Résilience et application aux protocoles de routage dans les réseaux de capteurs

Ochirkhand Erdene-Ochir

N d'ordre 2013-ISAL-XXXX Année 2013

Thèse

présentée devant
L'Institut National des Sciences Appliquées de Lyon
pour l'obtention
du Grade de Docteur

présentée et soutenue publiquement
le 05 07 2013

par

ERDENE-OCHIR Ochirkhand

Résilience et application aux protocoles de routage dans les réseaux de capteurs

Soutenue devant :

Bernard TOURANCHEAU (Rapporteur)	Professeur Université Joseph Fourier de Grenoble
Abdelmadjid BOUABDALLAH (Rapporteur)	Professeur Université de Technologie de Compiègne
Pascal LORENZ (Examinateur)	Professeur Université de Haute Alsace
Fabrice VALOIS (Encadrant)	Professeur INSA de Lyon
Marine MINIER (Encadrant)	Maître de conférence INSA de Lyon
Apostolos KOUNTOURIS (Encadrant)	Expert Orange Labs de Meylan

Les travaux présentés dans ce mémoire ont été financés partiellement par le projet ANR ARESA2. Ces travaux ont été réalisés en collaboration entre le laboratoire CITI sous la direction de Prof. Fabrice VALOIS et Dr. Marine MINIER et Orange Labs sous la direction de Dr. Apostolos KOUNTOURIS.

« A mon père. »

Remerciements

Je remercie Bernard Tourancheau, Professeur à Université Joseph Fourier de Grenoble et Abdelmadjid Bouabdallah, Professeur à Université de Technologie de Compiègne, d'avoir accepté d'être rapporteurs de ma thèse. Je remercie Pascal Lorenz, Professeur à Université Haute Alsace, d'avoir accepté d'évaluer mon travail.

J'exprime mes chaleureux remerciements et toute ma gratitude à mes directeurs de thèse, Fabrice Valois, Marine Minier et Apostolos Kountouris, qui m'ont offert l'opportunité de réaliser cette thèse, qui ont su me former et m'orienter avec rigueur et patience durant ces 4 années, qui m'ont encouragé tout au long de ce travail. Je remercie pour leur disponibilité, leurs conseils très précieux et leurs qualités pédagogiques et humaines à l'élaboration de cette thèse.

Cette thèse a été réalisée à Orange Labs de Meylan et au laboratoire CITI de l'INSA de Lyon. Je tiens donc à remercier Vincent Gimeno et Pierre Madillo respectivement, l'ancien et le nouveau manager de l'équipe TECH-MATIS-CITY d'Orange Labs, ainsi que tous mes collègues du travail, et en particulier mes camarades thésards Camille, Zheng, Ahmed, Bilel et Quentin.

Je tiens à remercier Jean-Marie Gorce, le directeur du CITI, Hervé Rivano, le responsable de l'équipe URBANET, ainsi que les enseignants et chercheurs du CITI, et mes camarades, les anciens thésards de CITI, Wassim, Elyes, Amira, Ioan, Virgile, Paul, Loïc, Fatiha, Anya, Hajer, Anis, Ibrahim, Ahmad, Leila et en particulier Sana pour son soutien constant et ses conseils précieux. Je remercie les personnels administratifs de CITI qui font un travail indispensable à la réalisation de toute thèse, Gaëlle, Maguy, Joëlle et Stéphane.

Je remercie ma famille en Mongolie, mon père Erdene-Ochir Sanga, à qui j'ai dédié cette thèse, ma belle mère Enkhtuya Dorj, mes sœurs Oshima, Nachka, Susu, Sasu et mes frères Sandag, Ulzii pour leur inconditionnel amour et irremplaçable soutien. Je remercie mon fils Kherlen de m'avoir donné du courage et d'avoir été relativement compréhensif sur l'indisponibilité de sa maman les dernières semaines de la rédaction. Je remercie également ma famille en France, Pascal, Béatrice, Mathéo, Ulysse, Maud, Sam, Isyan, Pierre et en particulier Nicolette pour sa disponibilité, son soutien constant et ses encouragements. Je remercie tous mes amis et en particulier Nemekhbayar, Franck, Dominique, Corinne, Marie-Solange, Daniel, Liesel, Philippe, Alain, Colette, Nathalie, Frédéric, Mathilde, Maxim, Manon, Petra, Premila, Marc, Alex, Sabrina, Coline, Daarii, Bagi, Julien, Meryame, Mehdi, Maryam et Laetitia. Je remercie en particulier Françoise, qui a toujours été présente à mes côtés, qui a su m'épauler à toutes les épreuves et aussi à la réalisation de cette thèse.

Enfin, je remercie Karel, qui m'a apporté un soutient déterminant dans toutes les phases de ce travail, qui a toujours été présent et qui m'a encouragé durant ces 4 années de thèse et surtout durant la dernière ligne droite.

Resumé

Les réseaux de capteurs sans fil sont constitués d'un grand nombre de nœuds, déployés pour collecter des données du monde physique (température, humidité, pollution etc.) et les transmettre, de manière autonome, vers un ou plusieurs points de collectes appelés « puits ». Dans cette thèse, nous nous focalisons sur la sécurité des protocoles de routage multi-sauts, plus particulièrement, sur la notion de résilience[1] aux attaques.

Les domaines d'applications des réseaux de capteurs sont variés, allant du suivi médical à la surveillance environnementale en passant par le bâtiment intelligent ou le monitoring urbain (éclairage, pollution, relevé de compteurs d'eau/électricité/gaz etc.). Dans ces applications, les capteurs sont souvent déployés dans des environnements ouverts et accessibles permettant aux éventuels attaquants de les détruire ou de les capturer afin d'en extraire les données sensibles (clés de chiffrement, identité, adresse, etc.). La compromission des nœuds est un problème majeur pour la sécurité de réseaux de capteurs, puisqu'un adversaire peut s'introduire à l'intérieur du périmètre de sécurité. Les méthodes traditionnelles, basées sur la cryptographie, permettent d'obtenir une sécurité de base (authentification, confidentialité, intégrité, non répudiation etc.), mais ne permettent pas toujours de se prémunir contre les attaques dues à la compromission des nœuds (réplication des nœuds, Sybil, Selective forwarding, Blackhole, Sinkhole, Wormhole etc.). Dans le but d'apporter des solutions algorithmiques complémentaires aux solutions cryptographiques, nous étudions la résilience des protocoles de communication en présence d'adversaires internes visant à perturber le routage de l'information à travers le réseau.

Dans un premier temps, nous introduisons le concept de résilience. Notre objectif est de proposer une définition explicitant le terme de résilience dans notre contexte et une métrique, permettant de comparer efficacement les protocoles de routage. L'originalité de cette métrique est d'utiliser à la fois une représentation graphique et une méthode de calcul quantitative liée à celle-ci. La représentation graphique à deux dimensions permet une vision synthétique de la résilience des protocoles selon plusieurs paramètres de performance. La méthode de calcul quantitative liée à cette représentation graphique agrège les valeurs des paramètres et permet de classifier les protocoles en termes de résilience. Grâce à cet outil, nous avons évalué la résilience de plusieurs protocoles de routage classiques de différentes catégories. Cette étude nous a permis d'identifier les mécanismes permettant d'améliorer la résilience des protocoles.

Dans un second temps, nous proposons les mécanismes résilients de routage pour les réseaux de capteurs. Les mécanismes résilients que nous proposons consistent en trois éléments :

1. La résilience d'un protocole de routage est sa capacité à absorber la dégradation des performances en présence des attaques internes dues aux compromissions des nœuds. C'est la capacité de continuer à livrer des messages avec l'augmentation du nombre de nœuds non fiables ou compromis dans le réseau. En d'autres termes, c'est la capacité d'un protocole de routage à subir et à surmonter la présence de nœuds non fiables ou compromis.

(i) introduire un comportement aléatoire (ii) limiter la longueur des routes (iii) ajouter de la réplication de paquets. Les comportements aléatoires augmentent l'incertitude pour les adversaires, rendant les protocoles moins prévisibles, les réplications des données permettent de bénéficier la diversification des routes créées entre les sources et le puits, en améliorant ainsi le succès et l'équité de livraison et la limitation de la longueur des routes est nécessaire pour diminuer la probabilité qu'un paquet tombe sur un nœud attaquant en route. La connexité entre les capteurs et le puits est ainsi augmentée. Grâce à notre métrique de résilience, nous avons proposé une nouvelle taxonomie de résilience. Selon cette taxonomie, le routage par gradient et la marche aléatoire biaisée avec les mécanismes proposés sont les plus résilients.

Nous avons donc évalué par la suite le routage par gradient en cas d'attaques combinées pour approfondir notre étude, mais aussi pour savoir si ces mécanismes proposés permettent d'augmenter la résilience même en cas d'attaques plus complexes, visant différents aspects du routage (construction des routes, paquets de contrôle, etc.). Nous avons introduit plusieurs valeurs de biais aux variantes aléatoires du routage par gradient pour étudier l'influence de l'entropie et nous les avons comparées à sa version classique. Nous avons également évalué leur résilience en introduisant deux types de réplications (uniformes et adaptatives). Sans attaques, ce sont les variantes les plus biaisées sans réplications qui sont les plus performantes. En cas d'attaques peu importantes, les réplications uniformes sont plus efficaces, tandis qu'en cas d'attaques plus intenses, ce sont les réplications adaptatives qui se montrent les plus efficaces. Les études menées jusqu'à ici étaient produites par des simulations et nous avions donc besoin d'une justification théorique.

Nous avons donc proposé une étude théorique de la marche aléatoire biaisée en cas d'attaques de non-retransmission des paquets. Nous avons évalué l'influence du biais, mais aussi les deux réplications que nous avions évaluées précédemment par des simulations. En premier lieu, nous avons étudié le succès de livraison et la consommation d'énergie pour tous les scénarios. Ensuite, nous les avons évalués selon notre métrique de résilience. Cette étude a permit de confirmer les résultats d'étude par simulations et elle a montré que le biais est indispensable pour la résilience et le seuil d'entropie bénéfique à la résilience est $\varepsilon = 0.7$ *quand la réplication de données est introduite. En dessous de cette valeur, la marche aléatoire est inefficace à cause de la longueur de chemins trop importante.*

L'ensemble des travaux réalisés dans cette thèse se concentre autour de la résilience. Ce concept reste assez nouveau, en particulier dans le domaine des réseaux et télécommunications. À travers cette thèse, nous avons voulu donner notre vision sur ce thème en nous concentrant sur les problématiques de sécurité des protocoles de routage dans le contexte des réseaux de capteurs.

Abstract

Wireless Sensor Networks (WSNs) are composed of a large number of low-cost, low power, and multifunctional sensor nodes that communicate at short distances through wireless links. In many cases these sensor nodes are deployed over a large geographic area in order to collect physical world data (temperature, humidity, pollution, etc.) and route them towards one or few destinations called the "sinks". This thesis focuses on the security issues of multi-hop routing protocols, especially on the resiliency[2] concept.

The rapid deployment capabilities, due to the lack of infrastructure, as well as the self organized and potentially fault-tolerant nature of WSNs make them attractive for multiple applications spanning from environmental monitoring (temperature, pollution, etc.) to building industrial automation (electricity/gas/water metering, event detection, home automation, etc.). Security is particularly challenging in WSNs. Because of their open and unattended deployment, in possibly hostile environments, powerful adversaries can easily launch Denial-of-Service (Dos) attacks, cause physical damage to sensors, or even capture them to extract sensitive information (encryption keys, identities, addresses, etc.). Consequently, the node compromise poses severe security and reliability concerns, since it allows an adversary to be considered as a legitimate node inside the network. After node compromise, an adversary can seek to disrupt the functionality of network layer by launching attacks such as node replication, Sybil, Selective forwarding, Sinkhole, Wormhole, etc. To cope with these "insider" attacks, stemming from node compromise, "beyond cryptography" algorithmic solutions must be envisaged to complement the traditional cryptographic solutions.

Firstly, we propose the resiliency concept. Our goal is to propose a definition of the resiliency in our context (security of WSNs routing protocols) and a new metric to compare routing protocols. The originality of this metric is that we combine the graphical representation (qualitative information) with the aggregation method (quantitative information). We introduce a two dimensional graphical representation with multiple axes forming an equiangular polygon surface. This method allows to aggregate meaningfully several parameters and makes it easier to visually discern various tradeoffs, thus greatly simplifying the process of protocol comparison.

Secondly, we propose the protocol behaviors enhancing resiliency. Our proposition consists in three elements : (i) introduce random behaviors (ii) limit route length (iii) introduce data replication. Random behaviors increase uncertainty for an adversary, making the protocols unpredictable. Data replication allows route diversification between the sources and the sink, thus improving the delivery success and fairness. Limitation of the route length is necessary

2. The resiliency of a routing protocol is its ability to absorb the performance degradation under some failure pattern (random or intentional) and to continue delivering messages with an increasing number of k compromised nodes. In other words, its the ability of a routing protocol to endure and overcome the presence of unreliable and/or compromised components.

to reduce the probability of a data packet to meet a malicious insider along the route. The quantitative metric enables to propose a new resiliency taxonomy of WSNs routing protocols. According to this taxonomy, the gradient based routing is the most resilient when it is combined with the proposed behaviors.

Thirdly, several variants of the gradient-based routing (classical and randomized) under more complex and realistic adversary model (several combined attacks) are considered to extend our simulations. Several values of bias are introduced to the randomized variants and two data replication methods (uniform and adaptive) are considered. Without attacks, the most biased variants without replications are the most efficient. However, under moderate attacks, the replication uniform is the most adapted, while under intense attacks, the replication adaptive is the most suitable.

Finally, a theoretical study of the resiliency is introduced. We present an analytical study of the biased random walk routing under attacks. The influence of bias is evaluated and two replication methods that previously evaluated by simulations are considered. After presenting the delivery success and the energy consumption of all scenarios, we evaluate them with our resiliency metric. This study permits to confirm the results obtained with simulations and it shows that the bias is essential to enhance the resiliency of random routing.

The work presented in this thesis focuses on the resiliency. This concept is relatively new, especially in the network protocol engineering. Through this thesis, we wanted to give our view on this topic by focusing on the "beyond cryptography" algorithmic solutions of WSNs routing protocols.

Table des matières

1 Introduction **2**
- 1.1 Contexte . . . 3
  - 1.1.1 Réseaux de capteurs sans fil . . . 3
  - 1.1.2 Sécurité des réseaux de capteurs . . . 7
  - 1.1.3 Applications et scénarios . . . 9
- 1.2 Problématiques et motivations . . . 10
- 1.3 Contributions et organisation de la thèse . . . 11

2 État de l'art **13**
- 2.1 Problématiques de sécurité dans les réseaux de capteurs . . . 14
  - 2.1.1 Attaques et défenses des réseaux de capteurs par couche OSI . . . 14
  - 2.1.2 Protocoles de routage sécurisés . . . 25
- 2.2 Étude préliminaire des protocoles de routage en cas d'attaques . . . 27
  - 2.2.1 Hypothèses du réseau, modèles d'adversaires et protocoles étudiés . . . 28
  - 2.2.2 Évaluation des performances . . . 33
- 2.3 Marches aléatoires . . . 38
  - 2.3.1 Marche aléatoire . . . 38
  - 2.3.2 Marche aléatoire avec mémoire . . . 39
  - 2.3.3 Marche aléatoire biaisée . . . 40
- 2.4 Conclusion . . . 41

3 Concept de résilience : Définition et Métrique **42**
- 3.1 Introduction . . . 43
- 3.2 Motivations . . . 43
- 3.3 Définition de la résilience . . . 44
- 3.4 Métrique de résilience . . . 46
  - 3.4.1 Méthode d'agrégation des multiples paramètres . . . 47
  - 3.4.2 Paramètres de résilience pour les réseaux de capteurs . . . 49
- 3.5 Application de la métrique aux protocoles de routage classiques . . . 52
  - 3.5.1 Modèle d'adversaires . . . 52
  - 3.5.2 Évaluation des performances . . . 53
- 3.6 Conclusion . . . 56

4 Proposition des mécanismes résilients pour le routage **58**
- 4.1 Introduction . . . 59
- 4.2 Mécanismes résilients pour le routage . . . 59
  - 4.2.1 Introduction de comportements aléatoires . . . 60

4.2.2 Limitation de la longueur des routes . . . 60
4.2.3 Réplication des paquets . . . 61
4.3 Introduction des mécanismes résilients aux protocoles de routage classiques . 61
4.3.1 Protocoles aléatoires sans réplication . . . 61
4.3.2 Protocoles aléatoires avec réplications . . . 62
4.3.3 Évaluation des performances . . . 63
4.3.4 Taxonomie de la résilience des protocoles de routage . . . 71
4.4 Conclusion . . . 72

5 Étude de la résilience du routage par gradient en cas de plusieurs attaques combinées 74
5.1 Introduction . . . 75
5.2 Mécanismes résilients appliqués au routage par gradient (GBR) . . . 76
5.2.1 GBR classique . . . 76
5.2.2 GBR aléatoires sans réplication . . . 76
5.2.3 GBR aléatoires avec réplications . . . 77
5.3 Modèles d'adversaires . . . 78
5.3.1 *Selective forwarding* basique . . . 79
5.3.2 *Sinkhole* combinée . . . 79
5.3.3 *Sybil* combinée . . . 79
5.3.4 *Wormhole* combinée . . . 81
5.4 Évaluation des performances . . . 81
5.4.1 Résultats en cas d'attaque *Selective forwarding* basique . . . 82
5.4.2 Résultats en cas d'attaque *Sybil* combinée . . . 90
5.4.3 Résultats en cas d'attaque *Wormhole* combinée . . . 94
5.4.4 Résultats en cas d'attaque *Sinkhole* combinée . . . 98
5.5 Conclusion . . . 100

6 Étude théorique de la résilience des marches aléatoires biaisées 102
6.1 Introduction . . . 103
6.2 Définitions, notations et hypothèses . . . 103
6.2.1 Marche aléatoire biaisée . . . 104
6.2.2 Hypothèses et paramètres de calcul . . . 105
6.2.3 Longueur moyenne des chemins . . . 106
6.3 Marche aléatoire biaisée sans réplication . . . 107
6.3.1 Succès de livraison . . . 108
6.3.2 Consommation d'énergie . . . 110
6.4 Marche aléatoire biaisée avec réplications . . . 113
6.4.1 Réplications uniformes . . . 113
6.4.2 Réplications adaptatives . . . 117
6.5 Évaluation de la résilience . . . 119
6.5.1 Marche aléatoire sans réplication . . . 119
6.5.2 Marche aléatoire avec réplications uniformes . . . 120
6.5.3 Marche aléatoire avec réplications adaptatives . . . 122
6.6 Conclusion . . . 124

7 Conclusion **125**
7.1 Bilan 126
7.2 Perspectives 129
7.2.1 Expérimentations 129
7.2.2 Mensonges sur les paquets de contrôle 129
7.2.3 Mécanismes de détection d'attaques 129
7.2.4 Codage réseau pour la résilience 130

Liste des publications **131**

Bibliographie **132**

Table des figures

1.1 Capteur sans fil. 3
1.2 Architecture d'un nœud capteur. 4
1.3 Les réseaux de capteurs. 5
1.4 Principales caractéristiques de quelques capteurs les plus courants. 6

2.1 Attaque *Sybil*. 19
2.2 Attaque par réplication des nœuds. 20
2.3 Attaque *Selective forwarding*. 21
2.4 Attaque *Sinkhole*. 22
2.5 Attaque *Wormhole*. 23
2.6 Distribution des nœuds compromis (cercle) uniformément : (a) dans le réseau (b) autour du puits (carré). 29
2.7 Catégories des protocoles de routage dans les réseaux de capteurs [1]. 30
2.8 DSR : Envoi des paquets (a) `RREQ` (b) `RREP` (c) `DATA` 31
2.9 GBR : Envoi des paquets (a) `INIT` (b) `DATA` 32
2.10 GF : Envoi des paquets (a) `HELLO` (b) `DATA` 32
2.11 RWR : Envoi des paquets (a) `HELLO` (b) `DATA` 33
2.12 BRWR : Envoi des paquets (a) `INIT` (b) `DATA` 34
2.13 Scénario 1 : (a) Taux de livraison moyen (b) Longueur moyenne des chemins. 35
2.14 Scénario 2 : (a) Taux de livraison moyen (b) Longueur moyenne des chemins. 37

3.1 Essai de résilience de Charpy. 45
3.2 Différentes grandeurs de surface de résilience : (a) cas ordinaire (b) mauvaise résilience (c) bonne résilience. 48
3.3 Différentes caractéristiques avec une surface de résilience identique : (a) dominé par 1 et 5 (b) dominé par 2 et 3 (c) équilibré. 49
3.4 Différents ordres des axes avec un ensemble de valeurs identiques : (a) importance primaire d'énergie (b) importance secondaire d'énergie. 52
3.5 Surface de la résilience des protocoles classiques sans attaque et avec attaques ($k = 30\%$). 53
3.6 Évaluation quantitative de la résilience des protocoles classiques en cas d'attaques. 54
3.7 Différentiel de la résilience des protocoles classiques en cas d'attaques. 55

4.1 Surface de la résilience de DSR sans attaque et avec attaques ($k = 30\%$) : (a) Version classique DSR (b) Variante aléatoire RS-DSR (c) Variante aléatoire avec réplication RM-DSR. 63

4.2 Surface de la résilience de GBR sans attaque et avec attaques ($k = 30\%$) : (a) Version classique GBR (b) Variante aléatoire RS-GBR (c) Variante aléatoire avec réplication RM-GBR. 63
4.3 Surface de la résilience de GF sans attaque et avec attaques ($k = 30\%$) :(a) Version classique GF (b) Variante aléatoire RS-GF (c) Variante aléatoire avec réplication RM-GF. 64
4.4 Surface de la résilience de RWR sans attaque et avec attaques ($k = 30\%$) : (a) Version classique RWR (b) Variante aléatoire avec réplication RM-RWR. 64
4.5 Surface de la résilience de BRWR sans attaque et avec attaques ($k = 30\%$) : (a) Version classique BRWR (b) Variante aléatoire avec réplication RM-BRWR. 64
4.6 Évaluation quantitative de la résilience des variantes de DSR. 66
4.7 Évaluation quantitative de la résilience des variantes de GBR. 66
4.8 Évaluation quantitative de la résilience des variantes de GF. 67
4.9 Évaluation quantitative de la résilience des variantes de RWR. 67
4.10 Évaluation quantitative de la résilience des variantes de BRWR. 68
4.11 Évaluation quantitative de la résilience des protocoles de routage aléatoires en cas d'attaques. 69
4.12 Évaluation quantitative de la résilience des protocoles aléatoires avec réplication. 70
4.13 Taxonomie de la résilience des protocoles de routage. 71

5.1 Exemple de fonctionnement de GBR-Rd $p = 0.8$. 77
5.2 Attaque *Selective forwarding* basique. 79
5.3 Attaque *Sinkhole* combinée. 80
5.4 Attaque *Sybil* combinée. 80
5.5 Attaque *Wormhole* combinée. 81
5.6 Surface de résilience en cas d'attaque *Selective forwarding* : sans réplication. 82
5.7 Surface de résilience en cas d'attaque *Selective forwarding* : avec réplications uniformes. 83
5.8 Surface de résilience en cas d'attaque *Selective forwarding* : avec réplications adaptatives. 83
5.9 Distribution de taux de livraison en cas de $k = 10\%$ d'attaquants : Cinq classes $c1$ à $c5$ et quatre distances $h1$ à $h4$ en nombre de sauts. 85
5.10 Résilience en cas d'attaque *Selective forwarding* : sans réplication. 85
5.11 Résilience en cas d'attaque *Selective forwarding* : avec réplications uniformes. 86
5.12 Résilience en cas d'attaque *Selective forwarding* : avec réplications adaptatives. 86
5.13 Consommation totale d'énergie (normalisée par GBR-Dt sans réplication) en cas d'attaque *Selective forwarding* : sans réplication. 87
5.14 Consommation totale d'énergie (normalisée par GBR-Dt sans réplication) en cas d'attaque *Selective forwarding* : avec réplications uniformes. 88
5.15 Consommation totale d'énergie (normalisée par GBR-Dt sans réplication) en cas d'attaque *Selective forwarding* : avec réplications adaptatives. 88
5.16 Taux de livraison moyen en cas d'attaque *Selective forwarding* : sans réplication. 89
5.17 Taux de livraison moyen en cas d'attaque *Selective forwarding* : avec réplications uniformes. 89

5.18 Taux de livraison moyen en cas d'attaque *Selective forwarding* : avec réplications adaptatives. 90
5.19 Surface de résilience en cas d'attaque *Sybil* : sans réplication. 90
5.20 Surface de résilience en cas d'attaque *Sybil* : avec réplications uniformes. . . 91
5.21 Surface de résilience en cas d'attaque *Sybil* : avec réplications adaptatives. . 91
5.22 Résilience en cas d'attaque *Sybil* : sans réplication. 92
5.23 Résilience en cas d'attaque *Sybil* : avec réplications uniformes. 92
5.24 Résilience en cas d'attaque *Sybil* : avec réplications adaptatives. 93
5.25 Surface de résilience en cas d'attaque *Wormhole* : sans réplication. 94
5.26 Surface de résilience en cas d'attaque *Wormhole* : avec réplications uniformes. 95
5.27 Surface de résilience en cas d'attaque *Wormhole* : avec réplications adaptatives. 95
5.28 Résilience en cas d'attaque *Wormhole* : sans réplication. 96
5.29 Résilience en cas d'attaque *Wormhole* : avec réplications uniformes. 96
5.30 Résilience en cas d'attaque *Wormhole* : avec réplications adaptatives. 97
5.31 Surface de résilience en cas d'attaque *Sinkhole* : sans réplication. 97
5.32 Surface de résilience en cas d'attaque *Sinkhole* : avec réplications uniformes. 98
5.33 Surface de résilience en cas d'attaque *Sinkhole* : avec réplications adaptatives. 98
5.34 Résilience en cas d'attaque *Sinkhole* : sans réplication. 99
5.35 Résilience en cas d'attaque *Sinkhole* : avec réplications uniformes. 99
5.36 Résilience en cas d'attaque *Sinkhole* : avec réplications adaptatives. 100

6.1 Tore $\Im$ [2] 104
6.2 Tore avec un seul puits. 106
6.3 Longueur moyenne des chemins en fonction du biais ε pour chaque position r_1 des nœuds. 107
6.4 Succès de livraison d'un nœud à la position r_1 en cas d'attaque (P_c). 109
6.5 Succès de livraison d'un nœud à la position r_1 en cas d'attaque (P_c). 109
6.6 Succès de livraison d'un nœud à la position r_1 en cas d'attaque (P_c). 109
6.7 Succès de livraison moyen en cas d'attaque (P_c). 110
6.8 Consommation d'énergie totale en cas d'attaques (P_c). 112
6.9 Part du gaspillage d'énergie à cause des paquets perdus sur l'énergie totale. . 112
6.10 Succès de livraison moyen en cas d'attaque (P_c) : avec 2 réplications 114
6.11 Succès de livraison moyen en cas d'attaque (P_c) : avec 3 réplications 114
6.12 Gain en taux de livraison en fonction de taux de réplications x avec $P_c = 0.1$ d'attaques. 115
6.13 Consommation d'énergie totale en cas d'attaque : avec 2 réplications 116
6.14 Consommation d'énergie par paquets reçus. 117
6.15 Succès de livraison moyen en cas d'attaques (P_c) : avec réplications adaptatives 118
6.16 Consommation d'énergie totale en cas d'attaque : avec réplications adaptatives 118
6.17 Surface de résilience des marches aléatoires biaisées sans réplication. 120
6.18 Résilience des marches aléatoires biaisées sans réplication. 121
6.19 Surface de résilience des marches aléatoires avec réplications uniformes. . . . 121
6.20 Résilience des marches aléatoires biaisées avec réplications uniformes. 122
6.21 Surface de résilience des marches aléatoires avec réplications adaptatives. . . 123
6.22 Résilience des marches aléatoires biaisées avec réplications adaptatives. . . . 123

Liste des tableaux

2.1 Résumé des paramètres de simulations. 34

5.1 Notations 77

1 Introduction

Sommaire

1.1 Contexte 3
1.1.1 Réseaux de capteurs sans fil 3
1.1.2 Sécurité des réseaux de capteurs 7
1.1.3 Applications et scénarios 9
1.2 Problématiques et motivations 10
1.3 Contributions et organisation de la thèse 11

1.1 Contexte

Les avancées de la technologie de télécommunications ont favorisé le développement des réseaux sans fil dans des domaines variés d'applications. Le progrès technique permet aujourd'hui de produire des capteurs miniaturisés à faible coût, dotés de la capacité de récolter des grandeurs physiques et de les acheminer d'un point à un autre grâce au médium radio et à du routage multi-sauts. Dans cette thèse, nous étudions les problématiques de sécurité des réseaux de capteurs, plus particulièrement au niveau de la couche réseau.

