar and O. N. Acharya, "K-means based energy aware clustering algorithm in wireless sensor network," *International Journal of Scientific and Engineering Research*, vol. 5, no. 5, pp. 156-161, 2014
- [28] A. Mahboub, M. Arioua, and E. M. En-Naimi, "Energy-Efficient Hybrid K-Means Algorithm for Clustered Wireless Sensor Networks," *International Journal of Electrical and Computer Engineering (2088-8708)*, vol. 7, no. 4, 2017.
- [29] R. Guigourès, "Utilisation des modèles de co-clustering pour l'analyse exploratoire des données," Université Panthéon-Sorbonne-Paris I, 2013
- [30] E. G. N. Gast, "k-Moyennes," Cours L3 math/info.
- [31] D. Arthur and S. Vassilvitskii, "k-means++ : The advantages of careful seeding (Tech. Rep.)," ed : Stanford, CA : Stanford InfoLabo, 2006
- [32] . Chaturvedi, P. E. Green, and J. D. Carroll, "K-modes clustering," *Journal of classification*, vol. 18, no. 1, pp. 35-55, 2001.
- [33] F. Xiangning and S. Yulin, "Improvement on LEACH protocol of wireless sensor network," in *2007 International Conference on Sensor Technologies and Applications (SENSORCOMM 2007)*, 2007, pp. 260-264 : IEEE

- [34] Y. Lei, F. Shang, Z. Long, and Y. Ren, "An energy efficient multiple-hop routing protocol for wireless sensor networks," in 2008 First International Conference on Intelligent Networks and Intelligent Systems, 2008, pp. 147-150 : IEEE
- [35] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, 1967, vol. 1, no. 14, pp. 281-297 : Oakland, CA, USA
- [36] L. Kaufman and P. J. Rousseeuw, Finding groups in data : an introduction to cluster analysis. John Wiley and Sons, 2009

Résumé

L'utilisation des réseaux de capteurs sans fil dans plusieurs domaines a poussé les chercheurs à proposer des protocoles de routage économes en énergie pour garantir une longue longévité des réseaux vu que ces derniers sont considérés comme des réseaux présentant une certaine autonomie en termes d'énergie.

Dans ce travail, on a proposé un schéma de routage basé sur l'approche de clustering. Ce protocole permet d'organiser le réseau en clusters en tenant compte de la distance entre les nœuds et utilise un schéma de routage pour acheminer les données d'un nœud source à la station de base. Ce schéma de routage permet d'établir des chemins de clusterhead à clusterhead (CH-to-CH) jusqu'à l'aboutissement à la station de base ce qui réduit la consommation d'énergie et par conséquent augmente la durée de vie du réseau.

Mots clés :

Clustering, Greedy Forwarding, K-means, Réseaux de capteurs, Routage

Abstract

The use of wireless sensor networks in several areas has prompted researchers to propose energy efficient routing protocols to ensure long network lifetime as the networks are considered energy autonomy networks.

In this work, we proposed a routing scheme based on the clustering approach. This protocol organizes the network into clusters by considering the distance between the nodes and uses a routing scheme to route data from a source node to the base station. This routing scheme establishes paths from clusterhead to clusterhead (CH-to-CH) until the base station, which reduces energy consumption and therefore increases network lifetime.

Keywords :

Clustering, Greedy Forwarding, K-means, Réseaux de capteurs, Routage