Cette thèse a été réalisée en collaboration avec Orange Labs à Meylan, sous la direction de Dr. Apostolos Kountouris, et le laboratoire de recherche CITI de l'INSA de Lyon sous la direction de Prof. Fabrice Valois et de Dr. Marine Minier. Elle a été financée par le projet de recherche ANR ARESA2 [3]. L'objectif du projet ANR ARESA2 est de raccorder au monde IP des réseaux radio de capteurs dynamiques et en garantir la sécurité à un coût énergétique maîtrisé. Dans cette thèse, nous nous sommes donc intéressés aux réseaux de capteurs urbains considérés par le projet ANR ARESA2 et nous apportons des solutions de sécurité, en particulier, nous proposons le concept de la résilience des protocoles de routage.

1.1.1 Réseaux de capteurs sans fil

Les réseaux de capteurs sans fil sont composés d'un grand nombre de capteurs multifonctionnels (Fig. 1.1).

FIGURE 1.1: Capteur sans fil.

Ce sont des entités électroniques simples, peu coûteuses et dont la puissance est limitée en termes de calcul, de mémoire et d'énergie. Malgré la diversité des capteurs sur le marché, leur l'architecture matérielle reste similaire (Figure 1.2). Ils sont composés de :

- une unité de mesure : qui consiste à mesurer une grandeur physique et à la transformer en une grandeur numérique ;
- une unité de traitement : composée d'un processeur et de mémoire ;
- une source d'énergie (batterie, photovoltaïque, énergie cinétique ou thermique) ;
- une unité de transmission sans fil ;

Les capteurs sont capables de communiquer entre eux grâce à une connexion radio, formant ainsi un réseau de capteurs sans fil (WSNs - *Wireless Sensor Networks*). Une des fonctionnalités importantes des réseaux de capteurs est leur capacité à être déployée en grand nombre dans des zones géographiques étendues pour recueillir des données du monde physique (température, humidité, pression, gaz, etc.). Les données recueillies peuvent être ensuite collectées de manière autonome vers un ou plusieurs nœuds destination appelés "puits" (Fig. 1.3).

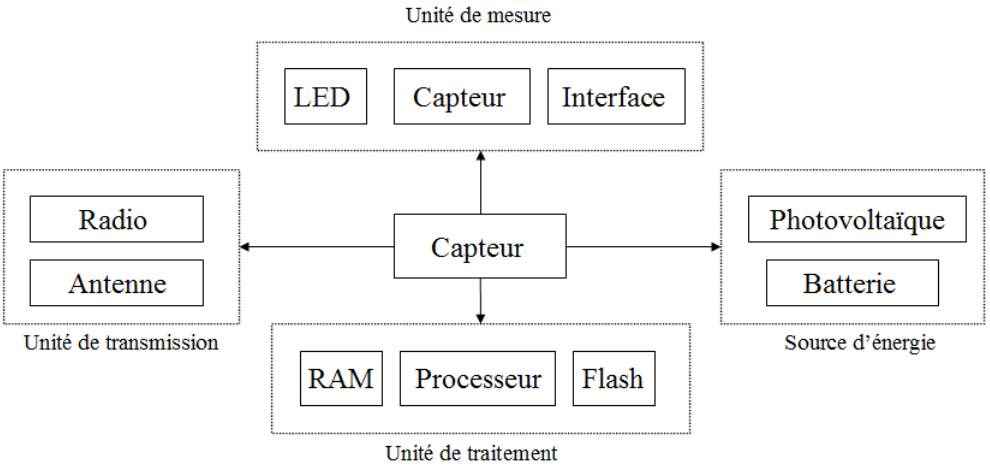


FIGURE 1.2: Architecture d'un nœud capteur.

Les réseaux de capteurs ont une large gamme d'applications telles que le contrôle industriel, la supervision et la surveillance des paramètres environnementaux [4], dans l'agriculture [5], les bâtiments intelligents [6], la domotique [7], les applications militaires [8], le suivi médical [9], etc. Les réseaux de capteurs présentent de multiples avantages tels que le déploiement rapide et le fait qu'ils soient peu onéreux, auto-organisés et tolérants aux pannes.

Les réseaux de capteurs sont considérés comme proches des réseaux ad-hoc (MANETs - *Mobile Ad-hoc NETworks*). Ils partagent effectivement certains points communs tels que la communication radio, l'architecture décentralisée, auto-organisée, auto-configurée ainsi que le routage multi-sauts. Les réseaux ad-hoc sont considérés comme limités en ressources, mais les réseaux de capteurs ont des limitations de ressources beaucoup plus importantes, plus particulièrement une contrainte forte en consommation d'énergie. De plus, les réseaux ad-hoc utilisent généralement un profil de trafic "point-à-point" (communication *any-to-any*), tandis que les réseaux de capteurs utilisent de préférence un profil de trafic *convergecast* (*many-to-one*). Enfin, les réseaux ad-hoc peuvent être plus dynamiques avec les nœuds mobiles, alors que dans les réseaux de capteurs les nœuds sont souvent considérés comme statiques.

Nous détaillons les caractéristiques des réseaux de capteurs comme suit :

- Profil du trafic :
  - *many-to-one* : plusieurs capteurs envoient leurs données recueillies à un point de collecte. En présence de plusieurs puits, le profil de trafic est *many-to-few*.
  - *one-to-many* : le puits peut envoyer des informations de routage par un mécanisme d'inondation (*flooding*) des paquets de contrôle aux capteurs. S'il y a plusieurs points de collecte, le trafic est *few-to-many*.
  - *one-to-any* : le puits peut interroger un nœud capteur spécifique.
  - *any-to-one* : un nœud capteur spécifique peut envoyer des données au puits.
- Contraintes matérielles : Les réseaux de capteurs ont des caractéristiques et contraintes spécifiques comparées aux réseaux traditionnels. Les capteurs sont des entités électroniques très simples limitées en mémoire, en énergie, mais également en capacité de calculs. Le tableau 1.4 résume les principales caractéristiques de plusieurs capteurs ac-

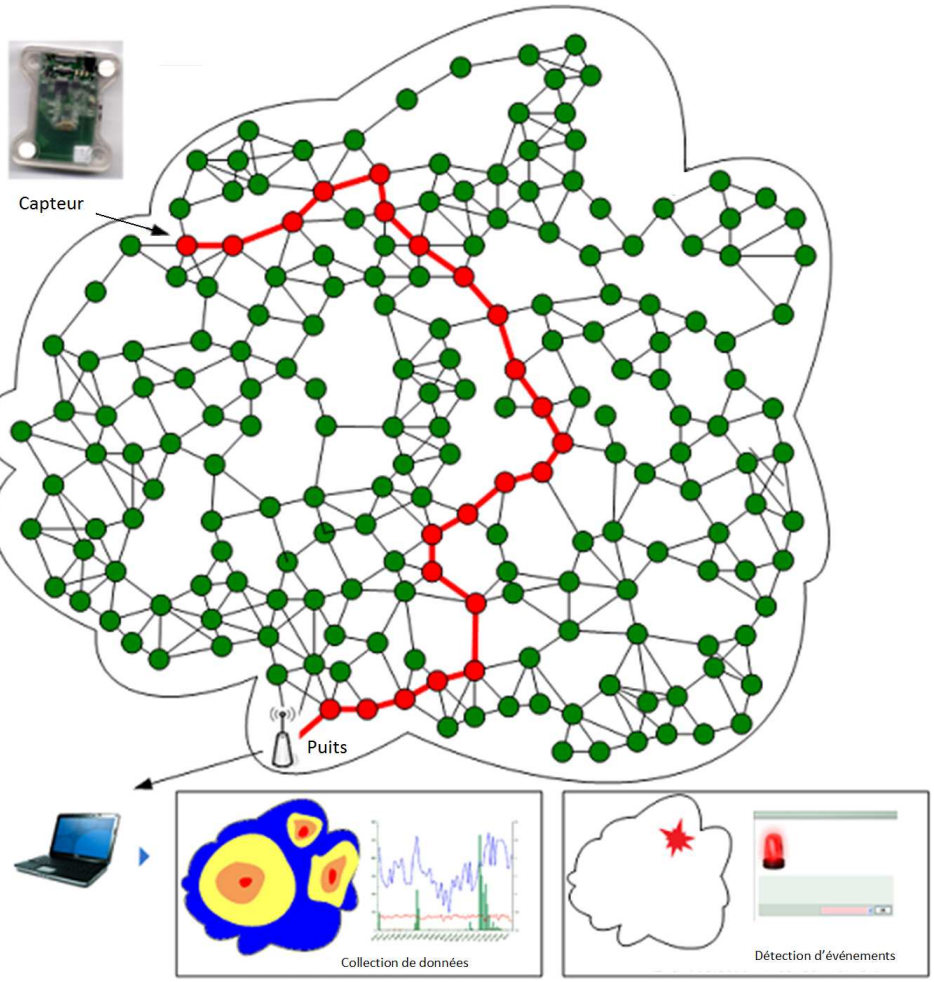


FIGURE 1.3: Les réseaux de capteurs.

tuels.

- Limitations en mémoire et en puissance de calcul : un capteur est une entité très simple avec seulement une petite quantité de mémoire et de puissance de calcul. Par exemple, le capteur MicaZ, couramment utilisé, ne dispose que de 4Ko de RAM, 128Ko de ROM, 512Ko de mémoire flash et d'un microcontrôleur de 8MHz - Atmel AVR Atmega.
- Limitations en énergie : un capteur ne peut être équipé que d'une batterie limitée. Parce que changer les piles dans un grand nombre des capteurs (de l'ordre de quelques centaines à plusieurs milliers) déployé dans une zone géographique étendue demande une intervention de maintenance importante et donc peu pratique. La durée de vie d'un capteur est donc fortement dépendante de la durée de vie de sa batterie.
- Limitations en puissance de transmission : un capteur a des capacités limitées en transmission radio. Les communications du réseau dépendent souvent de la coopération locale des capteurs en se basant sur l'acheminement des données via un routage multi-sauts.

– Environment :

Capteurs	MicaZ [10]	TelosB [10]	Imote2 [10]	WSN430 [11]	Shimmer [12]
Processeur	Atmel ATMega 128L	TI MSP430	Intel PXA271 XScale	TI MSP430	MSP430 F1611
Vitesse du processeur	16 MHz	8 MHz	13 - 416 MHz	8 Mhz	8 Mhz
RAM	4 Ko	10 Ko	256 Ko	10 Ko	10 Ko
Espace programme	128 Ko	48 Ko	32 Mo	48 Ko	48 Ko
Flash	512 Ko	256 Ko	32 Mo	1 Mo	48 Ko
Communication série	UART	DIO, SPI, I2C, UART	UART, GPIO	DS2411	DS2411
Batterie	2xAA	2/3A	3x AAA	PoLiFlex	50mAh Li-ion
Voltage	2.7 V	1.8 - 3.6 V	3.2 - 4.5 V	2.2 V	2.2 - 3.6 V
Radio	TI CC2420 802.15.4	TI CC2420 802.15.4	TI CC2420	TI CC1100	TI CC2420 802.15.4
Fréquence (MHz)	2400-2483	2400-2483	2400-2483	315/433/ 868/915	2400
Débit de données (Kb/s)	250	250	500	250	250
Système supporté	TinyOS, SOS, MantisOS, Nano-RK	Contiki, TinyOS, SOS, MantisOS	Microsoft .NET Micro, Linux, TinyOS	TinyOS, Contiki, FreeRTOS	TinyOS

FIGURE 1.4: Principales caractéristiques de quelques capteurs les plus courants.

Selon les applications, les capteurs sont souvent déployés dans des environnements ouverts, accessibles et potentiellement hostiles, contrairement à un système de réseau traditionnel dans un immeuble sécurisé.

- Supports de communication :
  Les capteurs sont reliés par une connexion sans fil, en général la communication radio. Bien que la communication par des technologies telles que l'infrarouge ou l'optique peut également être considérée comme une alternative, ces médias ne sont pas toujours les plus adéquats pour les réseaux de capteurs. Pour les deux exemples donnés plus haut, il est nécessaire que l'émetteur et le récepteur ne soient pas séparés par des obstacles opaques. Cette contrainte est difficilement réalisable dans la plupart des réseaux de capteurs.
- Topologie :
  Un grand nombre de capteurs sont déployés à travers une large zone géographique aléatoirement ou de façon régulière avec une densité des nœuds qui varie selon les applications. Le nombre de capteurs peut être de l'ordre de centaines ou de milliers. Selon l'application, ce nombre peut atteindre une valeur extrême de millions et la densité peut être élevée (20 nœuds/m^3) [8].
  La position des capteurs n'est pas toujours connue ou prédéterminée. La topologie peut aussi changer suite à l'apparition et la disparition de nœuds causés par le manque de puissance, des dommages physiques, ou des perturbations de l'environnement. Des capteurs supplémentaires peuvent également être redéployés.

1.1.2 Sécurité des réseaux de capteurs

Les besoins de base en sécurité (authentification, confidentialité, intégrité, non-répudiation, protection de vie privée, etc.) pour les réseaux de capteurs sont similaires aux réseaux traditionnels. Cependant, les techniques de sécurité utilisées dans les réseaux traditionnels ne peuvent pas être appliquées directement aux réseaux de capteurs en raison des caractéristiques spécifiques présentées à la section 1.1.1. Ces caractéristiques entraînent un ensemble de vulnérabilités et donc des besoins de sécurité spécifiques. En particulier, l'environnement ouvert et accessible et les contraintes matérielles les rendent vulnérables aux attaques physiques.

Définitions des modèles d'attaques

Nous décrivons dans cette section les définitions principales concernant les modèles d'adversaires [13].

Les attaques peuvent être caractérisées en fonction de l'intention :

- l'"attaque passive" est produite par des adversaires tentant de collecter de l'information du réseau sans affecter son fonctionnement. Par exemple, une écoute passive qui rassemble des informations peut compromettre la vie privée et la confidentialité [14, 15].
- l'"attaque active" est produite par des adversaires tentant d'affecter le fonctionnement du réseau de manière partielle ou totale. Comparé à une attaque passive, le but de l'adversaire actif est de produire des attaques de déni de services (DoS *Denial-of-Service*)

[16], de perturber la communication en détruisant des liens, d'épuiser les ressources disponibles telles que la bande passante ou l'énergie, etc.

Les attaques peuvent également être caractérisées selon la position de l'adversaire par rapport au réseau :

- l'"attaque externe" est lancée depuis l'extérieur du périmètre de sécurité, par un utilisateur non autorisé ou illégitime du réseau. De nombreuses attaques externes existent telles que le brouillage radio, l'écoute des messages transmis dans le réseau, l'injection des messages falsifiés ou fabriqués, le rejeu des messages collectés, etc. [17]
- l'"attaque interne" est initiée à l'intérieur du périmètre de sécurité, c'est-à-dire par une entité qui est autorisée à accéder aux ressources du réseau, mais les utilise de manière non approuvée. Les attaques internes telles que *Selective forwarding*, *Sybil* et *Black-Grey-Worm-Sinkholes* sont décrites à la section 2.1.1.

Les attaques peuvent être caractérisées selon les capacités des adversaires :

- l' "attaquant puissant" peut avoir accès à des appareils avec plus de ressources, telles que les ordinateurs portables avec des médiums radio sophistiqués ou des machines avec des capacités de calcul importantes. Un seul attaquant pourrait être en mesure d'écouter et de brouiller l'ensemble du réseau.
- l'"attaquant ordinaire" est un nœud ordinaire avec des ressources identiques aux autres nœuds du réseau. Ils n'ont pas plus de ressources que les nœuds légitimes.

Dans cette thèse, nous considérons les attaquants "ordinaires" qui produisent des attaques "actives" et "internes".

Besoins de sécurité

Comme nous l'avons décrit précédemment, les capteurs doivent être peu onéreux et ont donc des ressources très limitées (mémoire, énergie, puissance de calcul, etc.). Les capteurs sont souvent déployés dans des environnements ouverts et accessibles présentant des risques d'attaques physiques contrairement à des réseaux déployés dans un bâtiment fermé.

La sécurité est basée sur les éléments suivants : l'authentification, la confidentialité, l'intégrité, la non-répudiation et la protection de vie privée [14, 15]. À cause du support de communication sans fil, un attaquant passif peut espionner les communications et un attaquant actif peut modifier, rejouer et reproduire des messages. La confidentialité, l'intégrité et la non-répudiation des messages sont donc nécessaires. La disponibilité doit être également requise contre les attaques du type déni de service (DoS) telles que le brouillage radio, les collisions, l'épuisement, etc. [16]. En raison de l'environnement sans surveillance, ouvert et accessible, un adversaire peut compromettre les nœuds en modifiant sa structure interne et en permettant ainsi à un adversaire d'extraire les informations secrètes telles que les clés de cryptage partagées entre les capteurs, les reproduire ou les déplacer. Nous avons donc besoin d'authentification des nœuds.

Les méthodes cryptographiques permettent d'apporter des protections de base (confidentialité, authentification, intégrité des données et non-répudiation). Cependant, en cas de compromission des nœuds, elles ne sont plus efficaces puisque les nœuds malveillants ont la connaissance des informations secrètes [18]. L'environnement de déploiement dans des zones accessibles rend les capteurs vulnérables aux attaques physiques et leur capacité limitée

facilite la compromission des nœuds. Nous avons donc besoin de solutions algorithmiques complémentaires aux solutions cryptographiques. Ce besoin de solutions algorithmiques est la principale motivation de cette thèse.

1.1.3 Applications et scénarios

La taille réduite des capteurs, leur faible coût et leur communication sans fil dans des environnements variables sans infrastructure préétablie, rendent les réseaux de capteurs très attractifs dans de nombreux domaines d'applications. Leur tâche principale est de mesurer des données relevées dans leur environnement et de les transmettre au point de collecte. Ils sont donc particulièrement adaptés pour des applications de surveillances environnementales, d'agricultures, de détection des feux de forêts, de relevés des compteurs, de bâtiments intelligents, de surveillance du trafic, etc. [8]. Dépendant du domaine d'application, certains réseaux de capteurs ont besoin de sécurité plus élevée. Les principales familles de scénarios d'applications sont les suivantes :

- la surveillance : les télé-relèves des compteurs d'eau, d'électricité, de gaz périodiquement à distance [19], ou la surveillance de température, d'humidité, de pression destinée à l'agriculture [5] ou à la surveillance environnementale telle que la pollution urbaine [4] sont des exemples connus. Dans la surveillance et la détection des événements rares tels que les feux de forêt, ou la fuite d'eau dans des bâtiments, les alertes doivent arriver en temps réel. Dans ces scénarios, les adversaires peuvent profiter des vulnérabilités qui caractérisent la communication sans fil pour empêcher des nœuds d'accéder au canal pour transmettre leurs données, ou tenter de modifier, d'injecter, ou de générer de fausses données. Les environnements sans surveillance permettent aux adversaires de causer des dégâts physiques sur les capteurs, de les compromettre ou de les détruire pour perturber les activités du réseau.
- la domotique : les détections d'instruisions dans un bâtiment, la surveillance de la température, de l'humidité et de la luminosité sont des cas de scénarios visés par la domotique [7]. La connexion sans fil est en expansion ces dernières années, ce qui permet de faire disparaître des câbles et des connexions filaires dans la maison pour apporter plus de confort, mais également pour contrôler les équipements à distance. Ces applications ont besoin de sécurité pour une communication à distance et pour la protection de la vie privée.
- les applications médicales (*Body Area Networks*) : la taille réduite des capteurs permet aujourd'hui de développer des applications de surveillance de l'état de santé des personnes à domicile. Les capteurs peuvent être déployés à l'extérieur ou à l'intérieur du corps humain pour surveiller le rythme cardiaque, la pression sanguine, la température, etc. [9]. Ensuite les données peuvent être transmises à un hôpital. Dans ces scénarios d'applications, la sécurité et la fiabilité de la transmission des données et la protection de vie privée sont des problématiques essentielles.

Les contraintes et les besoins de sécurité dépendent donc de la nature des applications, nécessitant ainsi différents mécanismes de sécurité.

Dans cette thèse, nous nous intéressons plus particulièrement aux réseaux urbains (télé-relève des compteurs de gaz/électricité/eau, bâtiments intelligents, etc.) considérés par le

projet ANR ARESA2 [3], où les capteurs peuvent être exposés aux menaces physiques telles que la destruction ou la compromission des nœuds.

1.2 Problématiques et motivations

Les réseaux de capteurs sont vulnérables à divers types d'attaques. La communication sans fil et le déploiement en environnement ouvert facilitent les attaques passives (écoute des messages échangés, analyse du trafic, etc.) et permettent de lancer des attaques actives (modification, injection, le rejeu des messages, attaques de déni de services) comme nous l'avons décrit dans la section 1.1.2.

Traditionnellement, les méthodes cryptographiques sont utilisées pour se protéger contre ce type d'attaques. Toutefois, comme nous avons discuté précédemment, en cas de compromission des nœuds ces protections ne sont plus efficaces. En compromettant un nœud, un adversaire entre à l'intérieur du périmètre de sécurité du réseau. Il peut donc en extraire des informations sensibles (clés de chiffrement, identité, adresse, etc.). Ensuite, un adversaire peut changer le logiciel ou le matériel pour produire des attaques plus complexes. L'impact d'un tel comportement malveillant peut être grave parce que la coopération des nœuds est essentielle pour l'acheminement des données. Un nœud malveillant peut empêcher la communication des nœuds en refusant la retransmission des paquets de données (*Selective forwarding*) ou en produisant des attaques réseau plus complexes (*Sybil*, *Sink-Worm-Black-Grey-holes*, etc.).

De manière générale dans cette thèse, nous considérons la compromission des capteurs et nous étudions des solutions algorithmiques, des mécanismes pour rendre le routage résilient, qui sont complémentaires aux solutions basées sur les méthodes cryptographiques.

La compromission des nœuds est l'un des problèmes majeurs de la sécurité des réseaux de capteurs, plus particulièrement pour les protocoles de communications qui se basent sur la collaboration des nœuds pour acheminer les données. Notre approche a donc pour objectif non seulement de soulever ces problématiques en proposant un concept de résilience composé d'une définition et d'une métrique, mais aussi de proposer des mécanismes pour rendre le routage résilient. La résilience reste un terme relativement nouveau dans le domaine des télécommunications, et elle est souvent confondue avec les termes proches tels que la survie du réseau [20] et la robustesse [21]. Notre objectif est donc de clarifier ce terme en soulignant les problématiques de sécurité due aux compromissions des nœuds.

À cause de la présence des nœuds compromis à l'intérieur du réseau, il devient difficile de différencier les adversaires et les nœuds légitimes du réseau. Les erreurs de communications causées par des liens instables, des fautes et des défaillances des capteurs, rendent encore plus ardu de détecter les erreurs provoquées par des attaques. Les systèmes de détection d'intrusions (IDS *Intrusion Detection System*) [22, 23, 24] sont des domaines à part entière permettant d'étudier les comportements des attaquants. Dans cette thèse, nous ne considérons pas les mécanismes réactifs basés sur la détection des intrusions, parce que cela oblige les capteurs à modifier leurs stratégies du routage en cas de détection des attaques, les rendant ainsi particulièrement prévisibles. Dans notre contexte, cela engendre trois inconvénients majeurs : (i) les capteurs ont des algorithmes de routage identiques et les nœuds

malveillants ont déjà connaissance des stratégies implémentées. (ii) la détection des attaques exige une longue phase d'observation. (iii) le faux positif peut être important pour un IDS, en particulier dans le contexte des réseaux de capteurs, en raison des liens instables et non fiables, des défaillances et des fautes des nœuds fréquentes, etc.

Cette situation rend nécessaire le développement de mécanismes de routage dynamiques et imprévisibles, permettant de supporter de tels comportements malveillants. Les mécanismes que nous proposons consistent à introduire des comportements aléatoires dans la construction et le choix des routes. Cela permet d'augmenter l'entropie [25] du réseau pour rendre les protocoles non prévisibles, mais permet aussi pour diversifier les routes. De plus, nous proposons d'ajouter de la réplication de paquets afin de profiter de cette diversification des routes et d'améliorer le succès de transmission de chaque paquet.

1.3 Contributions et organisation de la thèse

Cette thèse est formée de 7 chapitres en incluant ce premier chapitre d'introduction.

Le chapitre 2 est consacré à l'état de l'art général. Tout d'abord, nous présentons les problématiques de sécurité des réseaux de capteurs, en regroupant les attaques et les défenses connues par couche OSI et nous décrivons les protocoles de routage sécurisés basés sur les solutions cryptographiques et les solutions algorithmiques. Ensuite, nous présentons une étude préliminaire de 5 protocoles de routage classiques en cas d'attaques afin d'étudier les comportements des protocoles face aux attaques. La fin de ce chapitre est consacrée à l'état de l'art général sur la marche aléatoire et en particulier, sur la marche aléatoire biaisée, car nous verrons que le comportement aléatoire est bénéfique à la résilience.

Le chapitre 3 aborde notre première contribution, le concept de résilience des protocoles de routages des réseaux de capteurs. L'objectif est de proposer une définition pour clarifier le terme de résilience dans notre contexte et un outil, une métrique, permettant de mesurer la résilience afin de comparer efficacement les protocoles de routage. L'idée est d'utiliser une représentation graphique à deux dimensions pour avoir une vision synthétique de la résilience des protocoles selon plusieurs paramètres de performance, mais aussi de lier cette représentation graphique à une méthode de calcul quantitative pour agréger les valeurs des paramètres. Pour illustrer l'application de cette métrique dans la pratique, nous évaluons la résilience des cinq protocoles de routage classiques par des simulations.

Le chapitre 4 est consacré à notre deuxième contribution : les mécanismes résilients de routage pour les réseaux de capteurs. Suite à notre étude préliminaire des protocoles de routages classiques, nous identifions les mécanismes bénéfiques à la résilience. Les mécanismes résilients que proposons consistent en trois éléments : (i) introduire du comportement aléatoire (ii) limiter la longueur des routes (iii) ajouter de la réplication de paquets. Nous introduisons ces mécanismes aux protocoles de routages classiques et nous les évaluons par simulations selon notre métrique de résilience.

Le chapitre 5 décrit notre troisième contribution. Nous présentons une étude approfondie des variantes du protocole de routage par gradient en cas de plusieurs attaques combinées. Cette étude a pour but d'étendre l'évaluation des protocoles par simulations sur les trois points suivants : (i) l'étude de l'impact de plusieurs attaques combinées (*Sybil*, *Wormhole* et

Sinkhole avec *Selective forwarding*) (ii) l'étude des différentes valeurs de biais pour introduire des comportements aléatoires à GBR (iii) l'étude des différentes façons de répliquer les paquets.

Le chapitre 6 est consacré à notre quatrième contribution, une étude théorique de la résilience des marches aléatoires biaisées en cas d'attaques. L'objectif principal est de confirmer les résultats obtenus par simulations, mais aussi de répondre à la question suivante : jusqu'à quel seuil d'entropie l'introduction des comportements aléatoires peut bénéficier à la résilience ? Nous évaluons analytiquement la marche aléatoire biaisée en cas d'attaques selon le succès de livraison et la consommation d'énergie. Ensuite, nous introduisons deux types de réplications (uniforme et adaptative). Enfin, nous évaluons tous les scénarios selon notre métrique de résilience pour avoir une vision complète.

Le dernier chapitre conclura cette thèse. Nous présentons les principaux apports de ce travail et nous évoquons les perspectives de nos travaux.

État de l'art

2

Sommaire

2.1 Problématiques de sécurité dans les réseaux de capteurs 14
2.1.1 Attaques et défenses des réseaux de capteurs par couche OSI . . . 14
2.1.2 Protocoles de routage sécurisés 25
2.2 Étude préliminaire des protocoles de routage en cas d'attaques 27
2.2.1 Hypothèses du réseau, modèles d'adversaires et protocoles étudiés . 28
2.2.2 Évaluation des performances 33
2.3 Marches aléatoires 38
2.3.1 Marche aléatoire 38
2.3.2 Marche aléatoire avec mémoire 39
2.3.3 Marche aléatoire biaisée 40
2.4 Conclusion 41

Ce chapitre présente l'état de l'art concernant l'ensemble des domaines abordés dans cette thèse. Il est organisé en trois parties :

1. l'état de l'art des stratégies de sécurité pour les réseaux de capteurs en regroupant les attaques et les défenses par couche (modèle OSI), et en particulier de la couche réseau.
2. une étude préliminaire des protocoles de routage de différentes catégories des réseaux de capteurs soumis à des attaques réseau. Notre objectif est d'étudier les comportements des protocoles vis-à-vis des attaques. Les protocoles considérés sont : Dynamic Source Routing (DSR) [26], Gradient-Based Routing (GBR) [27], Greedy Forwarding (GF) [28], Random Walk Routing (RWR) [29] et Biased Random Walk Routing (BRWR) [2]. Ces protocoles seront étudiés en cas d'attaques de non-retransmission des paquets par les nœuds compromis (*Selective forwarding*) et de *Sinkhole*.
3. l'état de l'art de la marche aléatoire, et en particulier de la marche aléatoire biaisée, car nous montrons par la suite que les comportements aléatoires sont bénéfiques pour la sécurité du routage des réseaux de capteurs.

2.1 Problématiques de sécurité dans les réseaux de capteurs

Les stratégies de sécurité sont essentiellement liées aux attaques connues des réseaux de capteurs. Nous proposons donc de lister les attaques courantes et les mécanismes de défense prévus en les regroupant par couche ciblée, même si ces attaques peuvent influencer les autres couches de la pile protocolaire.

2.1.1 Attaques et défenses des réseaux de capteurs par couche OSI

Attaques et défenses sur les composants matériels

La compromission des nœuds est un des problèmes majeurs de la sécurité des réseaux de capteurs. Un nœud compromis permet à l'adversaire d'entrer à l'intérieur du périmètre de sécurité du réseau.

En raison du déploiement des capteurs dans un environnement ouvert et accessible, les adversaires peuvent facilement capturer les capteurs ou les détruire. La modification des capteurs (*tampering*) [18, 16] est une attaque bien connue sur les composants matériels d'un capteur. Elle consiste en la modification de sa structure interne. Cette attaque permet à un adversaire d'extraire les informations sensibles telles que les clés de cryptage partagées entre les capteurs, ou même de changer le programme du capteur. Les capteurs sont des dispositifs électroniques très simples qui peuvent faire face à des machines bien plus puissantes. Les attaques via l'interface JTAG (*Joint Test Action Group*) sont décrites dans [18]. JTAG est l'interface utilisée pour les tests de composants, tels que le port d'accès de test TAP (*Test Access Port*). Il peut donc permettre à l'adversaire de prendre le contrôle complet d'un capteur. Des attaques sont également possibles en exploitant le BSL(*Bootstrap Loader*), où l'adversaire peut lire et écrire dans la mémoire du microcontrôleur [18]. Une autre forme d'attaques est l'écoute des données échangées sur les fils de raccordement de la mémoire

externe (EEPROM) et le microcontrôleur, ce qui permet à un adversaire de lire toutes les données transférées [18].

Les adversaires peuvent ainsi répliquer les capteurs et les placer aux endroits stratégiques du réseau [30]. Ils peuvent également utiliser des équipements de détection de capteurs pour les localiser et les détruire [18, 14].

Mesures de défenses :
Les capteurs peuvent être protégés contre les attaques physiques en améliorant les composants matériels, ou en utilisant des solutions algorithmiques basées sur la coopération des capteurs.

L'utilisation de matériels inviolables (*tamper-proof*) [18] permet à chaque capteur de rendre leurs données sensibles dans sa mémoire inaccessibles, mais ces dispositifs restants relativement chers, ils ne sont en général pas présents dans les réseaux de capteurs. Cependant, pour limiter les dommages et lutter contre ces attaques, les auteurs décrivent dans [18], un ensemble des solutions. Cela consiste par exemple à désactiver l'interface JTAG, et utiliser un mot de passe pour le BSL. Une autre technique possible est d'utiliser un logiciel spécial, ou du matériel spécial placé à l'extérieur du capteur pour détecter les intrusions. Une autre technique consiste à utiliser l'auto-destruction des capteurs, si un capteur détecte une intrusion il se détruit lui même ou il efface ses données et ses clés pour éviter de dévoiler ses informations secrètes.

Les approches algorithmiques consistent à utiliser des techniques telles que le contrôle du voisinage, la vérification de la localisation et les mécanismes de routage (multi-chemins, réplication des paquets, routage aléatoire, etc.) contre la compromission des nœuds [31]. La technique basée sur la localisation consiste à s'assurer que les informations de localisation sont légitimes. Contre les attaques avec des équipements de détection des capteurs, les nœuds peuvent travailler en coopération. Ils peuvent détecter les attaquants et avertir les autres capteurs pour empêcher qu'ils soient détectés par les attaquants [14].

Attaques et défenses de la couche physique

La couche physique est responsable de la sélection de fréquence, la génération de la fréquence de la porteuse, la détection du signal, la modulation, et le codage des données [8].

Dans la littérature, nous avons recensé un ensemble d'attaques ciblant la couche physique que nous présentons ici.

- **L'écoute, l'analyse du trafic et la corruption des messages** :
  La communication par radio permet à un adversaire d'écouter passivement et d'acquérir des informations de capteurs si les messages sont diffusés en clair. Même si la communication est chiffrée, les attaquants équipés de machines puissantes peuvent collecter suffisamment d'informations pour préparer des attaques, essayer de décrypter les messages, ou causer d'autres perturbations. Un attaquant actif peut modifier le contenu du message et donc compromettre l'intégrité [15, 14].
  Mesures de défenses :
  Différentes méthodes cryptographiques peuvent être utilisées pour se défendre contre ce type d'attaques [14]. Par exemple, nous pouvons utiliser le chiffrement des messages pour assurer la confidentialité, les fonctions de hachage pour assurer l'intégrité, la si-

gnature pour la non-répudiation et l'établissement des clés pour l'authentification. La cryptographie à clé publique est considérée comme coûteuse pour les réseaux de capteurs. De manière générale, la cryptographie symétrique est considérée comme mieux adaptée. Toutefois, cette dernière apporte deux problématiques majeures concernant la distribution des clés et la préservation du secret partagé en cas de compromission des nœuds [32].

- **Brouillage radio :**
  Le brouillage radio (*jamming*) est une attaque puissante de type déni de service. Elle consiste à créer des interférences sur les fréquences radio utilisées par les capteurs, perturbant ainsi la communication du réseau. Différentes stratégies de brouillage sont présentés dans [33] comme suit :
  - constant : l'adversaire émet un signal radio en continu.
  - trompeuse : au lieu d'envoyer des bits aléatoires, l'adversaire injecte des données régulièrement sur le canal.
  - aléatoire : au lieu d'envoyer un signal radio en continu, l'adversaire alterne entre les états silencieux et d'émission pour échapper à la détection de brouillage.
  - réactive : l'adversaire émet seulement quand il détecte une activité sur le canal et s'arrête lorsque le canal est à nouveau inactif.

  Au lieu de brouiller toute la bande des fréquences, un adversaire peut simplement brouiller le canal de contrôle. C'est une stratégie plus économe en dépense énergétique de l'adversaire et une méthode de brouillage très efficace.

  Mesures de défenses :
  Les méthodes de défense classiques contre le brouillage radio se basent principalement sur l'étalement de spectre à séquence directe (DSSS *Direct Sequence Spread Spectrum*) ou le saut de fréquences (FHSS *Frequency Hopping Spread Spectrum*) [34]. Pour attaquer le saut de fréquence, les adversaires doivent être capables de suivre la séquence précise de sauts, ou de brouiller une partie assez large de la bande [16].
  Pour les attaques par brouillage du canal de contrôle, [35] propose une défense en utilisant une technique basée sur les arbres binaires. Chaque émetteur construit un arbre binaire basé sur des fréquences choisies au hasard. L'émetteur transmet à tous les récepteurs, la fréquence correspondante à la feuille de l'arbre binaire et de ses ancêtres. Plus tard, l'émetteur envoie deux messages simultanément sur deux canaux différents. Le brouillage est détecté quand l'émetteur reçoit le premier message, mais pas le second. Une approche présentée dans [36] consiste à cacher la location du canal de contrôle par l'introduction d'un codage binaire en affectant des clés à tous les utilisateurs pour se défendre contre le brouillage du canal de contrôle. Cette méthode garantit à chaque utilisateur d'obtenir un accès au canal de contrôle dans un intervalle du temps. Une autre méthode proposée dans [37] consiste à utiliser une affectation aléatoire des clés. Si chaque nœud dispose d'une clé, la compromission d'un nœud ne permet pas d'affecter les autres, mais cela demande un grand nombre de canaux. Les auteurs soulignent l'importance du compromis entre le nombre de canaux et la robustesse contre le brouillage.
  La solution présentée dans [38] contre le brouillage radio est une défense de la couche réseau. Les capteurs coopérèrent pour déterminer la région brouillée et l'isoler du reste du réseau. Quand les capteurs détectent un brouillage localement, ils alertent leurs voisins. Les voisins ayant échappé au brouillage et qui ont reçu le message d'alerte

avertissent à leur tour les autres nœuds du réseau. Grâce aux messages échangés entre les capteurs, le réseau détermine la zone brouillée et les capteurs routent les messages en évitant la zone brouillée.

Attaques et défenses de la couche de liaison de données

La couche de liaison de données est responsable du multiplexage des flux de données, de la détection de trame de données, d'accès au support et du contrôle d'erreur [8]. La couche MAC (*Media Access Control*) permet l'arbitrage de l'accès au canal des nœuds voisins. Pour éviter de transmettre en même temps et d'éviter les collisions, les nœuds écoutent la porteuse pour savoir si d'autres nœuds transmettent.

- **Collisions** :
  Les adversaires peuvent intentionnellement violer le protocole de communication, et continuer à transmettre des messages afin de créer de l'interférence et générer des collisions. Pour être plus efficaces, ils peuvent envoyer leur propre signal au moment où un nœud légitime transmet. Les collisions engendrent la retransmission d'un paquet. En théorie, provoquer des collisions dans un seul octet est suffisant pour créer une erreur de CRC (*Cyclic Redundancy Check*) et donc paralyser le message entièrement. L'avantage d'une attaque de collision par rapport au brouillage radio est qu'un attaquant va consommer moins d'énergie et sa détection est plus difficile.
  Mesures de défenses :
  Toutes les mesures de défenses utilisées contre le brouillage radio peuvent être appliquées aux attaques par collisions. Une autre méthode consiste à utiliser des codes correcteurs d'erreurs [39], qui sont efficaces dans une situation où des erreurs se produisent sur un nombre limité d'octets, mais cette solution présente également un surcoût en termes de communications et de traitements supplémentaires.
  Une méthode pour détecter les attaques par déni de service de la couche MAC dans un réseau CSMA/CA (*Carrier Sense Multiple Access with Collision Avoidance*) est proposé dans [40]. Cette méthode est basée sur le calcul de la probabilité de collisions sur un réseau qui peut être déduit de l'observation du réseau. Cela consiste à surveiller les transmissions réussies et les collisions produites et à déterminer si le nombre de collisions est normal dans ce contexte.
- **Épuisement :**
  Les attaques par l'épuisement consistent à introduire des collisions et à forcer les capteurs à retransmettre en continu jusqu'à épuisement de leur batterie. Dans [41], les auteurs décrivent les attaques par épuisement. Ces attaques consistent en la violation de l'algorithme BEB (*Binary Exponential Backoff*), la manipulation des paramètres tels que SIFS, DIFS, SIFE et la manipulation des paquets de contrôle tels que RTS (*Ready-To-Send*) et CTS (*Clear-To-Send*) dans le but de perturber l'accès au canal. Un adversaire peut demander de façon répétitive l'accès au canal avec un paquet RTS, provoquant une réponse CTS de ses voisins jusqu'à l'épuisement de leurs ressources. On appelle également cette attaque le brouillage virtuel (*virtual jamming*).
  Mesures de défenses :
  La plupart des solutions proposées dans la littérature tentent de réduire l'impact de cette attaque et non pas l'éliminer définitivement. Une mesure de défense possible est

de limiter le taux de paquets de contrôle MAC pour que le réseau puisse ignorer les demandes excessives de l'adversaire et donc d'éviter l'épuisement de la batterie des nœuds. Une autre technique est de limiter le temps d'accès au canal de chaque capteur à un temps court et de leur permettre de transmettre des données seulement durant le créneau qu'on leur a attribué. Cela permet de limiter l'utilisation prolongée du canal.

- **Brouillage de la couche liaison de données** :
  L'attaque de brouillage de la couche de liaison de données est présentée dans [42]. Dans cette attaque, l'adversaire essaie de trouver un paquet de données en transmission et il crée du brouillage. Mais comme les paquets sont générés spontanément, pour un adversaire cela devient de plus en plus difficile de prédire l'arrivée d'un paquet de données. Pour résoudre cette difficulté, un adversaire peut prévoir la distribution de la probabilité du temps entre deux paquets. Cette attaque peut être appliquée à différents protocoles MAC tels que S-MAC, B-MAC et L-MAC.
  Mesures de défenses :
  Certaines mesures de défense contre les brouillages de la couche de liaison de données sont présentées dans [42]. Ces méthodes consistent par exemple à l'utilisation d'une fonction pseudo-aléatoire basée sur le temps pour changer la taille des paquets ou au raccourcissement du préambule afin de rendre sa détection plus difficile.

Attaques et défenses de la couche réseau

La couche réseau est responsable de l'adressage, de la découverte du voisinage et du routage. Les réseaux de capteurs se basent souvent sur le routage multi-sauts, les messages peuvent traverser plusieurs nœuds avant d'atteindre leur destination. Les attaques sur la couche réseau sont résumées dans [17].

- **Attaques *Sybil*** :
  L'attaque *Sybil* décrite dans [43], consiste à prendre illégitimement de multiples identités (Figure 2.1). Le but de l'attaquant est de perturber le réseau en créant des identités non existantes ou en volant des identités à d'autres nœuds. L'attaque *Sybil* permet de remplir la table de voisinage des nœuds légitimes par de fausses identités pour simuler l'existence de plusieurs voisins alors qu'il n'y a qu'un seul nœud physique. Les nœuds *Sybil* peuvent créer une fausse topologie du réseau, attirer plus de trafic vers les nœuds *Sybil*, perturbant ainsi le bon déroulement du routage.
  Mesures de défenses :
  Pour se défendre contre l'attaque Sybil, le réseau a besoin d'un mécanisme pour s'assurer que l'identité d'un nœud physique est sa seule identité déclarée. Dans [43], les auteurs décrivent deux méthodes pour valider les identités des nœuds : directe et indirecte. Dans la validation directe, un nœud de confiance déclare directement si l'identité du nœud est valide ou pas. Dans la validation indirecte, un nœud de confiance est autorisé à témoigner pour ou contre la validité de l'identité d'un nœud. Les auteurs de [43] ont proposé des techniques de validation directes, les tests radio. Dans cette méthode, le nœud testeur attribue un canal de communication différent pour chacun de ses voisins. Ensuite, il choisit aléatoirement un canal et il l'écoute. Si le nœud détecte une transmission sur le canal, il suppose que le nœud transmetteur sur ce canal est un nœud physique. Si le nœud ne détecte pas de transmission sur le canal spécifié, le nœud suppose que l'identité

affectée au canal n'est pas une identité physique.
Une autre technique pour se défendre contre l'attaque Sybil est décrite dans [14]. Il s'agit d'utiliser des techniques de pré-distribution de clés aléatoires. L'idée principale derrière cette technique est qu'avec un nombre limité de clés, un nœud qui génère aléatoirement des identités de nœuds ne peut posséder suffisamment de clés pour créer de nombreuses identités. Il ne peut donc pas échanger de messages sur le réseau, car la fausse identité créée ne permet pas de crypter ou décrypter des messages.

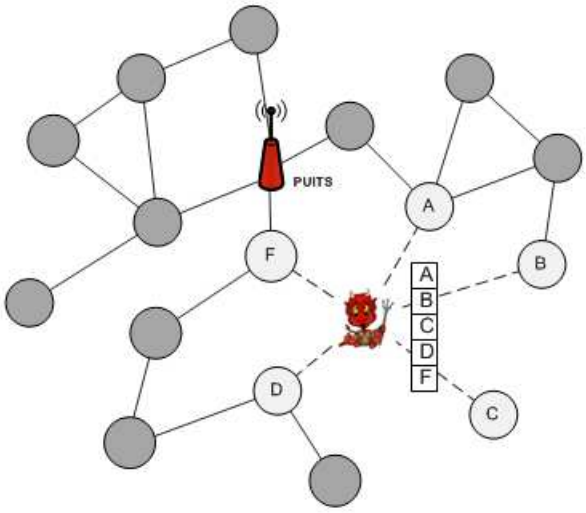


FIGURE 2.1: Attaque *Sybil*.

– **Attaques par réplication des nœuds** :
Un adversaire peut capturer des nœuds, les analyser et les reproduire. Ensuite, il peut placer les nœuds répliqués dans des endroits stratégiques du réseau (Figure 2.2). La réplication des nœuds peut causer des conséquences graves. Elle permet à l'adversaire de corrompre des données du réseau ou de déconnecter une partie du réseau [30].
Mesures de défenses :
Les défenses contre les attaques par réplication des nœuds peuvent être distinguées selon deux approches : centralisées et distribuées. Dans l'approche centralisée, chaque nœud envoie la liste de ses voisins avec leur emplacement à un nœud central, la station de base. Dès la réception, la station de base vérifie qu'aucun nœud n'est déclaré sur deux emplacements en même temps. Le principal inconvénient de cette approche est la nécessité de la présence d'une station de base en permanence qui doit vérifier et alerter l'existence des nœuds répliqués.
Dans l'approche distribuée, les nœuds peuvent compter sur leurs voisins pour détecter la réplication des nœuds au lieu d'avoir une station de base en permanence. L'utilisation d'un mécanisme de vote des voisins peut parvenir à un consensus sur la légitimité d'un nœud donné.
Dans l'approche *node-to-network broadcast*, chaque nœud diffuse à l'ensemble du réseau une déclaration de sa propre position signée. Chaque nœud conserve la déclaration de l'emplacement de ses d voisins. S'il reçoit une déclaration de position signée qui entre en conflit avec celle d'un autre voisin, il diffuse une preuve de révocation contenant les

revendications contradictoires à l'ensemble du réseau. Si n est le nombre de nœuds et d le nombre de témoins, le coût de la communication de cette méthode est $O(n)$ et utilisation de la mémoire est $O(d)$ [30].
Dans l'approche *deterministic multicast*, il y a une fonction publique déterministe F qui pour un nœud i, donne les $F(i)$ nœuds témoins. Le coût de communication est $O(d\ln dn\sqrt{n})$ utilisation de la mémoire est $O(d)$ [30].
Dans [30] les auteurs proposent deux algorithmes distribués : *randomized multicast* et *line-selected multicast*. *Randomized multicast* améliore la sécurité de *deterministic multicast* en choisissant les témoins aléatoirement. Le paradoxe des anniversaires suggère qu'il y aura au moins une collision. Le coût de la communication est $O(n)$ et l'utilisation de la mémoire est $O(\sqrt{n})$ [30].
L'algorithme de *line-selected multicast* cherche à réduire le coût de communications en choisissant comme témoins des nœuds intermédiaires entre la source et la destination. Dans les protocoles proposés dans [30], les nœuds détectent la réplication de façon passive. S'il y la réplication des nœuds, un témoin reçoit passivement deux déclarations d'emplacements contradictoires et il les utilise pour interdire les nœuds répliqués. Le coût de communications est $O(\sqrt{n})$ et l'utilisation de la mémoire est $O(\sqrt{n})$ [30].
Une nouvelle approche active est proposée dans [44]. Dans cette approche, chaque nœud teste activement si d autres nœuds choisis aléatoirement sont répliqués (on appelle ces nœuds, les nœuds examinés (*scrutinized nodes*)). L'approche active a besoin d'un nombre constant de nœuds examinés par nœud. L'utilisation de la mémoire par nœud est $O(1)$ et le coût de communications est $O(\sqrt{n})$ [44].

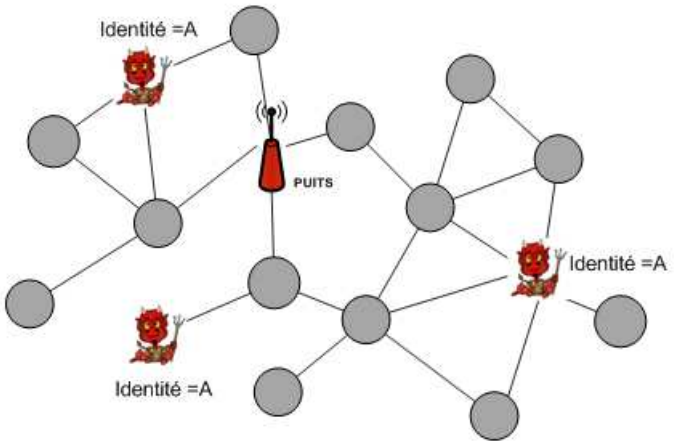


FIGURE 2.2: Attaque par réplication des nœuds.

– **Attaques *Selective forwarding*** :
Dans un routage multi-sauts les messages peuvent traverser plusieurs nœuds avant d'atteindre leur destination. Dans une attaque *Selective forwarding*, les nœuds malveillants ne remplissent pas leur rôle de routeur en jetant certains messages au lieu de les transmettre tous (Figure 2.3). Le but de l'attaquant est donc d'éliminer le plus de paquets de données possible pour perturber le bon fonctionnement du réseau. Il en existe différentes versions : le *Blackhole* consiste à attirer d'abord le trafic vers un nœud malveillant et à

éliminer ensuite tout le trafic qui passe par ce nœud et le *Greyhole* consiste à éliminer des paquets de façon sélective pour ne pas être repéré facilement.

Mesures de défenses :

Une solution possible est d'utiliser la surveillance du trafic du réseau pour s'assurer que les nœuds voisins retransmettent des messages. Dans [45], les auteurs proposent d'utiliser un système de surveillance qui identifie les nœuds égoïstes et un système *Pathrater* qui aide les protocoles de routage à éviter ces nœuds. Le système de surveillance est complété par un système de réputation proposée dans [46], où tous les voisins d'un nœud note collectivement si ce nœud exécute les fonctions demandées et s'il retransmet les messages. Une autre possibilité est d'utiliser le routage multi-chemins [47, 48]. Cependant, ces méthodes adressent essentiellement les problématiques de défaillances ou les fautes des nœuds ou des liens.

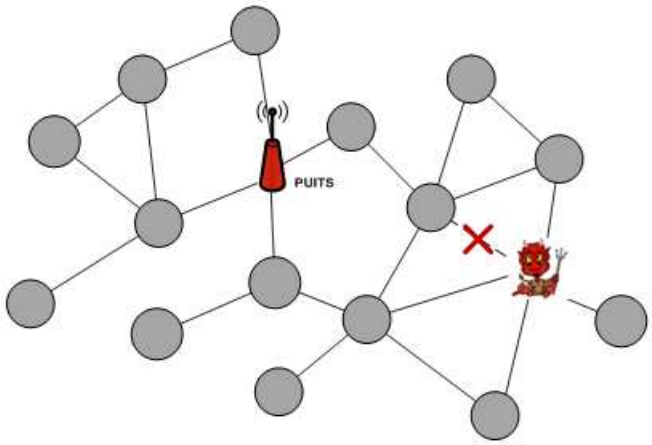


FIGURE 2.3: Attaque *Selective forwarding*.

- **Attaques *Sinkhole* :**
  L'attaque *Sinkhole* consiste à faire croire qu'un nœud malveillant est très proche du puits pour attirer la plus grande partie du trafic et l'éliminer (Figure 2.4). Cette attaque est d'autant plus pertinente que, dans le contexte de réseaux de capteurs, le trafic de données se dirige principalement vers un ou peu de puits, les points de collecte.
  Mesures de défenses :
  Une approche pour éviter les attaques *Sinkhole* est décrite dans [17]. Cela consiste à utiliser des protocoles de routage qui permettent de vérifier la fiabilité bidirectionnelle d'une route avec l'envoi des paquets ACK de bout en bout contenant l'information de latence et de qualité de la route.
- **Attaques *Wormhole* :**
  L'attaque *Wormhole* [49] consiste à créer des tunnels entre deux nœuds attaquants pour faire transiter des informations via une connexion privée (Figure 2.5). Un nœud malveillant reçoit des paquets à un endroit donné du réseau et les envoie à un autre endroit par l'intermédiaire de ce tunnel. Des nœuds, physiquement loin dans le réseau, peuvent ainsi se faire passer pour voisins. L'attaque *Wormhole* permet de créer une fausse topologie du réseau ou d'attirer plus de trafic vers un endroit donné.

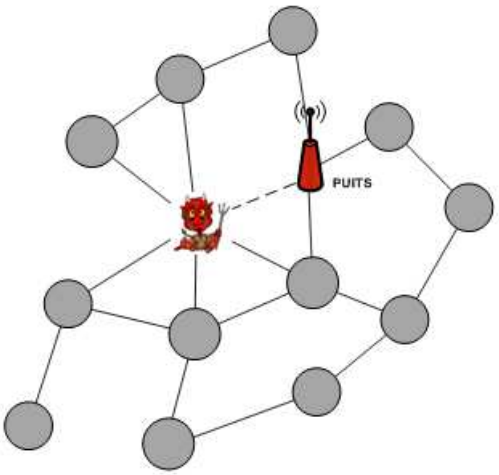


FIGURE 2.4: Attaque *Sinkhole*.

Mesures de défenses :
Les attaques *Wormhole* sont très difficiles à détecter, surtout lorsqu'elles sont combinées avec une attaque *Sinkhole*. Un mécanisme de détection du *Wormhole*, appelé *packet leashes* est introduit dans [49] consistant en deux mécanismes : *geographical leashes* and *temporal leashes*. L'idée principale de ces deux mécanismes est d'ajouter de l'information dans le paquet pour limiter leur distance maximale de transmission. Les *geographical leashes* garantissent que le destinataire du paquet est à une certaine distance de l'expéditeur. Les *temporal leashes* font en sorte que le paquet ait une limite maximale de sa durée de vie, ce qui limite la distance maximale de son déplacement.
Ces deux méthodes peuvent chacune empêcher l'attaque *Wormhole*, car elles permettent au récepteur d'un paquet de détecter si le paquet a voyagé plus loin que ce que la limitation maximale lui permet. Cependant, ces deux méthodes nécessitent l'authentification et l'intégrité des messages. Dans le cas contraire, un adversaire peut modifier ou falsifier les informations de distance se trouvant dans le paquet. Le principal inconvénient du *geographical leashes* est que les nœuds doivent être en mesure de déterminer leur position. Le principal inconvénient de *temporal leashes* est qu'il nécessite une synchronisation très précise, qui n'est pas toujours possible selon l'environnement et le matériel.
Une autre technique pour se défendre contre les attaques *Wormhole* consiste à utiliser des antennes directionnelles [50]. Quand deux nœuds sont à portée de communication, ils doivent entendre la transmission de l'autre dans la direction opposée. Équiper les nœuds par des antennes directionnelles coûte moins cher comparé aux nombreux mécanismes de localisation tel que GPS. Avec cette méthode, l'utilisation de l'énergie est plus efficace. Cependant, l'utilisation d'antennes directionnelles ne peut être mise en oeuvre dans toutes les applications. Un autre inconvénient vient du fait qu'un lien peut être perdu entre les nœuds, car la probabilité de perdre des liens dépend de la densité du réseau. De plus, cette méthode permet de détecter une seule attaque *Wormhole*.
Dans [51], reprenant le principe de QLoP [52], les auteurs utilisent des informations de voisinage pour détecter des attaques *Wormhole*. Si les deux nœuds sont déclarés comme des voisins directs et si le réseau est suffisamment dense, ces deux nœuds doivent avoir

une certaine similarité de voisinage. Dans le cas d'attaque *Wormhole*, ce n'est pas le cas puisque les deux nœuds sont physiquement éloignés et donc ils n'ont pas les mêmes voisins.

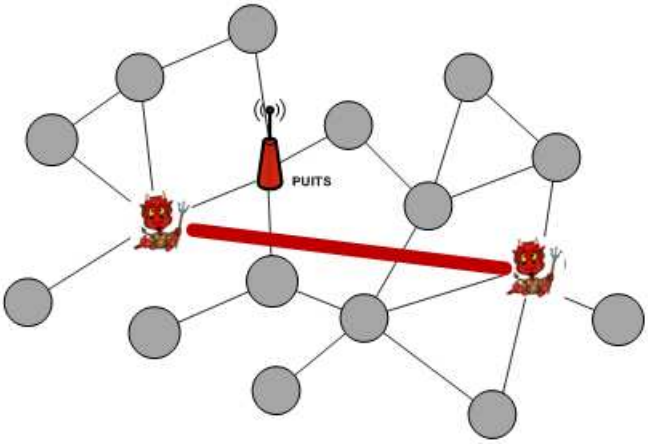


FIGURE 2.5: Attaque *Wormhole*.

- **Attaques par inondation de paquets `HELLO` :**
  De nombreux protocoles de routage utilisent des paquets `HELLO` pour la découverte de voisinage. L'attaque par inondation de paquets `HELLO` [17] consiste à convaincre, à l'aide d'un signal radio particulièrement puissant, un ensemble de nœuds de le choisir comme prochain saut. Un attaquant peut ainsi atteindre une grande partie du réseau avec ses propres messages `HELLO` annonçant de fausses informations de voisinage. Les nœuds légitimes vont tenter de transmettre leurs messages vers le nœud attaquant se trouvant hors de leur portée.
  Mesures de défenses :
  Si l'adversaire a la même capacité de réception, une façon d'éviter les attaques par inondation de paquets `HELLO` est de vérifier si les liens sont bidirectionnels [17]. Par ailleurs, l'authentification peut être une solution pour que les nœuds puissent vérifier l'identité de leurs voisins.

Attaques et défenses de la couche application

- **Attaques contre l'agrégation de données**
  L'agrégation des données peut aider à réduire la consommation d'énergie en éliminant les données redondantes dans les réseaux de capteurs. Certains nœuds du réseau agrègent les données à l'aide de fonctions d'agrégation adaptées [53] puis retransmettent le résultat agrégé à un autre nœud ou au puits.
  Les capteurs sont vulnérables aux attaques de compromission des nœuds. Avec un seul nœud compromis, l'adversaire peut injecter de fausses données pour que le résultat d'agrégation soit complètement faux.
  Mesures de défenses :
  De nombreuses techniques sont proposées dans la littérature pour sécuriser l'agrégation des données [14, 54, 15].

Dans [55], les auteurs proposent une méthode, où les données de capteurs sont retransmises directement au premier saut sans être modifiées, puis elles ne sont agrégées qu'au second saut.
Cette méthode est améliorée dans [56] en utilisant des clés paire-à-paire (*pairwise key*) à un saut et à deux sauts au lieu d'utiliser μTESLA (la version "micro" de *Timed, Efficient, Streaming, Loss-tolerant Authentication Protocol*) [57] permettant l'authentification de la source d'un message diffusé en introduisant une asymétrie par la divulgation tardive des clés symétriques à condition de considérer la synchronisation minimale d'horloges des nœuds.
Une autre méthode, présentée dans [58], est d'utiliser l'agrégation sécurisée saut par saut ce qui permet de tolérer plus d'un nœud compromis.
Une attaque, appelée *stealthy* attaque, est décrite dans [59], où l'adversaire donne de faux résultats d'agrégation sans révéler sa présence. Une technique appelée *aggregate-commit-prove* est proposée dans [59]. Ce travail est étendu par [60] aux modèles de réseaux distribués au lieu d'un modèle centralisé.
Un modèle mathématique sur la sécurité des fonctions d'agrégation est proposé dans [53]. Ce travail se base sur la théorie de l'estimation et des statistiques robustes pour quantifier la résilience des fonctions d'agrégation en cas de nœuds compromis. L'auteur affirme que certaines fonctions d'agrégation telles que la moyenne, la somme, le minimum et le maximum ne sont pas robustes par nature. Il suffit d'avoir au moins une valeur fausse pour falsifier le résultat d'agrégation. Tandis que d'autres fonctions telles que la médiane sont des fonctions d'agrégation plus robustes. Il faut falsifier au moins la moitié des valeurs pour compromettre le résultat d'agrégation. L'auteur suggère également d'éliminer des valeurs extrêmes. Cette approche se concentre uniquement sur la fonction d'agrégation à la station de base sans tenir compte de l'architecture du réseau. Cette méthode peut également fausser certains résultats par élimination des valeurs extrêmes dans certaines applications telles que la surveillance de feux de forêt, de fuite d'eau, etc., où les valeurs extrêmes (par exemple la température) portent de l'information primordiale pour détecter des événements rares.
D'autres méthodes telles que l'agrégation des données en utilisant des témoins sont proposées dans [61]. Cette méthode assure la validation des données envoyées. Dans [62], les auteurs assurent l'intégrité des données par l'utilisation de la signature pour les données agrégées. Une méthode dynamique permettant de filtrer l'injection de fausses données au long de la route est proposée dans [63].
Au cours des dernières années, le chiffrement homomorphique [54] a été l'objet de nombreuses études. Un système de chiffrement homomorphique est un cryptosystème qui permet des calculs directs des données chiffrées à l'aide d'un système efficace. Il peut être utilisé par l'agrégation de données pour fournir l'intégrité de données de bout en bout. Un protocole d'agrégation des données est proposé dans [64] où l'agrégation se base sur les chiffrements homomorphiques additifs et multiplicatifs. Ce travail est basé sur le mécanisme (PH *Privacy Homomorphism*) permettant d'effectuer des calculs directement sur les données chiffrées proposé dans [65]. Une nouvelle méthode pour sécuriser l'agrégation de données basée sur le chiffrement homomorphique est proposée dans [66]. Les auteurs utilisent une addition modulaire à la place d'un OU logique (XOR) que l'on trouve en général dans les chiffrements à flot. De cette façon, même si un agrégateur est

compromis, les messages originaux ne peuvent être obtenus par un attaquant.

2.1.2 Protocoles de routage sécurisés

Les réseaux de capteurs utilisent souvent le routage multi-sauts pour augmenter la capacité du réseau, économiser de l'énergie (augmentant ainsi la durée de vie du réseau) et réduire les interférences entre les nœuds. Dans le routage multi-sauts, les messages peuvent être acheminés par plusieurs nœuds avant d'atteindre leur destination.

Comme nous avons discuté précédemment, la compromission des nœuds permet à un adversaire d'accéder au réseau et il peut alors produire des attaques complexes pour perturber le routage. Ces attaques peuvent avoir des conséquences graves, elles peuvent permettre à l'adversaire de corrompre les données du réseau ou même de déconnecter des parties importantes du réseau. Nous identifions les défenses en deux catégories : (i) les solutions basées sur les mécanismes cryptographiques (ii) les solutions algorithmiques que nous détaillons dans les sections suivantes.

Solutions basées sur la cryptographie

Pour sécuriser les protocoles de routage, plusieurs solutions ont été proposées dans la littérature pour les réseaux ad-hoc et de capteurs. Ici, nous nous intéresserons à des solutions basées sur les méthodes cryptographiques pour se défendre contre les attaques au niveau de la couche réseau en introduisant des protocoles sécurisés. Nous pouvons citer Secure Routing Protocol (SRP) [67], ARIADNE [68], ARAN [69] qui sont proposés pour se défendre contre les attaques des protocoles de routage réactifs.

SRP [67], se focalise sur la défense contre les attaques visant le processus de découverte des routes et il garantit l'acquisition d'informations topologiques correctes. SRP permet à la source qui initie la découverte des routes de détecter et de rejeter les réponses falsifiées. Cependant, SRP ne permet pas de se défendre contre des attaques *Wormhole*, et des ententes entre nœuds malveillants qui peuvent mal acheminer les paquets de routage.

ARIADNE [68] est un protocole de routage sécurisé pour les réseaux ad hoc. Il permet de sécuriser les protocoles de routage par la source tels que le Dynamic Source Routing (DSR) [26], qui est un protocole réactif, où le processus de découverte des routes se fait uniquement au moment où la source en fait la demande. Il garantit que chaque nœud sur la route puisse authentifier la source qui initie la découverte des routes et réciproquement (que la source puisse authentifier à son tour chaque nœud intermédiaire sur le chemin vers la destination). ARIADNE fournit une authentification point à point d'un message de routage en utilisant des codes d'authentification des messages (MAC *Message Authentication Code*) et les clés partagées entre les deux parties. Pour l'authentification des paquets de contrôle tels que `RREQ` inondés par la source, ARIADNE se base sur le protocole d'authentification des diffusions des paquets TESLA [70].

ARAN [69] est un protocole de routage réactif sécurisé qui détecte et protège contre les actions malveillantes. ARAN introduit l'authentification, l'intégrité et la non-répudiation en utilisant un serveur de certificats de confiance. Le but de ARAN est de permettre de vérifier qu'une destination prévue est atteinte. Cependant, l'utilisation de la cryptographie

asymétrique rend ARAN très coûteux en termes de consommation d'énergie et de puissance de calcul. Il n'est donc pas adapté pour les réseaux de capteurs.

SEAD [71] est un protocole de routage proactif sécurisé qui se base sur le protocole Destination Sequenced Distance Vector (DSDV) [72]. L'idée de SEAD est d'authentifier le numéro de séquence et les métriques utilisées pour le routage à l'aide des chaînes de hachage à sens unique. Le récepteur de l'information de routage authentifie l'expéditeur, en s'assurant que les informations de routage proviennent du nœud légitime. Les auteurs proposent d'utiliser TESLA [70] pour l'authentification des diffusions, ou d'utiliser MAC (*Message Authentication Code*) en supposant qu'il y ait des clés partagées entre chaque couple des nœuds dans le réseau. Cependant, SEAD ne permet pas de se défendre contre les attaques *Wormhole*.

SPINS [57], est composé des deux protocoles μTESLA et SNEP. μTESLA introduit une asymétrie par la divulgation tardive des clés symétriques résultant d'un système d'authentification de diffusion efficace et adaptée aux réseaux de capteurs. SNEP assure la confidentialité des données, l'authentification bipartite et la fraîcheur des données. L'objectif de SNEP est de protéger la communication entre la station de base et les capteurs ou entre deux capteurs dans le réseau. SNEP nécessite une clé symétrique initialement partagée entre les nœuds et la station de base. Cette clé partagée permet à chaque capteur de déduire les clés de chiffrement et d'authentification. SNEP propose d'utiliser un compteur partagé entre les nœuds et la station de base pour garantir la fraîcheur des données.

Dans [73], les nœuds sont divisés en différents niveaux en fonction de leur consommation d'énergie et de leur fiabilité. Les nœuds de niveaux bas ont le rôle de mesurer tandis que les nœuds de niveaux supérieurs sont responsables du routage et de l'agrégation des données. Les auteurs proposent d'utiliser la clé symétrique basée sur la gestion des clés de groupe, où chaque nœud apporte des clés partielles pour calculer la clé de groupe.

Le principal inconvénient de ces solutions réside dans le fait qu'elles se basent sur des primitives cryptographiques pour sécuriser le routage. Comme nous l'avons souligné précédemment, non seulement elles ne sont pas efficaces en cas de compromission des nœuds (les secrets sont connus des adversaires), mais certains de ces protocoles sont conçus pour les réseaux ad-hoc et ne sont donc pas adaptés aux réseaux de capteurs (coût énergétique ou complexité algorithmique). De notre point de vue, il est nécessaire de compléter ces mécanismes par des solutions algorithmiques que nous décrivons dans la section suivante.

Solutions algorithmiques

Les solutions algorithmiques sont essentiellement basées sur la réputation des nœuds ou sur le routage multi chemins. Les solutions basées sur la réputation des nœuds consistent à choisir les nœuds avec une bonne réputation pour la construction des routes. Watchdog et Pathrather [45] sont des méthodes de routage basées sur la confiance des nœuds. Watchdog permet d'identifier les nœuds qui se comportent mal et Pathrather permet d'éviter ces nœuds. Une autre méthode *Nuglets* [74] est basée sur l'idée de récompenser les nœuds dont le comportement est conforme par une monnaie virtuelle. Les nœuds reçoivent un paiement virtuel pour transmettre un message et ce paiement est déduit de l'expéditeur. Toutefois, ces solutions basées sur la réputation des nœuds sont principalement conçues pour les réseaux

ad-hoc et ils ne sont pas efficaces quand les nœuds malveillants s'entendent pour augmenter leurs réputations mutuellement.

Les méthodes basées sur les multi-chemins améliorent la fiabilité du routage, mais ils introduisent des coûts de communications supplémentaires. Dans la littérature, certaines techniques de routage multi-chemins sont proposées afin de résister contre les pannes de nœuds telles que l'utilisation de multi-chemins disjoints [75] et non disjoints (*braided multi-path*) [47]. Une autre technique décrite dans [76] consiste à réparer les liens rompus en utilisant des informations de localisation. Ces techniques peuvent être bénéfiques pour la sécurité même si elles s'adressent la tolérance aux fautes et la défaillance des nœuds. Cependant, [77] montre que les protocoles de routage multi-chemins disjoints apportent un surcoût important en termes de consommation d'énergie même s'ils améliorent le taux de livraison des paquets.

Un protocole de routage sécurisé et configurable SIGF a été proposé dans [78], apportant de la sécurité au routage géographique IGF [79] contre les attaques externes. Dans SIGF, le choix du prochain saut se base sur le choix aléatoire, apportant ainsi le non-déterminisme au routage. Puis, les auteurs considèrent la réputation des nœuds et enfin la cryptographie. Comme nous avons décrit précédemment, les mécanismes basés sur la réputation ne permettent pas de se défendre contre l'entente des nœuds malveillants et les primitives cryptographiques ne sont pas efficaces en cas de compromission des nœuds.

Dans cette thèse, nous travaillons dans cette direction, dans le but de proposer des solutions algorithmiques pour compléter les défenses traditionnelles basées sur les méthodes cryptographiques. Notre but est de contribuer d'une manière similaire en considérant, comme dans SIGF, le non-déterminisme comme base du comportement du protocole. Dans notre cas, nous considérons des mécanismes de routage pour améliorer la sécurité des protocoles en cas d'attaques internes dues à la compromission des nœuds, où les adversaires peuvent éventuellement s'entendre entre eux.

2.2 Étude préliminaire des protocoles de routage en cas d'attaques

Dans cette section, nous étudions les comportements des protocoles de routage en présence des attaques. Les protocoles de routage étudiés appartiennent à différentes catégories afin de couvrir un ensemble représentatif. Nous introduisons des attaques réseau afin d'observer leur comportement face aux attaques internes. Notre but est de déterminer les mécanismes de routage qui permettent aux protocoles d'être plus résistants aux attaques par nature, et cela même quand les protections cryptographiques sont inefficaces.

Nous présentons d'abord les hypothèses du réseau et les modèles d'adversaires que nous considérons. Ensuite, nous décrivons les protocoles de routage étudiés. Enfin, nous présentons les résultats des simulations et les analyses.

2.2.1 Hypothèses du réseau, modèles d'adversaires et protocoles étudiés

Hypothèses du réseau

Nous considérons deux types de nœuds dans le réseau : les capteurs (nœuds sources) et le puits (point de collecte). Les capteurs sont des nœuds simples permettant de mesurer des grandeurs du monde physique (luminosité, humidité, concentration d'un gaz spécifique dans l'air, etc.). Les capteurs sont considérés comme physiquement identiques (même puissance de transmission ou de calcul, et mêmes paramètres de configuration). Ils sont déployés de manière aléatoire et uniforme avec une forte densité sur une région $N \times N$. Le puits est placé au centre de la région de déploiement (aux coordonnées $\{N/2, N/2\}$). Nous prenons les hypothèses suivantes :

- le puits est considéré comme robuste, ayant des ressources suffisantes en termes de mémoire, de calcul et d'énergie pour supporter des dispositifs cryptographiques. Les adversaires ne pourront pas compromettre le puits durant un temps fini.
- les capteurs sont des entités simples et limitées en termes de ressources (mémoire, calcul et énergie). Nous prenons l'hypothèse que les nœuds capteurs sont vulnérables aux attaques physiques et les adversaires peuvent donc les détruire ou les compromettre.

Nous considérons un graphe connexe pour représenter la topologie physique du réseau. Les paquets sont acheminés entre les sources et le puits sur cette topologie physique. Nous considérons $G(\Omega, E)$ un graphe aléatoire. Ω est l'ensemble des nœuds. E est l'ensemble des arêtes représentant les liens de communication entre les nœuds. Dans ce modèle, les nœuds sont placés aléatoirement sur un plan $N \times N$ selon une distribution aléatoire et uniforme. Un lien existe entre deux nœuds i et j si la distance euclidienne entre ces deux nœuds est inférieure à la portée de communication. Tous les nœuds ont le même rayon de transmission. Ce modèle implique que les liaisons sans fil sont bidirectionnelles, c'est-à-dire si le nœud i entend le nœud nœud j, alors la réciproque est vraie : le nœud j peut aussi entendre le nœud i.

Modèles d'adversaire

Dans cette étude, notre modèle d'adversaires correspond à des attaques "actives", "internes" et à des attaquants "ordinaires" selon les définitions présentées dans 1.1.2.

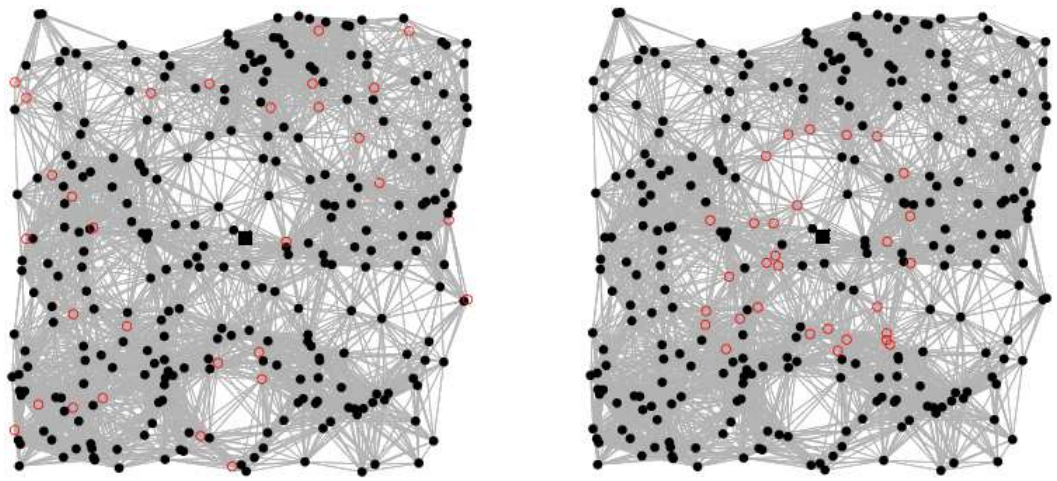
Nous considérons l'attaque *Selective forwarding*, où les nœuds malveillants ne retransmettent pas les paquets de données. Cette attaque a la particularité d'être réalisable avec tous les protocoles de routage étudiés. De plus, s'attaquant au processus de retransmission des paquets de données, elle est particulièrement pertinente pour étudier le succès de transmission des protocoles de routage.

Dans le routage multi-sauts, les paquets peuvent traverser plusieurs nœuds avant d'atteindre leur destination. Avec le *Selective forwarding*, les nœuds malveillants ne remplissent plus leur rôle de routeur et jettent une partie ou la totalité des messages transitant par eux au lieu de les retransmettre. Dans le modèle d'attaques *Selective forwarding*, la non retransmission des paquets s'applique uniquement aux paquets de données `DATA`. Nous ne considérons donc pas la non retransmission des paquets de contrôle (`RREQ`, `RREP`, `INIT` etc.).

2.2 Étude préliminaire des protocoles de routage en cas d'attaques

Dans nos simulations, nous nous focalisons sur deux scénarios de distribution des nœuds compromis. De manière aléatoire et uniforme dans l'ensemble du réseau $N \times N$ (i) dans un premier temps, et dans une petite zone $M \times M, M = N/4$ autour du puits (ii) dans un second temps.

- Scénario 1 :
  Dans ce scénario, nous supposons que l'adversaire n'a pas d'information sur l'emplacement du puits et nous considérons donc des nœuds compromis distribués à des positions aléatoires dans l'ensemble du réseau. Ainsi, un pourcentage k des nœuds est compromis (10% à 50% de la population des nœuds du réseau) et est aléatoirement et uniformément réparti sur une zone de déploiement $N \times N$. Un exemple est donné sur la Figure 2.6 (a), où les nœuds compromis sont indiqués en rouge. Les nœuds malveillants jettent les paquets de données en provenance de leurs voisins. Cependant, les nœuds malveillants continuent de générer leurs propres données au puits afin d'éviter d'être repérés.
- Scénario 2 :
  Dans ce scénario, nous supposons que l'adversaire a des informations sur l'emplacement du puits. Par souci d'efficacité, un adversaire peut tenter de compromettre les nœuds à proximité du puits. Ainsi, une zone de la taille $M \times M$ est définie (1/4 de $N \times N$), autour du puits, où un pourcentage k des nœuds compromis est distribué aléatoirement comme indiqué sur la Figure 2.6 (b). Les nœuds compromis se comportent de la même façon que dans le scénario précédent.



(a) (b)

FIGURE 2.6: Distribution des nœuds compromis (cercle) uniformément : (a) dans le réseau (b) autour du puits (carré).

Protocoles de routage étudiés

Nous classons dans [1] les protocoles de routage dans les réseaux de capteurs dans quatre catégories principales : (a) avec découverte des routes (ou gradient), (b) probabilistes, (c)

géographiques et (d) hiérarchiques (Figure 2.7).

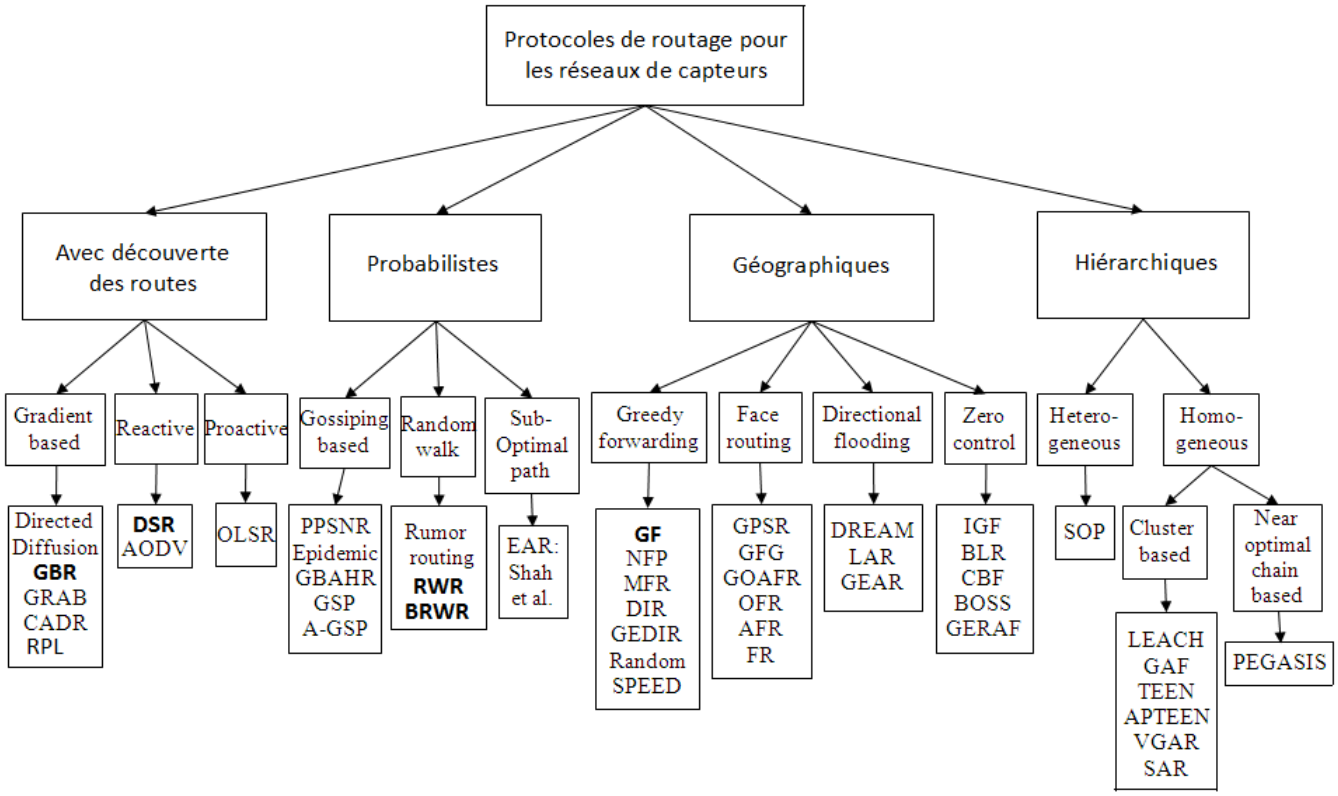


FIGURE 2.7: Catégories des protocoles de routage dans les réseaux de capteurs [1].

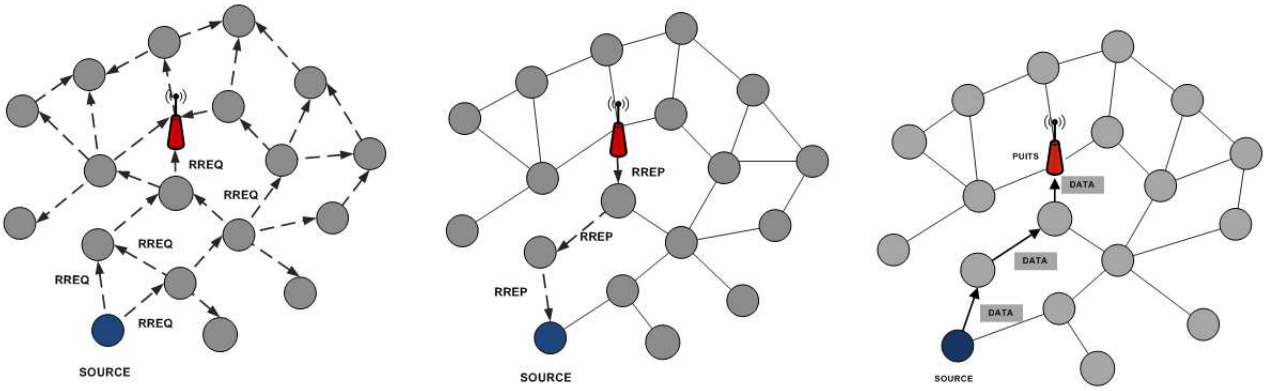
Les protocoles de routage de la catégorie (a) utilisent le mécanisme d'inondation pour découvrir les routes, maintenir l'information topologique ou configurer les gradients. Les protocoles de routage de la catégorie (b) permettent aux nœuds de choisir le prochain saut en fonction d'une certaine probabilité ou de façon totalement aléatoire. Les protocoles de routage géographique (c) utilisent l'information de localisation pour créer et maintenir une topologie de routage : chaque nœud doit connaître l'emplacement de la destination, son propre emplacement et parfois, ceux de ses voisins. Les protocoles hiérarchiques (d) sont basés sur une hiérarchie entre les nœuds : certains nœuds peuvent être physiquement différents (réseau hétérogène), ou bien, dans le cas où les nœuds sont physiquement identiques (réseau homogène), ils peuvent attribuer des rôles particuliers à certains d'entre eux avec un mécanisme d'élection.

Pour comparer le comportement des protocoles de routage confrontés à des attaques, nous avons choisi cinq protocoles candidats à partir des différentes catégories : Dynamic Source Routing (DSR)[26], Gradient Based Routing (GBR) [27], Greedy Forwarding (GF)[28], Random Walk Routing (RWR)[29] et Biased Random Walk Routing (BRWR)[2].

2.2 Étude préliminaire des protocoles de routage en cas d'attaques

Dynamic Source Routing (DSR)

DSR est un protocole de routage basé sur le mécanisme d'inondation, considérant trois types de paquets : `RREQ` (*route request*), `RREP` (*route reply*) et `DATA` (Figure 2.8). Le paquet `RERR` (*route error*) n'est pas considéré dans cette étude. Les nœuds sources inondent le réseau avec un paquet `RREQ` afin de découvrir les routes vers le puits. Les identités des nœuds traversés en chemin sont sauvegardées dans l'entête du paquet `RREQ`. Quand le puits reçoit un paquet `RREQ` d'une source pour la première fois, il répond avec un paquet `RREP` par le chemin inverse. Quand les nœuds en chemin reçoivent le paquet `RREP`, ils mettent à jour leur table de routage. Quand la source reçoit un paquet `RREP`, elle envoie ses paquets `DATA` au puits. La topologie du réseau étant considérée comme statique, le processus de découverte des routes est prévu uniquement au début de la simulation, quand un nœud source envoie son premier paquet `DATA`.



(a) (b) (c)

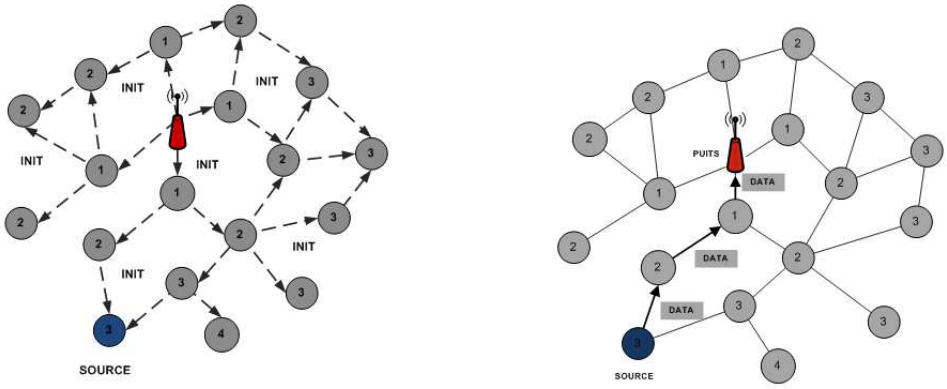
FIGURE 2.8: DSR : Envoi des paquets (a) `RREQ` (b) `RREP` (c) `DATA`

Gradient-Based Routing (GBR)

GBR est un protocole de routage basé sur le mécanisme d'inondation, considérant deux types de paquet : `INIT` et `DATA` (Figure 2.9). Le puits inonde le réseau par un paquet `INIT` afin de configurer les gradients pour chaque nœud. Le paquet `INIT` enregistre le nombre de sauts effectués à partir du puits. Chaque nœud a connaissance ainsi du nombre minimum de sauts au puits, appelé "hauteur". La différence de hauteur entre un nœud et un de ses voisins est le "gradient" de ce lien. Le processus de découverte des gradients est prévu uniquement au début de la simulation. Ensuite, à chaque saut un nœud transmet le paquet `DATA` au voisin avec le gradient minimum jusqu'à ce que le puits reçoive le paquet.

Greedy Forwarding (GF)

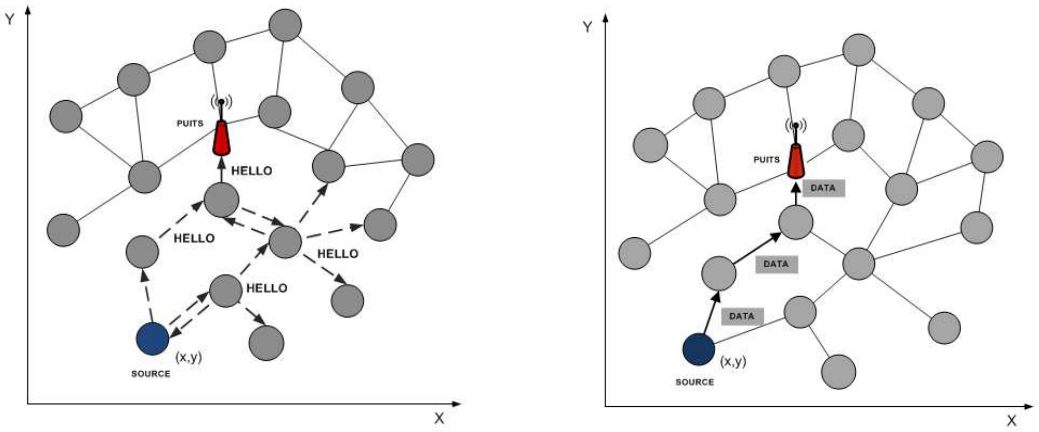
GF est un protocole de routage géographique, considérant deux types de paquet : `HELLO` et `DATA` (Figure 2.10). Chaque nœud connaît son propre emplacement et l'emplacement du



(a) (b)

FIGURE 2.9: GBR : Envoi des paquets (a) `INIT` (b) `DATA`

puits. Chaque nœud diffuse à ses voisins un paquet `HELLO` avec son identité et ses informations de localisation. Tous les voisins qui reçoivent ce paquet `HELLO` mettent à jour leur table de voisinage. Ensuite, chaque nœud transmet ses paquets `DATA` vers le voisin géographiquement le plus proche au puits en réalisant ainsi un progrès maximum en direction du puits. Nous considérons la version basique sans aucun mécanisme de contournement des trous, puisque nous considérons une topologie dense sans trou.



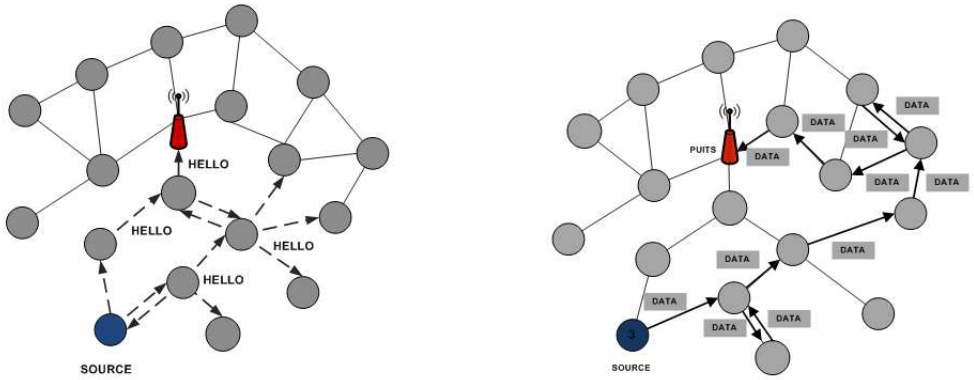
(a) (b)

FIGURE 2.10: GF : Envoi des paquets (a) `HELLO` (b) `DATA`

Random Walk Routing (RWR)

RWR est un protocole de routage probabiliste, considérant deux types de paquet : `HELLO` et `DATA` (Figure 2.11). Nous considérons une marche aléatoire basique. Chaque nœud diffuse

un paquet `HELLO` avec son identité. Tous les voisins qui reçoivent ce paquet `HELLO` mettent à jour leur table de voisinage. Un nœud envoie ou retransmet des paquets `DATA` à un voisin choisi aléatoirement et uniformément jusqu'à ce que le puits reçoive le paquet ou que la durée de vie TTL *time-to-live* du paquet soit atteinte. TTL est fixé à 32 sauts. Il est à noter que RWR peut produire des boucles de routage.



(a) (b)

FIGURE 2.11: RWR : Envoi des paquets (a) `HELLO` (b) `DATA`

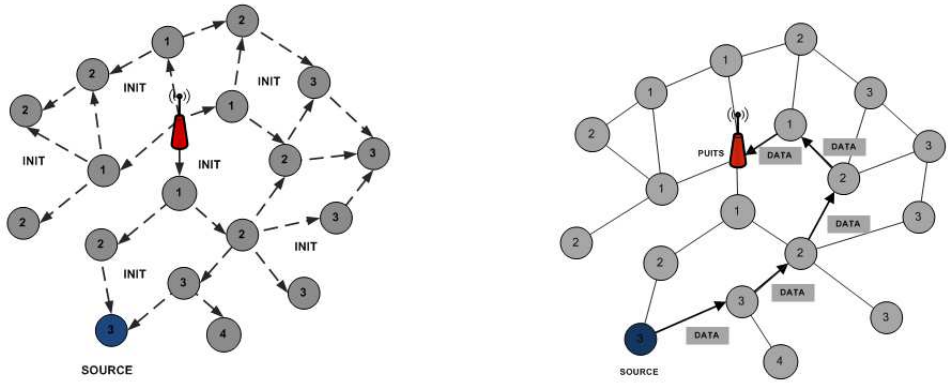
Biased Random Walk Routing (BRWR)

BRWR est un protocole de routage probabiliste, considérant deux types de paquet : `INIT` et `DATA` (Figure 2.12). Nous considérons une marche aléatoire biaisée vers le puits grâce à l'information sur la distance au puits en nombre de sauts. Cette information peut être obtenue de différentes manières. Dans cette étude, nous considérons l'envoi d'un paquet `INIT` par le puits pour établir les informations de distance en nombre de sauts (de type gradient). À chaque saut, le paquet est envoyé de manière équiprobable, soit à un voisin de niveau inférieur (le paquet progresse donc vers le puits), soit à un voisin de même niveau (le paquet ne progresse ni ne s'éloigne du puits), jusqu'à ce que le puits reçoive le paquet ou que la durée de vie TTL du paquet soit atteinte.

2.2.2 Évaluation des performances

Paramètres de simulations

Les simulations sont effectuées sur le simulateur WSNet [80]. Nous déployons 300 capteurs sur une zone 100×100 mètres de manière aléatoire et uniforme. Un unique puits est placé au centre. Les nœuds sont redéployés à différentes positions à chaque itération de simulation. La portée radio est de 20 mètres, entraînant un degré moyen par nœud de 31. La durée de vie (ou TTL) de chaque paquet est fixée à 32 sauts.



(a) (b)

FIGURE 2.12: BRWR : Envoi des paquets (a) `INIT` (b) `DATA`

Paramètres	Valeurs
Nombre de nœuds	300
Taille de la zone de déploiement	100×100m
Rayon de transmission	20m
Topologie	dist. uniforme des nœuds
Génération du trafic	dist. poissonnienne avec $\lambda = 1$ paquet/s
Nombre d'itérations	100
Temps de simulation	$100s$

TABLE 2.1: Résumé des paramètres de simulations.

WSNet donne la possibilité de définir un modèle de propagation des signaux radio avec un grand réalisme. Cependant, nous considérons des couches MAC/PHY idéales (sans interférence, sans perte due aux collisions) afin de se concentrer uniquement sur les aspects de routage et sur l'impact des attaques sur les performances des protocoles de routage.

Le trafic est généré avec une distribution poissonnienne, soit environ 1 paquet par seconde et par nœud. Le temps entre deux paquets générés suit ainsi une distribution exponentielle de paramètre λ. Le temps de simulation est de 100 secondes, et le nombre total de paquets générés est d'environ 30000. En moyenne, 100 itérations de simulations sont effectuées pour chaque scénario avec un intervalle de confiance à 95%. Le tableau 2.1 résume les paramètres de simulations.

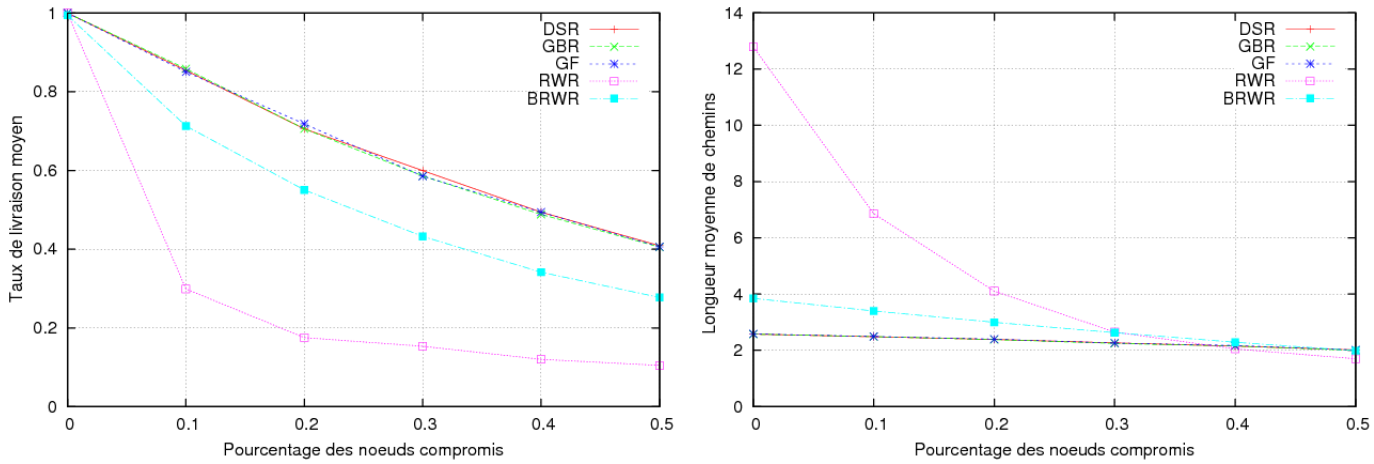
Métriques évaluées

La principale responsabilité du protocole de routage est de construire des routes afin d'assurer la transmission fiable des données dans le réseau. Pour déterminer l'impact des attaques sur des protocoles de routage, nous mesurons les deux métriques de performance suivantes :

- **Le taux de livraison moyen - ADR** *Average Delivery Ratio = nombre total des paquets reçu au puits / nombre total des paquets envoyés par les capteurs.* Il s'agit de la métrique la plus importante pour évaluer le succès de transmission des données d'un protocole de routage. Sans attaque, sans interférences et sans collisions, tous les paquets de données doivent être reçus au puits et le taux de livraison doit être égal à 1. Nous mesurons le taux de livraison moyen pour les cinq protocoles selon les deux scénarios d'attaques en faisant varier le pourcentage des nœuds compromis.
- **La longueur moyenne des chemins - APL** *Average Path Length = nombre moyen de sauts effectué par les paquets reçus.* Cette métrique permet de déterminer le nombre de nœuds traversés sur la route.

Analyses des performances

Scénario 1 :



(a) (b)

FIGURE 2.13: Scénario 1 : (a) Taux de livraison moyen (b) Longueur moyenne des chemins.

La Figure 2.13(a) présente le taux de livraison moyen des cinq protocoles de routage (DSR, GBR, GF, RWR et BRWR) sous attaques par *Selective forwarding* pour le scénario 1. Le taux de livraison moyen diminue avec l'augmentation des nœuds compromis. GBR, DSR et GF ont des résultats similaires et ils ont de meilleurs taux de livraison moyens, tandis que BRWR et RWR sont moins performants en termes de taux de livraison. Toutefois, notons que le protocole BRWR est bien meilleur en taux de livraison comparé à RWR. La différence entre RWR et BRWR par rapport aux trois autres protocoles est due au choix du prochain saut pour retransmettre les paquets de données : ils se basent sur un choix aléatoire. RWR ne privilégie pas automatiquement les chemins les plus courts, puisque chaque nœud envoie des paquets de données à un voisin choisi aléatoirement. Par conséquent, les paquets de données peuvent prendre de longues routes. BRWR se base également sur le choix aléatoire

du prochain saut, mais il privilégie la direction du puits une fois sur deux en moyenne. De manière formelle, BRWR effectue une épreuve de Bernoulli de paramètre $p = 1/2$ (équivalent à un lancé d'une pièce équilibrée). Le prochain saut choisi est en direction d'un nœud de niveau inférieur (en direction du puits) avec une probabilité p, et en direction d'un nœud de même niveau (équidistant du puits) avec une probabilité $1 - p$. BRWR a un taux de livraison significativement meilleur à RWR grâce à l'utilisation de chemins moins longs. Mais il reste moins performant que les trois autres DSR, GBR et GF. Entre BRWR et RWR, nous constatons une différence d'environ 20% en termes de taux de livraison avec 20% d'attaquants.

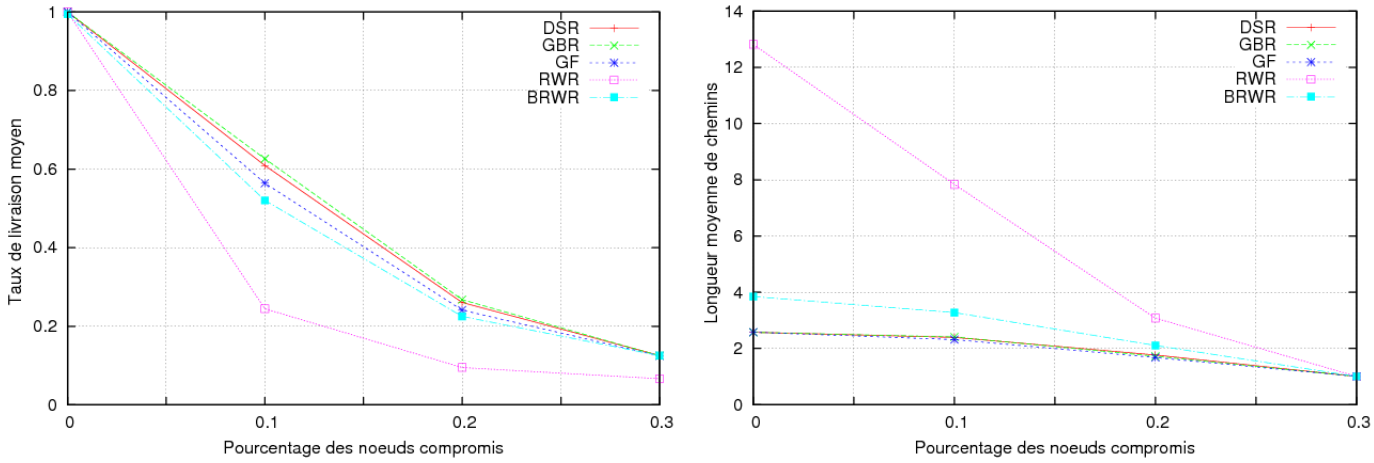
La Figure 2.13(b) présente la longueur moyenne des routes pour cinq protocoles de routage (DSR, GBR, GF, RWR et BRWR) confrontés à l'attaque *Selective forwarding* pour le scénario 1. RWR a une longueur de chemin beaucoup plus importante que les autres. Nous constatons que la longueur moyenne du chemin est inversement proportionnelle au taux de livraison moyen. Lorsque la longueur de chemin est élevée, le nombre de nœuds relais est élevé. La probabilité de rencontrer un nœud malveillant est ainsi augmentée. Nous pouvons remarquer que la longueur des chemins de RWR diminue lorsque le pourcentage de nœuds compromis augmente. Les paquets de données reçus au puits proviennent donc principalement des nœuds sources les plus proches au puits. Ce fait peut être expliqué par le fait que la probabilité de trouver un chemin sûr dans RWR décroît exponentiellement avec la longueur moyenne des chemins. BRWR a, quant à lui, une longueur de chemin significativement plus courte que RWR, puisque les paquets sont dirigés dans la direction du puits sans reculer. BRWR est donc un compromis intéressant, car il possède un meilleur taux de livraison que RWR tout en gardant un comportement aléatoire permettant de diversifier les routes.

Dans DSR, GBR et GF, les nœuds choisissent le prochain saut le plus proche du puits. Les paquets de données sont relayés, la plupart du temps, par les mêmes nœuds au centre du réseau, autour du puits. Nous observons que la stratégie de plus court chemin mène à un meilleur taux de livraison moyen, parce que les paquets ont une plus faible probabilité de rencontrer des nœuds malveillants. Cependant, ils utilisent les mêmes routes pour l'envoi de tous les messages et seulement un petit nombre de nœuds au centre du réseau participent au routage. De notre point de vue, ces protocoles basés sur les plus courts chemins ne permettent pas de résister aux attaques de non retransmission des paquets puisqu'un seul nœud compromis sur la route permet de déconnecter totalement un nœud du puits. En raison du choix déterministe du prochain saut, la redondance structurelle du réseau n'est pas exploitée. Les protocoles aléatoires, eux, ont la capacité d'exploiter cette redondance de la topologie physique.

Scénario 2 :

La Figure 2.14(a) présente le taux de livraison moyen de cinq protocoles de routage (DSR, GBR, GF, RWR et BRWR) sous attaques par *Selective forwarding* pour le scénario 2. Dans ce scénario, la répartition des nœuds compromis est localisée autour du puits menant à une attaque de type *Sinkhole*. L'impact des attaques est beaucoup plus important comparé au scénario 1. Lorsque les nœuds compromis sont près du puits, ils reçoivent plus de paquets à retransmettre que les autres nœuds du réseau. Ils peuvent ainsi attirer plus de trafic.

La Figure 2.14(b) montre la longueur moyenne des chemins des cinq protocoles pour le



(a) (b)

FIGURE 2.14: Scénario 2 : (a) Taux de livraison moyen (b) Longueur moyenne des chemins.

scénario 2. Lorsque tous les nœuds autour du puits sont compromis, le puits reçoit des paquets seulement des nœuds compromis. Les paquets de données des nœuds légitimes ne sont pas reçus. C'est pour cette raison que nous observons sur la Figure 2.14(b) une longueur moyenne des routes qui tend vers 1 dans le cas où les nœuds compromis jettent tous les paquets.

Le scénario 2 nous montre que les attaques peuvent avoir un impact très important sur le réseau si un adversaire capture du trafic des nœuds proches du puits. Quand un adversaire capture tous les nœuds autour du puits, il peut isoler le puits efficacement du reste du réseau. C'est équivalent à dire que le puits est compromis, même s'il est considéré comme physiquement robuste.

Cette étude nous montre que la plupart de ces protocoles (DSR, GBR et GF) utilisent tous les plus courts chemins pour l'efficacité en consommation énergétique. Ces protocoles sont donc déterministes. En présence d'un seul attaquant sur la route, l'utilisation d'une seule route performante conduit à la perte de tous les messages. Ainsi, les nœuds peuvent être complètement déconnectés du puits. D'un autre côté, nous avons les marches aléatoires qui réagissent différemment par rapport aux autres. RWR ne privilégie pas systématiquement les chemins les plus courts puisqu'il ne dispose pas d'information où se trouve le puits. En revanche, il choisit des chemins aléatoirement pour chaque message envoyé. Ainsi, RWR diversifie les routes et les sources ne sont pas complètement déconnectées du puis. La marche aléatoire contient donc une propriété intéressante pour la sécurité, mais complètement pénalisée par la longueur de chemins. Avec les chemins longs, la probabilité qu'un message tombe sur un attaquant augmente. La marche aléatoire biaisée BRWR est donc un compromis intéressant que nous étudierons par la suite, parce qu'elle permet à la fois de limiter la longueur moyenne des routes pour une meilleure performance et de bénéficier de comportements aléatoires.

2.3 Marches aléatoires

Notre étude préliminaire nous a permis de voir que les comportements aléatoires sont extrêmement intéressants pour la sécurité des réseaux de capteurs. Le choix aléatoire des routes rend les protocoles de routage non prévisibles pour les adversaires. Dans cette section, nous présentons donc les notions et les définitions générales des marches aléatoires qui sont essentielles pour notre étude théorique des protocoles probabilistes en cas d'attaques.

La théorie de la marche aléatoire est largement utilisée dans des domaines aussi variés que les mathématiques, la physique ou les réseaux. En partant d'un nœud du graphe, le paquet est envoyé à un voisin sélectionné aléatoirement, puis le voisin retransmet le paquet à un de ces voisins sélectionnés aléatoirement, etc. La séquence aléatoire de ces points ainsi sélectionnés est une marche aléatoire sur un graphe [81].

2.3.1 Marche aléatoire

Soit $G = (V, E)$, un graphe connexe, où V est l'ensemble des nœuds (sommets) et E est l'ensemble des arcs, i.e. les liens entre les nœuds $E \subseteq V \times V$. $n = |V|$ est le nombre des nœuds du graphe G. Nous considérons une marche aléatoire sur G : en partant du nœud v_0, si à l'étape t le paquet est au nœud v_t, le paquet marche vers un voisin de v_t avec la probabilité $1/d(v_t)$, où $d(v_t)$ est le degré de v_t. La séquence des nœuds aléatoires $v_t : t = 0, 1, 2...$ est une chaîne de Markov, M_G, où on note $P = (p_{ij})_{i,j \in V}$, la matrice de probabilités de transition de cette chaîne de Markov.

$$p_{ij} = \begin{cases} 1/d(i), & \text{si } (i,j) \in E \\ 0, & \text{sinon} \end{cases} \tag{2.1}$$

où, p_{ij} est la probabilité de se déplacer du nœud i au nœud j et $d(i)$ est le degré de i (le nombre de voisins directs de i). P permet d'étudier la probabilité de visiter chaque nœud à chaque saut. Cette probabilité est exprimée par le vecteur $q^l = \{q_1^l, q_2^l, ..., q_n^l\}$ des probabilités d'état, où q_i^l représente la probabilité que la marche aléatoire visite le nœud i au l-ème saut et $q^l = q^{l-1}P$. En prenant l'hypothèse que G est connexe et fini, alors M_G est irréductible, c'est-à-dire que n'importe quel nœud peut être atteint à partir de n'importe quel autre, et la longueur du chemin moyen entre deux nœuds est finie. Avec l'hypothèse que G est *connexe* et *non biparti*, nous sommes donc en mesure d'appliquer le théorème fondamental des chaînes de Markov [82]. Ce théorème affirme que M_G est *ergodique* et qu'il existe une unique distribution de probabilités d'état π, appelée distribution *stationnaire*, telle que $\pi P = \pi$, où π_i est définie par :

$$\pi_i = d(i)/2|E| \tag{2.2}$$

, où $d(i)$ est le degré de i et $|E|$ est la cardinalité de l'ensemble des arcs. Intuitivement, π représente l'état d'équilibre de M_G. C'est-à-dire que π représente la probabilité que le nœud i soit visité à n'importe quel saut de la marche aléatoire une fois la distribution stationnaire atteinte. Cette probabilité est proportionnelle au degré de i, $d(i)$ [83].

La théorie classique des marches aléatoires considère la plupart du temps des graphes simples, réguliers et infinis, afin d'étudier les comportements qualitatifs : la marche aléatoire retourne-t-elle sur son point de départ avec une probabilité égale à 1 [84, 85] ? Si oui, combien

de fois ? Les auteurs donnent une réponse affirmative à cette question, toutefois la réponse dépend fortement du graphe considéré. Par exemple, Pólya a prouvé en 1921 que si la marche aléatoire est effectuée sur une grille de dimension d, la marche aléatoire retourne sur son point de départ infiniment souvent si $d = 2$, mais si $d \geq 3$, seulement en nombre fini de fois [86].

Plus tard, des études ont été menées sur la marche aléatoire sur des graphes finis, afin d'étudier leurs propriétés de façon quantitative : combien de sauts effectue une marche aléatoire avant de retourner sur son point de départ ou sur un nœud particulier ; combien en fait-elle avant de visiter tous les nœuds [87, 88, 89]. Les résultats dépendent du type de graphe considéré, mais également de sa dimension.

2.3.2 Marche aléatoire avec mémoire

Une partie des études sur la marche aléatoire considère l'introduction de mémoire dans cette dernière (*self-avoiding walk*) [90, 91, 48]. Dans [91], un marcheur (ou paquet) est doté d'une mémoire de taille finie appelée *tabu list* contenant une partie des nœuds déjà visités. C'est une façon de diriger la marche aléatoire en évitant de faire marche arrière et de repasser par des nœuds déjà visités récemment. Cela permet d'améliorer l'efficacité de la marche aléatoire standard en diminuant la longueur moyenne des routes. L'avantage de ces méthodes est d'éviter l'utilisation d'information de contrôle (établissement du gradient, découverte de routes, information géographique, etc.) pour obtenir la direction de la destination, tout en améliorant la performance de la marche aléatoire standard. Toutefois, ces méthodes introduisent une nouvelle vulnérabilité à cause de l'information dans l'entête du paquet. Dans le contexte où la compromission des nœuds du réseau est considérée, les nœuds malveillants peuvent falsifier cette information. Le paquet peut voyager plus longtemps dans la mauvaise direction, dépensant ainsi les ressources énergétiques du réseau. Dans [48], les auteurs considèrent des nœuds menteurs dans le réseau et proposent des méthodes hybrides entre la marche aléatoire standard et la marche aléatoire dirigée. Ces méthodes sont rendues possibles grâce à l'information donnée par des voisins qui peut être potentiellement fausse.

Le principal inconvénient de ces méthodes réside dans le fait que l'information de direction n'est pas établie une fois pour toutes, mais dépend des autres nœuds du réseau tout au long de l'acheminement des données. Ces méthodes peuvent être plus adaptées aux réseaux distribués et dynamiques comme les réseaux ad-hoc, où les nœuds n'ont pas la possibilité d'avoir une information de direction de chaque destinataire en raison de la mobilité ou le nombre important des nœuds.

Dans le contexte de sécurité des réseaux de capteurs, où nous considérons des capteurs statiques et un seul ou peu de points de collecte, il est plus adapté d'établir l'information de la direction du puits une fois pour toutes (ou mise à jour peu fréquemment), si possible de façon sécurisée. Par exemple, le puits peut inonder le réseau par un paquet de contrôle avant le routage pour établir l'information de distance en nombre minimum du sauts (du type gradient). De cette manière, tout au long de l'acheminement des données, l'information de direction n'est pas mise en cause. En prenant l'hypothèse que les nœuds ont déjà une information fiable sur la direction du puits, la marche aléatoire biaisée est plus intéressante dans notre cas. Le biais permet de diriger les données vers le puits en améliorant l'efficacité du routage tout en gardant le comportement aléatoire. De plus, aucune information n'est

mise à jour dans l'entête du paquet de données tout au long de l'acheminent.

2.3.3 Marche aléatoire biaisée

Les travaux sur les marches aléatoires biaisées [92, 93, 94] étudient l'influence des biais sur les comportements aléatoires, ou processus stochastiques, afin de découvrir comment un biais influence la performance en termes de longueur des chemins (appelée également *hitting time* ou *commute time* quand l'aller et le retour sont considérés).

Une définition générale de la marche aléatoire biaisée est introduite dans [92] en considérant un graphe G régulier et fini de degré d. Chaque étape de la marche aléatoire est précédée par un lancé d'une pièce équilibrée $(\epsilon, 1-\epsilon)$. Avec une probabilité $1-\epsilon$, un des d voisins est choisi aléatoirement selon une distribution uniforme, et le marcheur se déplace vers ce voisin. Avec une probabilité ϵ le marcheur choisit le voisin vers lequel il se déplace. En d'autres termes, si la matrice de probabilité de transition initiale de la marche aléatoire est Q, alors la matrice de transition modifiée de la marche aléatoire biaisée P est définie comme suit :

$$P = (1 - \epsilon)Q + \epsilon B \tag{2.3}$$

où B est une matrice stochastique arbitraire restreinte aux arêtes de G choisie par le décideur. Si $\epsilon < 1/d$ alors le processus sera dominé par la stratégie du décideur. Ce processus est un cas particulier d'un processus de décision de Markov, où un décideur fait le choix dans un ensemble d'actions disponibles pour biaiser le comportement d'une chaîne de Markov.

La performance en longueur moyenne des chemins (*mean data gathering delay*, équivalent du *hitting time* dans le contexte de collecte de données dans un réseau de capteurs) de la marche aléatoire biaisée est étudié dans [2] en considérant un tore (une grille infinie) de dimension $N \times N$ comme support. La notion d'entropie y est introduite, associée à la probabilité de choisir le prochain saut. Plus la marche possède d'information, moins il y a d'entropie. L'entropie maximale correspond donc à la marche aléatoire sans biais, puisqu'aucune information d'état (distance géographique, gradient, etc.) n'est disponible et que la probabilité de choisir le prochain saut est uniforme.

Au contraire, plus il y a de l'information concernant la direction de la destination, plus nous introduisons du biais pour diriger la marche vers la destination. Le biais permet d'arriver plus rapidement à la destination tout en préservant le comportement aléatoire. Dans [2], les auteurs estiment quantitativement l'influence de la connaissance (biais) pour le routage. Ceci est motivé par le fait que la prise de décisions appropriées pour transmettre des données dépend de la quantité d'informations d'état qu'un nœud détient. Sans ces informations d'état, les nœuds transmettraient des données aveuglément à un voisin choisi aléatoirement. Si certaines informations d'état sont disponibles au niveau des nœuds, nous pouvons biaiser la marche aléatoire avec une direction privilégiée pour améliorer les performances. Les études dans [2] permettent donc d'obtenir la longueur moyenne des chemins entre les nœuds capteurs et le point de collecte en fonction du biais.

Notre étude théorique se base sur ces résultats pour étudier l'impact des attaques de non retransmission des paquets sur la performance des marches aléatoires biaisées afin de déterminer l'influence des biais en cas d'attaques.

2.4 Conclusion

Dans ce chapitre, nous avons présenté tout d'abord, l'état de l'art général sur la sécurité des réseaux de capteurs en regroupant les attaques et les défenses par couche de la pile protocolaire, en insistant en particulier sur celles de la couche réseau. Cette étude bibliographique identifie les problématiques majeures de la sécurité des réseaux de capteurs, dont les caractéristiques entraînent des vulnérabilités spécifiques. En particulier, l'environnement ouvert et accessible permet aux adversaires de compromettre des capteurs. Nous avons donc souligné le fait que la compromission des nœuds ouvre la porte aux nombreuses attaques complexes de la couche réseau telles que *Sybil*, *Selective forwarding*, *Sinkhole*, *Wormhole* etc. Dans ce cas, les protocoles sécurisés utilisant les primitives cryptographiques ne sont plus efficaces, puisque les adversaires détiennent déjà les informations secrètes. Il est donc nécessaire d'aller au-delà des méthodes cryptographiques traditionnelles et développer des solutions algorithmiques complémentaires.

Ensuite, nous avons présenté une étude préliminaire de plusieurs protocoles de routage de différentes catégories en cas d'attaques. Le but de cette étude était d'identifier les mécanismes qui permettent aux protocoles d'être plus résistants aux attaques. Cette étude souligne les éléments suivants :

- La plupart des protocoles (DSR, GBR et DF) se basent sur la critère du plus court chemin pour une meilleure efficacité. Ces protocoles sont donc déterministes, parce que les capteurs utilisent toujours les meilleures, mais les mêmes routes pour envoyer leurs données au puits. Ces comportements déterministes rendent les protocoles particulièrement prévisibles pour les adversaires.
- Les protocoles aléatoires (RWR, BRWR) sont, en revanche, extrêmement intéressants pour la sécurité parce que les comportements aléatoires les rendent non prévisibles. Cependant, la marche aléatoire classique (RWR) est trop pénalisée par la longueur de ses routes : en cas d'attaques, le succès de livraison est inversement proportionnel à la longueur des routes.
- La marche aléatoire biaisée (BRWR) est un bon compromis parce qu'elle permet à la fois d'obtenir des routes plus courtes que la marche aléatoire standard, tout en bénéficiant de comportements aléatoires.

Enfin, nous avons présenté l'état de l'art général des marches aléatoires et, en particulier, des marches aléatoires biaisées. Les notions et les définitions générales que nous introduisons dans cette dernière section sont essentielles à notre étude théorique des protocoles aléatoires en cas d'attaques que nous présentons dans le chapitre 6.

Ce travail préliminaire nous a permis de mettre en évidence, à la fois le besoin de définir un cadre d'approche algorithmique pour la sécurité de réseaux de capteurs (cadre que nous proposons dans le chapitre suivant 3), mais aussi la nécessité de proposer des mécanismes rendant les protocoles plus résistants aux d'attaques internes (mécanismes que nous présentons dans le chapitre 4).

Concept de résilience : Définition et Métrique

3

Sommaire

3.1 Introduction **43**
3.2 Motivations **43**
3.3 Définition de la résilience **44**
3.4 Métrique de résilience **46**
3.4.1 Méthode d'agrégation des multiples paramètres 47
3.4.2 Paramètres de résilience pour les réseaux de capteurs 49
3.5 Application de la métrique aux protocoles de routage classiques **52**
3.5.1 Modèle d'adversaires 52
3.5.2 Évaluation des performances 53
3.6 Conclusion **56**

3.1 Introduction

Dans le domaine des réseaux et télécommunications, la résilience des protocoles de routage est un concept encore nouveau. Dans la littérature, peu d'études [95, 96, 97, 98] tentent de définir une métrique pour quantifier la résilience des protocoles de routage. Nous présentons dans ce chapitre notre concept de la résilience en cas d'attaques internes. Nous proposons notre définition de la résilience dans le contexte de sécurité des protocoles de routage destinée aux réseaux de capteurs et un nouvel outil, une métrique permettant de mesurer la résilience. Grâce à cette métrique, nous pouvons comparer efficacement les différents protocoles de routage en terme de résilience et identifier les mécanismes de routage bénéfiques à la résilience.

Dans notre étude préliminaire 2.2, nous avons étudié par simulations des protocoles de différentes catégories selon plusieurs métriques connues telles que le taux de livraison moyen, la longueur moyenne des chemins, etc. Cependant, évaluer les protocoles selon plusieurs métriques séparément ne permet pas d'avoir une vision globale, synthétique et de discerner les différentes tendances et les compromis. Pour cette raison, nous avons besoin d'une métrique qui regroupe tous les paramètres importants pour la résilience. Notre motivation dans ce chapitre est donc celle-ci : obtenir une seule métrique quantifiable en agrégeant les paramètres essentiels à la résilience, sans perdre d'information sur chaque paramètre.

Nous proposons donc une métrique permettant d'agréger de multiples paramètres en se basant sur une représentation graphique à deux dimensions. Nous représentons la résilience par la surface d'un polygone régulier (équiangle), où chaque sommet représente un paramètre. Cette méthode donne la possibilité de lier la représentation graphique au calcul quantitatif. La représentation graphique nous donne une information qualitative permettant de visualiser le profil des protocoles et leurs tendances (plutôt économe en énergie, plutôt meilleur en délai, etc.), mais également l'impact des attaques sur chacun des paramètres. Le calcul de la surface nous donne une information quantitative en permettant de classifier les protocoles et de déterminer un ordre entre eux.

Pour illustrer la mise en œuvre de cette métrique, nous proposons son application à plusieurs protocoles de routage classiques de l'étude préliminaire (DSR [26], GBR [27], GF [28], RWR [29] et RWRB [2]). Ces protocoles seront étudiés en cas d'attaques de non-retransmission des paquets par les nœuds compromis (*Selective forwarding*).

3.2 Motivations

La résilience est un terme initialement défini en mécanique pour caractériser la capacité mécanique des matériaux à résister aux chocs [99]. Ce terme est également utilisé dans plusieurs domaines tels que la psychologie, l'écologie et l'économie [100]. Dans le domaine des réseaux et des télécommunications, la résilience a été initialement définie en considérant la tolérance aux pannes [101, 102]. Toutefois, ces études ont considéré principalement la connectivité de la topologie du réseau sans prendre en compte les mesures de performances liées au routage ni les aspects de sécurité. Des définitions de la résilience liées à la sécurité ont été proposées pour la distribution de clés de cryptage [103, 104]. Ces définitions traitent

spécifiquement des primitives cryptographiques.

Plus récemment, des travaux sur la résilience d'Internet [105, 106, 95] ont considéré conjointement plusieurs domaines tels que la tolérance aux pannes [101, 102], la sécurité [103, 104], la capacité de survie [20, 107, 21] et la sûreté de fonctionnement [108, 109]. Cette définition intègre plusieurs domaines et plusieurs termes, en rendant le cadre très général.

Notre but dans ce chapitre est de définir un cadre de résilience plus spécifique aux études des problématiques de sécurité de routage des réseaux de capteurs, en particulier quand les nœuds sont compromis par des adversaires. Comme discuté précédemment dans 1, pour résoudre les problématiques de sécurité de routage, les réseaux prévoient des protections contre les attaques venant de l'extérieur du réseau. Les protections de base consistant à recourir à des méthodes cryptographiques pour assurer l'authentification, l'intégrité, la confidentialité et la non-répudiation ne sont plus efficaces en cas de compromission des nœuds, puisque les informations secrètes sont divulguées. La compromission des capteurs devient donc la source de nombreuses attaques complexes du réseau difficilement détectables, puisqu'il devient difficile de distinguer quels sont les nœuds légitimes ou malveillants à l'intérieur du réseau.

Nous avons donc besoin d'une part, d'une définition concrète permettant d'éclaircir le terme de résilience et d'autre part, d'une métrique permettant de mesurer la résilience.

3.3 Définition de la résilience

Nous définissons la résilience d'un protocole de routage en nous inspirant de la définition initiale de la résilience en mécanique. En mécanique, la résilience est mesurée à l'aide d'un pendule muni d'un marteau de masse m fixé à son extrémité (essai de Charpy [110]). Ce pendule peut tourner dans le plan vertical autour d'un axe horizontal. Le matériau à tester se trouve au point le plus bas de la trajectoire du marteau. Pour effectuer un essai, le marteau est lâché d'une hauteur initiale h_0 sur le matériau d'essai (éprouvette normalisée). On effectue une série de lâchers en augmentant progressivement la masse jusqu'à la rupture du matériau (la hauteur initiale h_0 reste inchangée). Lors de la rupture du matériau (déformation de la structure interne), le pendule est remonté jusqu'à une hauteur h (Figure 3.1). La résilience est alors déterminée par : $W = mg(h_0 - h)$, où g est l'accélération gravitationnelle. Cela représente l'énergie absorbée par la rupture du matériau. Par conséquent, la résilience représente la capacité des matériaux à absorber le choc.

Dans le contexte de sécurité des réseaux de capteurs, nous introduisons une analogie liée à cette définition de la résilience en mécanique. Les "matériaux" à tester correspondent aux protocoles de routages que l'on veut évaluer. Le "choc" introduit par la masse m correspond aux attaques du réseau dues aux compromissions des nœuds. Nous soulignons cette analogie entre le choc qui modifie la structure interne des matériaux et les attaques internes qui, elles aussi, modifient la structure du réseau.

De la même manière que la résilience mécanique mesure la capacité des matériaux à subir et à absorber un choc, la résilience que nous définissons dans ce chapitre mesure la capacité du réseau à absorber la dégradation des performances par des attaques internes.

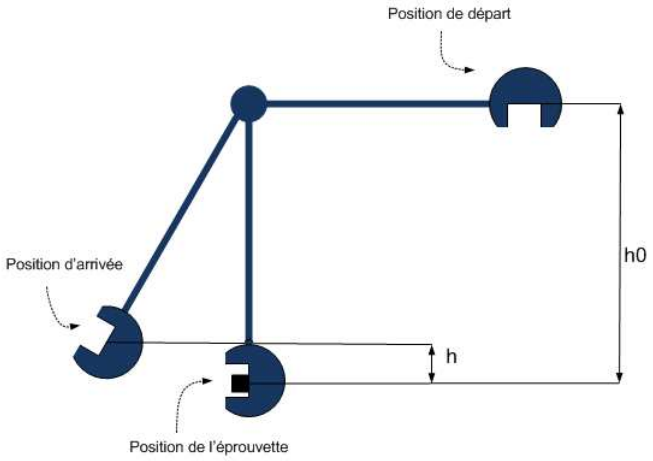


FIGURE 3.1: Essai de résilience de Charpy.

Definition 3.3.1. La résilience d'un protocole de routage est sa capacité à absorber la dégradation des performances en présence des attaques internes dues aux compromissions des nœuds. C'est la capacité de continuer à livrer des messages avec l'augmentation du nombre de nœuds non fiables ou compromis dans le réseau. En d'autres termes, c'est la capacité d'un protocole de routage à subir et à surmonter la présence de nœuds non fiables ou compromis.

Nous proposons de clarifier les définitions des termes proches en comparaison avec notre définition de la résilience.

La **capacité de survie** (*survivability*) est la capacité d'un système informatique à fournir des services essentiels en présence d'attaques ou de pannes, et de se remettre (*recover*) entièrement [20]. Dans [107], les auteurs appliquent cette capacité de survie à la tolérance aux intrusions [111]. Ils affirment que pour une meilleure survie, il est nécessaire de considérer ensemble les approches préventives, réactives et tolérantes.

La capacité de survie a des similitudes évidentes avec notre définition de la résilience. Cependant, elle diverge sur un point essentiel : la compromission des nœuds. Nous insistons sur la présence d'attaques internes lorsqu'une partie des nœuds légitimes est compromise. Cette présence de nœuds compromis engendre une modification de la structure interne du réseau. La définition de survie, quant à elle, ne différencie pas les attaques externes et internes. Elle correspond donc à une vision plus traditionnelle et générale des problématiques de sécurité. Sur cette base, les auteurs proposent des protections classiques telles que les mécanismes cryptographiques. De notre point de vue, il nous parait important de compléter ces protections cryptographiques par des mécanismes algorithmiques, indispensables dans le cas où les attaquants ont déjà les informations secrètes.

La **robustesse** (*robustness*) est définie dans [21] comme la capacité à résister aux défaillances matérielles et logicielles, aux liens asymétriques et unidirectionnels et à la limitation de la portée de communication. La définition de la robustesse se focalise principalement sur les défaillances du système et ne traite pas les aspects concernant la sécurité du réseau, ni les mécanismes résistants aux compromissions des nœuds.

Cependant, certaines attaques telles que la non-retransmission des paquets de données (*Selective forwarding*) peuvent engendrer les mêmes conséquences que la défaillance des nœuds, car les paquets de données n'arrivent pas à la destination. Nous avons donc un lien évident entre les mécanismes tolérants aux défaillances des nœuds et les mécanismes résilients aux attaques de non-retransmission des paquets parce que dans les deux cas les paquets de données sont perdus.

Les mécanismes robustes aux défaillances et aux fautes des nœuds peuvent être bénéfiques en cas d'attaques de non-retransmission des paquets et la réciproque est également valable. Toutefois, les problématiques de sécurité ne s'arrêtent pas à ce type d'attaques, puisque la compromission des nœuds ouvre la porte à des attaques plus complexes telles que *Wormhole*, *Sinkhole*, *Sybil,* etc.. Nous devons également considérer ces attaques pour évaluer leur impact sur la performance des protocoles. Nous ne pouvons donc nous limiter à cette définition de la robustesse.

Une définition de la résilience des fonctions d'agrégation est proposée dans [53]. Une fonction d'agrégation f est (k, α)-résiliente (en considérant une distribution paramétrée $p(X_i \mid \theta)$), si $rms^*(f, k) \leq \alpha \cdot rms(f)$, où $rms^*(f, k)$ est l'erreur quadratique moyenne (*root mean square error*) de la fonction f en cas de k attaquants et $rms(f)$ est l'erreur quadratique moyenne sans attaque. En d'autres mots, une fonction d'agrégation f est (k, α)-résiliente, pour une petite valeur de α, si elle peut être calculée de façon sécurisée en présence de k nœuds compromis.

Notre proposition de définition de la résilience est proche de cette définition même si le domaine d'application est différent. Comme l'auteur de [53] compare la résilience des fonctions d'agrégation en présence des nœuds compromis, notre but est de comparer la résilience des protocoles de routage en présence des nœuds compromis. L'auteur de [53] démontre par exemple que la médiane est une fonction plus résiliente que la moyenne. Parce que pour la fonction médiane, pour falsifier le résultat d'agrégation, il faut compromettre au moins la moitié des nœuds. Tandis que pour la moyenne, un seul nœud compromis suffit pour falsifier le résultat de l'agrégation. Dans notre contexte, nous n'avons pas la possibilité de représenter un protocole de routage par une seule fonction mathématique. Nous ne pouvons donc pas appliquer directement cette définition car le routage est un processus complexe. Nous pensons qu'il est nécessaire de considérer plusieurs paramètres de performance pour exprimer la résilience.

3.4 Métrique de résilience

Notre objectif est d'exprimer la résilience avec une seule valeur scalaire. Il ne s'agit pas d'un problème trivial en raison des nombreux paramètres entrant en compte pour exprimer la résilience des protocoles de routage. Quelles mesures définissent la résilience ? Comme nous l'avons vu dans la section précédente, peu d'études [95, 96, 97, 98] tentent de définir une métrique pour caractériser la résilience d'un protocole de routage.

Le taux de livraison moyen des paquets à lui seul n'est pas suffisant pour représenter la résilience. Nous devons également considérer l'équité de livraison, l'efficacité en termes de la consommation d'énergie, de débit et de délai de livraison. Nous détaillons ces paramètres

de résilience essentielles dans le contexte des réseaux de capteurs et nous en justifions la pertinence dans la section 3.4.2.

Pour obtenir une seule métrique de la résilience, tous ces paramètres doivent être agrégés de façon à obtenir une seule valeur. La somme, la somme pondérée ou une fonction logique permettrait d'agréger les valeurs des différents paramètres. Cependant, cela conduit à une perte d'information sur chacun des paramètres et empêche de discerner les différents compromis et les différentes tendances. Nous pouvons également recourir à une représentation graphique à n dimensions. Cependant, au-delà de trois dimensions cela limite la lisibilité et la possibilité d'avoir une vision globale.

Pour ces raisons, nous proposons une représentation graphique à deux dimensions (voir la Figure 3.2) et une méthode d'agrégation pour obtenir une seule valeur quantifiable. La résilience des protocoles en cas d'attaques est exprimée par la surface du polygone avec multiples axes. Plus la surface est large après les attaques plus le protocole est résilient. Une grande surface en cas d'attaque implique que le protocole est résilient et ses paramètres de performance souffrent peu et il a donc une bonne capacité à résister aux attaques (voir la Figure 3.2).

Nous montrerons que l'ordre des axes du polygone a un impact sur le calcul de la surface. En d'autres mots, pour un ensemble de valeurs identiques, leur ordonnancement donnera une surface différente. Le choix de l'ordre des axes est crucial dans le calcul de la résilience, car il en résulte une pondération implicite. Nous considérons donc un ordre spécifique des axes pour exploiter cette propriété et donner un sens à l'agrégation (voir la Figure 3.4).

Dans la section suivante, nous détaillons d'abord la méthode d'agrégation des multiples paramètres. Cette méthode est générale et elle peut s'appliquer à différents domaines. Ensuite, nous appliquons cette méthode dans le contexte spécifique des réseaux de capteurs et nous détaillons le contenu de notre métrique, les paramètres de résilience.

3.4.1 Méthode d'agrégation des multiples paramètres

Pour généraliser et prendre en compte les m métriques caractérisant les différents aspects de la performance du protocole, nous avons besoin d'un vecteur de performance P à m dimensions pour chaque protocole.

Pour l'évaluation de n protocoles de routage, nous prenons P_i représentant le vecteur de performance du protocole de routage i selon m métriques pertinentes : soit $P_i = \{p_{i,j}\}$, $i = 1, 2, ..., n$ et $j = 1, 2, ..., m$. Un protocole de routage résilient doit avoir la capacité d'absorber la dégradation de performance due à l'augmentation des attaquants dans le réseau. Intuitivement, une forte résilience correspond à une faible dégradation des performances face aux attaques, où k exprime le pourcentage de nœuds compromis dans le réseau.

Supposons $P_i(k) = \{p_{i,j}(k)\}$, le vecteur de performance du protocole i, en présence d'un pourcentage k de nœuds compromis. Pour chaque $p_{i,j}(k)$, exprimé dans sa propre unité, la formule suivante est utilisée pour obtenir une valeur normalisée :

$$p'_{i,j}(k) = \frac{p_{i,j}(k) - \min(p_{1,j}(k), ..., p_{n,j}(k))}{\max(p_{1,j}(k), ..., p_{n,j}(k)) - \min(p_{1,j}(k), ..., p_{n,j}(k))} \in [0; 1] \tag{3.1}$$

3.4 Métrique de résilience

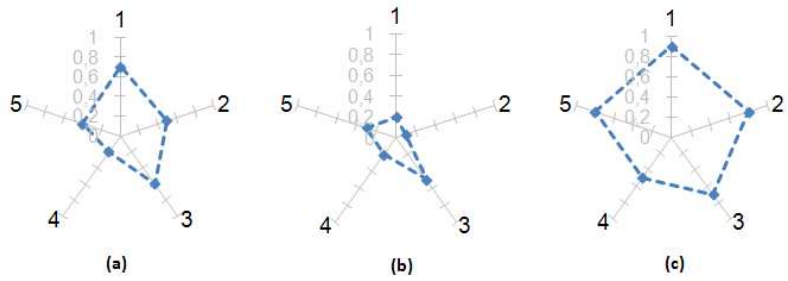


FIGURE 3.2: Différentes grandeurs de surface de résilience : (a) cas ordinaire (b) mauvaise résilience (c) bonne résilience.

où $k = \{0\%, 10\%, 20\%, 30\%, 40\%, 50\%\}$ représente le pourcentage des nœuds compromis dans le réseau.

Les deux raisons principales d'avoir choisi de normaliser les valeurs par rapport aux valeurs extrêmes (le maximum et le minimum) trouvées sur les protocoles étudiés sont les suivantes :

1. cela permet d'obtenir chaque valeur de chaque paramètre, exprimé dans sa propre unité, dans l'intervalle $\in [0; 1]$
2. cela permet de faire un "zoom" sur les différences des protocoles et de pouvoir les comparer efficacement.

Cette méthode est donc relative aux protocoles comparés et aux attaques étudiés, puisque les valeurs sont normalisées par rapport aux valeurs extrêmes des autres protocoles que l'on compare. Notez que nous ne calculons pas la résilience intrinsèque d'un protocole, mais nous calculons la résilience d'un protocole en comparaison avec d'autres. Ainsi, pour certains protocoles, et pour certains pourcentages d'attaquants, les valeurs normalisées seront nulles.

Il est possible de rendre cette méthode absolue en fixant les valeurs de référence pour chaque paramètre, puis en normalisant par ces valeurs de référence. Cependant, la difficulté reste à trouver les bonnes valeurs à prendre comme références. Si les valeurs de référence sont trop éloignées des valeurs réellement obtenues, il devient difficile de percevoir la différence entre les protocoles.

La résilience d'un protocole i est exprimée grâce à la représentation graphique de $P'_i(k)$ que l'on va appeler *surface de résilience*. Cela correspond à la surface d'un polygone (équiangle) avec m côtés ($m \geq 3$). La résilience est calculée par la formule suivante :

$$R_i(k) = (\sum_{j=1}^{m-1} p'_{i,j}(k)p'_{i,j+1}(k) + p'_{i,m}(k)p'_{i,1}(k))\frac{1}{2}\ \sin(\frac{2\pi}{m}) \tag{3.2}$$

La surface du polygone avant l'attaque représente la performance initiale d'un protocole et la diminution de la surface après les attaques représente la résilience des protocoles.

En mesurant la grandeur de cette surface en cas d'attaques, nous obtenons les valeurs quantitatives de la résilience correspondantes. Quelques exemples de la surface de résilience sont donnés sur la Figure 3.2, où nous illustrons les différentes grandeurs de surface de résilience.

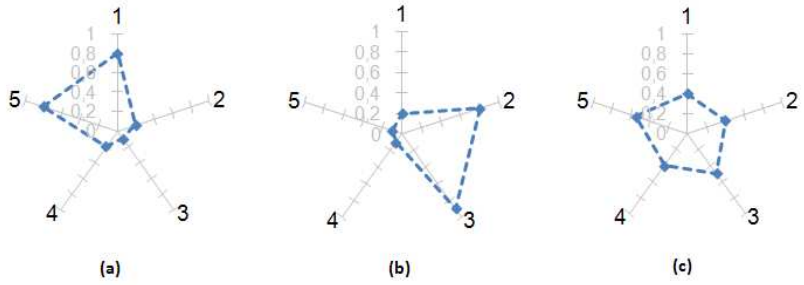


FIGURE 3.3: Différentes caractéristiques avec une surface de résilience identique : (a) dominé par 1 et 5 (b) dominé par 2 et 3 (c) équilibré.

Pour certains protocoles, il est possible que la surface de résilience soit la même. Cependant, la représentation graphique permet de saisir leurs différentes caractéristiques. Quelques exemples sont illustrés sur la Figure 3.3, où trois protocoles fictifs sont évalués selon $m = 5$ paramètres. Bien que ces protocoles ont la même grandeur de surface, ils aient des caractéristiques différentes. Le cas Figure 3.3 (c) est un protocole bien équilibré avec les valeurs équitablement réparties parmi tous les paramètres, alors que le cas Figure 3.3 (a) est nettement dominé par paramètres 5 et 1, et le cas Figure 3.3 (b) par 2 et 3. La représentation graphique de la surface nous permet donc de mettre en évidence ces aspects qualitatifs et de donner la possibilité de choisir les protocoles selon le compromis souhaité.

La surface de résilience nous permet de percevoir la dégradation des paramètres de performance en cas d'attaques. Une grande surface après les attaques indique une meilleure résilience. Toutefois, il est important de connaître également si la performance d'un protocole se dégrade progressivement avec l'augmentation d'attaquants. Une dégradation progressive et graduelle indique une meilleure résilience, puisque cela implique qu'un protocole absorbe bien l'impact des attaques.

Nous considérons le *différentiel de résilience* comme indicateur de l'intensité de la dégradation de performance avec l'augmentation des attaquants. Il est calculé par la formule suivante :

$$\frac{\Delta R_i(k)}{\Delta k} = \frac{R_i(k_{m+1}) - R_i(k_m)}{k_{m+1} - k_m} \tag{3.3}$$

Le *différentiel de résilience* permet donc de saisir si la résilience diminue de façon progressive ou brutale.

3.4.2 Paramètres de résilience pour les réseaux de capteurs

Nous illustrons maintenant comment cette méthode d'agrégation générale peut être appliquée dans le contexte spécifique du routage des réseaux de capteurs et nous justifions chacun des paramètres choisis.

Dans ce contexte, le routage des paquets doit être (1) réalisé avec succès, (2) efficace en terme de consommation énergétique, de débit et de délai et (3) réparti équitablement parmi les nœuds.

3.4 Métrique de résilience

Pour obtenir une métrique de résilience efficace, toutes ces exigences doivent être dûment prises en compte. Pour cela, les paramètres suivants sont sélectionnés afin d'évaluer la résilience des protocoles de routage. Ils sont regroupés en deux parties : primaires et secondaires. Nous avons choisi l'ordre suivant pour le placement des axes en fonction de leur importance dans notre contexte des réseaux de capteurs.

– Paramètres de résilience primaires :

  Les paramètres suivants sont primaires parce qu'ils caractérisent l'aboutissement, le but principal d'un protocole de routage des réseaux de capteurs et son coût.

  1. **ADR** (*Average Delivery Ratio*) - Taux de livraison moyen : il est défini comme la fraction du nombre des paquets reçus au puits et du nombre total de paquets envoyés par les capteurs.
  Justifications :
  Il s'agit de la métrique la plus importante pour le routage, puisque le principal rôle d'un protocole de routage est de construire des routes permettant d'acheminer des paquets de données des sources au puits. La livraison de données fiable caractérise donc le succès d'un protocole de routage. Dans les réseaux de capteurs, les nœuds sources doivent être en mesure d'atteindre le puits et délivrer leurs paquets avec succès même en présence d'attaquants.

  2. **EE** (*Energy Efficiency*) - Efficacité de la consommation énergétique : elle est définie comme l'efficacité en termes de dépense énergétique globale du réseau pour tous les paquets générés (`CONTROL` et `DATA`).
  Justifications :
  Il s'agit du coût des protocoles en consommation énergétique. Un réseau est résilient s'il a la capacité d'avoir un bon taux de livraison, pour un coût énergétique raisonnable. C'est un aspect crucial dans le contexte des réseaux de capteurs, puisque les capteurs ont une contrainte forte en énergie. Les capteurs sont équipés de batterie et l'intervention sur le terrain peut s'avérer ardu en raison de la taille de la zone de déploiement et du nombre important de capteurs. L'épuisement de l'énergie du réseau peut engendrer la mort des capteurs et la déconnexion d'une partie du réseau.

  3. **DF** (*Delivery Fairness*) - Équité de livraison du réseau : elle est définie comme l'écart type du succès de livraison de chaque nœud.
  Justifications :
  Elle caractérise la manière dont le succès de livraison est reparti parmi les nœuds et donc la distribution du succès de livraison des nœuds. Pour un même taux de livraison moyen, les protocoles peuvent différer en termes d'équité de livraison : certains nœuds peuvent être complètement déconnectés du puits alors que d'autres auront un taux de livraison très important. La distribution du succès de livraison est donc très importante, car un réseau est plus résilient si une grande partie des nœuds arrivent à atteindre le puits même en cas d'attaques. D'un point de vue applicatif, il est crucial que chaque capteur puisse envoyer ses données au puits en cas d'attaques, parce que cela permet d'avoir une vision de l'ensemble du réseau (même si elle est dégradée) et d'obtenir une

large couverture géographique.

– Paramètres de résilience secondaires :

Les paramètres suivants sont secondaires parce qu'ils caractérisent l'efficacité d'un protocole, mais ils ne caractérisent pas l'aboutissement du protocole de routage. Ils sont donc importants, mais pas cruciaux.

4. **AT** (*Average Throughput*) - Débit moyen : il est défini comme la quantité maximum de flux de données reçues au puits durant une unité de temps.
Justifications :
Dans le contexte spécifique des réseaux de capteurs, nous devons également considérer le phénomène de convergence du trafic vers le puits en raison du modèle de trafic *convergecast* (communication *many-to-one*), où tous les flux de données des capteurs sont collectés au puits. Cela peut engendrer d'éventuel goulots d'étranglement autour du puits. Il est donc important de considérer le débit au sens de capacité du réseau (*network capacity*) [112], lorsqu'on évalue le débit sur le réseau soumis à un trafic en régime saturé.

5. **DE** (*Delay Efficiency*) - Efficacité du délai moyen : elle est définie comme l'efficacité du délai de bout en bout d'un paquet généré par une source pour arriver au puits.
Justifications :
Cela comprend le temps passé dans les files d'attente, en retransmissions et en propagation. Dans nos simulations, elle est fonction de la longueur moyenne du chemin en nombre de sauts. Nous faisons abstraction des couches inférieures pour se concentrer uniquement sur l'impact des attaques à la couche routage. Elle caractérise l'efficacité du réseau en termes de vitesse de livraison. Certaines applications des réseaux de capteurs telles que les alarmes incendie, les applications de détection de fuites, etc. ont des contraintes de temps réel. Il est donc important que les protocoles puissent fournir un délai de livraison court.

Comme nous avons discuté précédemment, le choix de l'ordre des axes a une importance dans le calcul de la résilience. Nous exploitons la propriété de calcul de la surface du polygone pour discriminer les protocoles en fonction de certains paramètres.

La principale raison d'avoir choisi cet ordre précis : (ADR), (EE), (DF), (AT) et (DE) est que l'efficacité énergétique permet de pondérer simultanément le taux de livraison moyen et l'équité de livraison dans le groupe des paramètres les plus importants. Cela signifie qu'un protocole efficace en termes de succès de livraison et de l'équité de livraison ne doit pas être coûteux en consommation énergétique pour une bonne résilience. La consommation de l'énergie étant particulièrement pénalisante dans un réseau de capteurs, pouvant engendrer l'épuisement (et la mort) des capteurs et la déconnexion d'une partie du réseau, cette discrimination énergétique est souhaitable.

Dans l'exemple suivant sur la Figure 3.4, nous illustrons deux ordres d'axes différents. Supposons que nous mettons les paramètres dans l'ordre : (ADR), (EE), (DF), (AT) et (DE) côte à côte (Figure 3.4 (a)), puis nous calculons la surface du polygone qui nous donne la valeur $0,92247$. Cela signifie que le paramètre (EE) pondère simultanément les paramètres

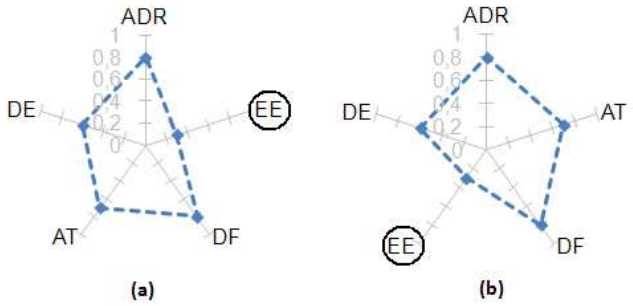


FIGURE 3.4: Différents ordres des axes avec un ensemble de valeurs identiques : (a) importance primaire d'énergie (b) importance secondaire d'énergie.

(ADR) et (DF). Pour le calcul de la surface, un (EE) très bas avec (ADR) et (DF) très hauts est pénalisant au sens de la résilience. Ensuite, supposons que nous mettons les paramètres dans l'ordre suivant : (ADR), (AT), (DF), (EE) et (DE) (Figure 3.4 (b)), la surface calculée nous donne la valeur $0,96051$. Cela revient à minimiser l'importance de l'énergie.

Cet ordre spécifique des axes doit également mettre en évidence la différence entre les protocoles et les conséquences des attaques sur la performance des protocoles.

3.5 Application de la métrique aux protocoles de routage classiques

Les hypothèses considérées ici sont identiques à celles de notre étude préliminaire de la section 2.2. Les notations sont également les mêmes. Pour comparer la résilience des protocoles de routage confrontés à des attaques, nous avons évaluons les cinq protocoles suivants : Dynamic Source Routing (DSR)[26], Gradient Based Routing (GBR) [27], Greedy Forwarding (GF)[28], Random Walk Routing (RWR)[29] et Biased Random Walk Routing (BRWR)[2].

3.5.1 Modèle d'adversaires

Dans cette étude, nous considérons l'attaque *Selective forwarding*, où les nœuds malveillants ne retransmettent pas les paquets de données. Comme nous avons discuté dans la section 2.2, cette attaque a la particularité d'être réalisable avec tous les protocoles de routage étudiés puisqu'elle cible les paquets de données. Elle est particulièrement pertinente pour étudier l'acheminement des paquets de données et donc le succès des protocoles de routage multi-sauts.

Nous faisons l'hypothèse que l'adversaire ne détient pas d'information sur l'emplacement du puits et donc les nœuds compromis (un pourcentage k des nœuds) sont déployés aléatoirement et uniformément sur un plan carré de taille $N \times N$.

Les nœuds compromis jettent tous les paquets provenant de leurs voisins, mais continuent de générer leurs propres paquets de données. Pour nos simulations, k varie entre 10% et 50% de la population des nœuds du réseau.

Nous avons choisi de montrer le comportement des protocoles face à des attaques modérées puis plus importantes, de manière progressive.

3.5.2 Évaluation des performances

Le but de cette étude par simulations est de montrer comment notre métrique de résilience peut s'appliquer en pratique aux protocoles classiques de différentes catégories.

Toutes nos simulations sont effectuées sur le simulateur WSNet [80] comme dans le cas de notre étude préliminaire et nous considérons les mêmes paramètres de simulations. Le tableau 2.1 du chapitre 2 résume les paramètres de simulations.

Analyses des performances

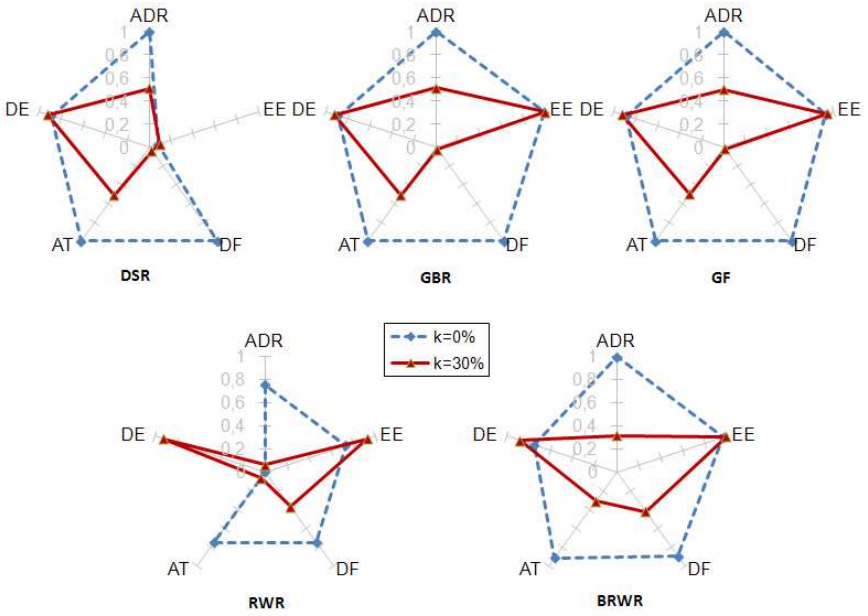


FIGURE 3.5: Surface de la résilience des protocoles classiques sans attaque et avec attaques ($k = 30\%$).

L'évaluation de la surface de résilience des protocoles de routage classiques (GBR, DSR, GF, RWR et BRWR) sans et avec attaques *Selective forwarding* ($k = 30\%$) est illustrée sur la Figure 3.5.

Comme présenté précédemment, nous avons considéré $m = 5$ paramètres de performances dans l'ordre suivant :

1. ADR - Taux de livraison moyen
2. EE - Efficacité de la consommation énergétique
3. DF - Équité de livraison du réseau
4. AT - Débit moyen

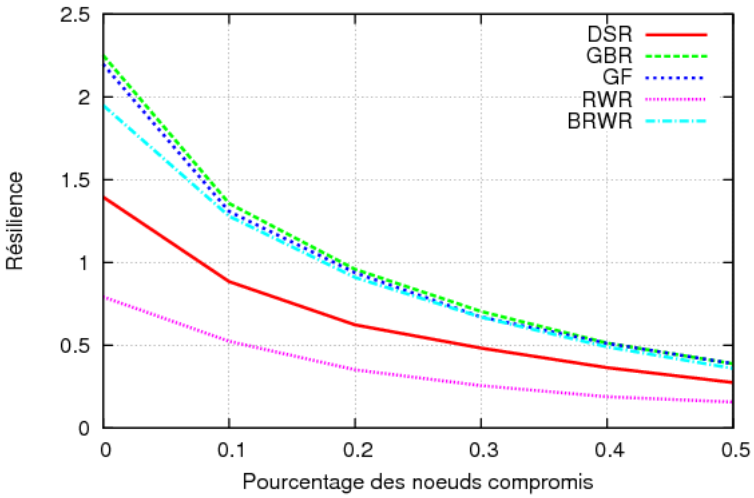


FIGURE 3.6: Évaluation quantitative de la résilience des protocoles classiques en cas d'attaques.

5. DE - Efficacité du délai moyen

Nous exploitons la propriété du calcul de la surface du polygone pour discriminer les protocoles en fonction de leur consommation énergétique.

La représentation graphique de la surface de résilience permet d'avoir une synthèse de la performance sur un plan, tout en donnant de l'information sur chaque axe avant et après les attaques. Sans attaque, cette surface nous montre le profil et la tendance des protocoles. En cas d'attaques, cette surface nous montre quels paramètres sont les plus impactés. Cette représentation graphique nous permet de mieux comprendre le fonctionnement des protocoles parce cela donne une information qualitative et donc cela permet de déterminer le profil des protocoles (plutôt économe en énergie, plutôt bon en livraison, etc.). Elle nous permet également de calculer la surface du polygone et donc quantifier la résilience.

L'évaluation quantitative de la résilience des protocoles classiques en cas d'attaques est donnée sur la Figure 3.6. La quantification de la surface nous donne la possibilité de classer et de donner une relation d'ordre entre les protocoles en termes de résilience. La classification des protocoles est présentée dans l'ordre décroissant de résilience :

1. GBR
2. GF
3. BRWR
4. DSR
5. RWR

Les protocoles GBR, GF et BRWR sont les plus performants parce que ces protocoles sont les mieux équilibrés pour tous les paramètres. Sans attaque, leur principale différence apparaît sur l'axe DE, l'efficacité de délai, et cela dû à la longueur de chemins qui sont les plus courts.

En cas d'attaques, les surfaces sont réduites pour tous les protocoles. Cette diminution est due à une diminution de tous les axes sauf l'axe EE, l'efficacité énergétique et l'axe

DE, l'efficacité en délai (Figure 3.5). Ce phénomène est contre-intuitif, parce que EE et DE s'améliorent avec les attaques. Cette amélioration de l'efficacité énergétique (EE) s'explique par la diminution du trafic du réseau à cause des paquets perdus. Les attaques ont également pour conséquence de réduire l'accès au puits des nœuds les plus éloignés. Les paquets reçus au puits proviennent ainsi principalement des nœuds sources proches et la longueur moyenne des chemins constatée diminue, entraînant une amélioration de l'efficacité en délai (DE). Nous constatons une diminution brutale de l'axe DF pour les protocoles GBR, GF et DSR. Les protocoles sont basés sur le critère du plus court chemin, leur comportement est donc déterministe. Le succès de livraison n'est pas équitablement distribué parmi les nœuds, puisque les nœuds ayant au moins un nœud compromis sur leur route sont complètement déconnectés du puits.

Les protocoles DSR et RWR sont les moins performants en raison de deux aspects importants : (i) la longueur des chemins et (ii) le mécanisme d'établissement des routes. DSR est pénalisé sur l'aspect (ii) tandis que RWR est pénalisé sur l'aspect (i). Sans attaque, la surface du polygone nous confirme le déséquilibre de DSR sur l'axe EE (Figure 3.5). DSR est basé sur le routage par la source, ce qui signifie que chaque nœud source construit sa route vers le puits en inondant le réseau par un paquet `RREQ`. Ce mécanisme de construction des routes par les sources est très coûteux en terme de consommation énergétique. Il n'est donc pas adapté dans le contexte de réseaux de capteurs, où nous considérions un nombre important des sources et un seul (ou peu de) puits. RWR, quant à lui, est très pénalisé par la longueur des chemins en raison de la marche aléatoire des paquets. Les paquets peuvent voyager longtemps, faire des boucles et revenir sur les nœuds déjà traversés. Sans attaque, la surface du polygone nous confirme le déséquilibre de RWR sur l'axe DE (Figure 3.5). L'efficacité en délai est directement liée à la longueur moyenne de chemins. Plus le chemin est long, plus le paquet voyage longtemps avant d'arriver au puits.

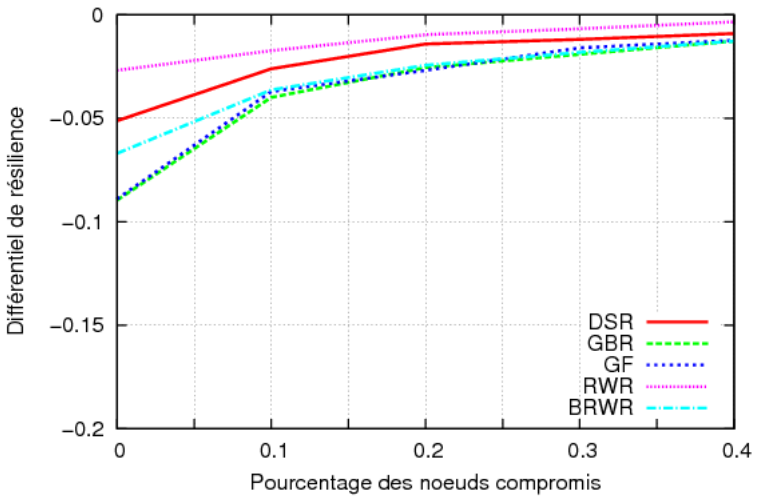


FIGURE 3.7: Différentiel de la résilience des protocoles classiques en cas d'attaques.

Cependant, si nous regardons le différentiel de résilience sur la Figure 3.7, nous constatons que la dégradation de résilience de RWR est la moins brutale comparée aux autres protocoles, même s'il reste le moins performant sans attaque. Cette propriété de RWR a at-

tiré notre attention, parce que la diminution de résilience progressive (*gracefull degradation*) avec l'augmentation d'attaques implique que le protocole absorbe bien le choc causé par des attaques.

En ce sens, la marche aléatoire a des propriétés extrêmement intéressantes pour la résilience sous condition d'améliorer sa performance initiale sans attaque. Ce phénomène s'explique par l'équité de livraison (DF) bien meilleure des protocoles aléatoires en cas d'attaques. Leur distribution du taux de livraison parmi les sources est bien meilleure, même avec un taux de livraison moyen plus bas. Il est important de constater qu'un comportement aléatoire dans le choix de la route va entraîner un gain de robustesse. Ce gain est réalisé grâce à une plus grande diversité des routes empruntées, permettant d'éviter des nœuds corrompus. Le fait de biaiser la marche aléatoire (BRWR) nous donne la possibilité d'améliorer la performance de la marche aléatoire standard (RWR) en diminuant la longueur de chemins excessive. Pour ces raisons, nous nous attacherons à exploiter ces mécanismes dans la suite de cette thèse.

3.6 Conclusion

Dans ce chapitre, nous avons présenté le concept de la résilience comprenant notre définition de la résilience dans le contexte de la sécurité de routage des réseaux de capteurs et la métrique permettant de mesurer la résilience des protocoles de routage. Notre métrique prend en compte plusieurs paramètres de performance essentiels (taux de livraison moyen, équité de livraison, efficacité en consommation d'énergie, débit moyen et efficacité en délai) pour exprimer la résilience des protocoles de routage. Nous avons ensuite appliqué cette métrique de résilience à des protocoles de routage dédiées aux réseaux de capteurs.

Grâce à cette métrique, nous avons évalué la résilience de cinq protocoles de routage classiques de différentes catégories. La représentation graphique de la métrique nous donne une information qualitative. Elle permet de discerner instantanément la différence des protocoles, d'avoir une vue synthétique sur tous les paramètres, et de percevoir les paramètres les plus impactés par des attaques. Le calcul de la surface nous donne une information quantitative. Elle permet d'établir une relation d'ordre entre les protocoles.

Cette étude nous permet de souligner les trois éléments suivants :

- Le choix de la stratégie de construction des routes reste un aspect important pour la performance des protocoles et donc également pour la résilience. Certains protocoles tels que DSR ou RWR sont pénalisés par la consommation d'énergie (EE) et par la longueur de chemins respectivement (DE).
- Les comportements aléatoires des protocoles RWR et BRWR ont permis une meilleure équité de livraison (DF) que les protocoles déterministes grâce aux diversifications des routes. La marche aléatoire permet une meilleure connexité entre les capteurs et le puits, même si le taux de livraison moyen est plus bas que les protocoles déterministes.
- La marche aléatoire biaisée est très sensiblement meilleure que RWR grâce à une longueur de chemins plus courte (DE). BRWR bénéficie également du comportement aléatoire (DF).

Le comportement aléatoire est donc bénéfique à la résilience sous condition de limiter la longueur des chemins comme dans le cas de BRWR. Grâce au choix aléatoire du prochain

saut, chaque paquet a la possibilité d'exploiter des chemins différents, apportant ainsi une grande diversité des routes pour chaque capteur pour atteindre le puits. Ceci est d'autant plus souhaitable dans les réseaux de capteurs pour exploiter leur densité naturelle. La variation des routes pour chaque paquet permet également une distribution plus équitable de la consommation énergétique du réseau. Les nœuds les plus sollicités par un routage basé sur les plus courts chemins sont soulagés en cas de routage aléatoire. Un autre avantage pour la sécurité vient du fait que le comportement aléatoire augmente l'incertitude vis-à-vis des adversaires. Il devient difficile pour un attaquant de prédire quel chemin un paquet va emprunter même en connaissant la stratégie de routage. C'est une façon de dissuader les adversaires de tenter d'attirer le trafic en falsifiant les paquets de contrôle et de diminuer l'impact de leurs attaques.

Pour faire suite à cette étude, nous présentons dans le prochain chapitre, notre proposition pour des mécanismes résilients de routage.

Proposition des mécanismes résilients pour le routage

4

Sommaire

4.1 Introduction 59
4.2 Mécanismes résilients pour le routage 59
4.2.1 Introduction de comportements aléatoires 60
4.2.2 Limitation de la longueur des routes 60
4.2.3 Réplication des paquets 61
4.3 Introduction des mécanismes résilients aux protocoles de routage classiques 61
4.3.1 Protocoles aléatoires sans réplication 61
4.3.2 Protocoles aléatoires avec réplications 62
4.3.3 Évaluation des performances 63
4.3.4 Taxonomie de la résilience des protocoles de routage 71
4.4 Conclusion 72

4.1 Introduction

Dans les réseaux de capteurs sans fils, économiser l'énergie est souvent considérée comme une priorité. Nous avons montré dans l'étude précédente que la plupart des protocoles de routage que nous avons étudié se basent sur le critère du plus court chemin. Le but de cette stratégie est d'atteindre le puits en limitant le nombre de nœuds relais et ainsi permettre un délai moindre et une plus faible consommation énergétique. Les protocoles de routage réactifs tels que DSR [26], le routage géographique GF [28] et le routage par gradient GBR [27], emploient tous ce principe du plus court chemin.

Le fait d'optimiser les routes sur un critère fixe tel que la longueur des routes implique un comportement déterministe, puisque les paquets sont acheminés le long des meilleures routes selon le critère choisi. Cette observation est valable en particulier dans le contexte des réseaux de capteurs, où la mobilité n'est pas considérée.

Malheureusement, cette stratégie, parfaitement valable dans un contexte sans attaques, n'est plus bénéfique en présence d'attaquants internes. Les limitations d'un routage déterministe en cas d'attaques sont les suivants :

- déconnexion des nœuds : une source utilise toujours le meilleur chemin pour transmettre tous ses paquets. Il suffit d'avoir au moins un nœud compromis sur la meilleure route pour que la source soit complètement déconnectée du puits. Cela engendre une distribution non équitable du succès de livraison parmi les sources.
- stratégie prévisible : un attaquant peut exploiter le critère de sélection des routes afin de manipuler les nœuds légitimes et attirer plus de trafic vers lui. Un seul nœud compromis peut faire d'importants dégâts dans son voisinage s'il est choisi comme prochain saut par tous ses voisins.
- surcharge non équitable : les mêmes nœds, situés sur les meilleures routes sont sollicités pour transmettre les paquets des autres sources. Ils risquent d'épuiser leur énergie plus rapidement que les autres capteurs. Cela empêche également de bénéficier des liens redondants des réseaux de capteurs, puisque beaucoup de routes existantes ne sont pas utilisées.

Le chapitre 3 tend à montrer que le comportement aléatoire est bénéfique à la résilience sous condition de limiter la longueur des routes. Nous allons donc approfondir notre étude dans cette direction en proposant des mécanismes résilients que nous introduisons dans les protocoles classiques.

Notre but est de démontrer l'utilité des mécanismes résilients dans le routage en cas d'attaques par compromission des nœuds, mais également d'illustrer la validité et la pertinence de la métrique de résilience que nous avons définie dans le chapitre 3.

4.2 Mécanismes résilients pour le routage

Comme discuté précédemment, la plupart des protocoles de routage sont basés sur le critère du plus court chemin ce qui a pour conséquences néfastes de limiter la diversité des routes entraînant une vulnérabilité aux attaquants internes et une faible équité. En raison de ces observations, nous présentons les mécanismes de routage qui améliorent la résilience

des protocoles classiques.

Notre proposition consiste en trois éléments : (i) introduire du comportement aléatoire (ii) limiter la longueur des routes (iii) ajouter de la réplication des paquets.

4.2.1 Introduction de comportements aléatoires

Le routage basé sur le critère du plus court chemin oblige au trafic de circuler sur un sous-ensemble réduit aux nœuds se trouvant sur les meilleures routes. Le succès de livraison n'est pas équitablement réparti entre les nœuds du réseau : certains nœuds ont un très bon taux de livraison et d'autres sont complètement déconnectés du puits. Il s'agit d'une limitation du protocole puisqu'il ne permet pas d'exploiter la redondance structurelle de la topologie physique du réseau pour bénéficier des routes alternatives existantes. L'introduction de comportement aléatoire permet donc de rendre plus dynamiques les protocoles classiques dans les mêmes conditions.

L'introduction d'un comportement aléatoire dans un protocole peut se faire de différentes manières en fonction de l'information que les nœuds détiennent (information des routes construites préalablement, informations géographiques ou gradient, etc.). Par exemple, dans DSR [26], les nœuds construisent les routes préalablement et ils ont donc la liste des prochains sauts. Dans GBR [27], les nœuds obtiennent de l' information de distance en nombre de sauts. Enfin, dans GF [28], les nœuds disposent d'information sur les coordonnées géographiques et grâce à la découverte de voisinage, ils disposent donc la distance du puits de leurs voisins.

4.2.2 Limitation de la longueur des routes

La marche aléatoire standard est trop pénalisée par sa longueur excessive des routes comme nous l'avons montré dans les chapitres 2 et 3. Supposons que les nœuds compromis (ne retransmettant pas les paquets) sont distribués uniformément dans le réseau. Si l est la longueur des routes en nombre de sauts d'une source du puits, si P_c est la probabilité qu'un nœud soit compromis et si P_n est la probabilité qu'un paquet soit livré, nous avons $P_n = (1 - P_c)^{l-1}$. La probabilité de trouver une route sans attaquants diminue exponentiellement avec la longueur des routes. Il faut donc limiter la longueur des routes pour une meilleure résilience. Nous proposons de biaiser la marche aléatoire pour la diriger vers le puits grâce aux informations d'états disponibles à chaque nœud.

D'une manière générale, l'introduction de comportements aléatoires et la limitation de la longueur des routes nécessitent deux choses. Tout d'abord, déterminer, pour chaque nœud, l'ensemble des candidats pour le prochain saut. Ensuite, il faut spécifier une probabilité de sélection sur cet ensemble. Par exemple, il est possible d'attribuer une probabilité plus élevée aux nœuds plus proches du puits. C'est une façon de biaiser la marche aléatoire.

Illustrons ce biais avec le routage par gradient GBR. Dans ce protocole, chaque nœud détient, pour chacun de ses voisins, sa distance du puits en nombre sauts. Un nœud peut déterminer un sous-ensemble de voisins plus proches du puits que lui-même grâce à cette information du gradient. Il peut donc choisir le prochain saut aléatoirement dans ce sous-ensemble des voisins.

Il est possible de faire varier les paramètres pour régler la taille de l'ensemble de candidats pour le prochain saut et la probabilité de sélection afin de contrôler le biais que l'on introduit.

4.2.3 Réplication des paquets

Afin d'exploiter efficacement la diversité des routes crée par les protocoles aléatoires, nous répliquons x fois chaque paquet de données pour augmenter son succès de livraison. Chaque copie répliquée est une itération indépendante de la marche aléatoire, et elle suivra donc son propre chemin pour accéder au puits.

C'est une façon de créer des chemins multiples sans toutefois garantir des routes disjointes. Dans un protocole avec construction de routes préalable (comme DSR), nous avons la possibilité de construire plusieurs routes disjointes pour chaque source et de choisir aléatoirement une route pour chaque envoi de message. Cependant, la construction et la maintenance des routes disjointes sont très coûteuses en termes de consommation d'énergie. De plus, leur construction se fait préalablement et elle rend donc le protocole non flexible aux changements et aux adaptations.

Les protocoles tels que GBR et GF ne construisent pas de routes préalables. Pour obtenir des routes disjointes, il faut prévoir des échanges de paquets de contrôle supplémentaires entre le puits et les nœuds, augmentant ainsi considérablement le coût des protocoles. Avec notre méthode de construction des routes aléatoires, nous donnons la possibilité de diversifier les routes, mais également la possibilité d'ajuster les valeurs du biais que l'on introduit, sans paquets de contrôle supplémentaires.

4.3 Introduction des mécanismes résilients aux protocoles de routage classiques

Dans cette étude, nous introduisons les mécanismes résilients aux protocoles classiques étudiés dans les chapitres précédents (DSR [26], GBR [27], GF [28], RWR [29] et BRWR [2]). Nous les évaluons selon la métrique de résilience présentée au chapitre 3.

Notre objectif est double :
- étudier le gain possible en termes de résilience de mécanismes tels que l'introduction d'aléa et de la réplication
- montrer la pertinence de notre métrique définie dans le chapitre 3 dans l'évaluation de la résilience des protocoles de routage.

4.3.1 Protocoles aléatoires sans réplication

Les variantes aléatoires des protocoles sans réplication seront nommées en ajoutant le préfixe **RS** (*Random Single path*).

RS-DSR : pour un protocole nécessitant un processus de découverte de routes préalable tel que DSR, nous proposons sa variante aléatoire RS-DSR, qui permet de découvrir plusieurs routes (les plus courtes) entre un nœud et le puits au lieu d'une seule pour la version originale.

Dans RS-DSR, la notion de plus court chemin est ainsi conservée. Quand un nœud source inonde le réseau par un paquet `RREQ`, le puits attend de recevoir plusieurs `RREQ` par des chemins différents. Le puits envoie par le chemin inverse un paquet `RREP` pour chaque `RREQ` reçu. Chaque nœud relais sur la route enregistre les informations de routage sur leur table de routage. Ensuite, pour chaque paquet de données envoyé, un nœud sélectionne aléatoirement un relais parmi les candidats enregistrés pour transmettre le paquet.

RS-GF et **RS-GBR** : pour les protocoles sans construction des routes préalables tels que GF et GBR, nous proposons leurs variantes aléatoires RS-GF et RS-GBR, où chaque nœud détermine un sous-ensemble de voisins plus proches du puits que lui-même. Ensuite, un nœud choisit aléatoirement un voisin comme prochain saut dans ce sous-ensemble. Pour RS-GF, les nœuds disposent d'informations géographiques, tandis que RS-GBR possède des informations de gradient. Dans RS-GF, un nœud choisit le prochain saut aléatoirement dans le sous-ensemble des nœuds géographiquement plus proches du puits. Dans RS-GBR, un nœud choisit le prochain saut aléatoirement dans l'ensemble des voisins qui ont des gradients plus petits que le sien.

Dans ces variantes aléatoires des protocoles, les nœuds sources ne considèrent pas de réplication des paquets. Un paquet est donc acheminé par une seule route, construite de façon aléatoire. C'est équivalent à l'itération d'une marche aléatoire biaisée par l'information qu'un nœud détient (information géographique, gradient, information des routes, etc.).

RS-BRWR et **RS-RWR** : sont identiques à leurs variantes classiques BRWR et RWR, puisqu'ils sont déjà des protocoles aléatoires.

4.3.2 Protocoles aléatoires avec réplications

Les variantes aléatoires des protocoles avec réplication des paquets par les sources seront nommées en ajoutant un préfixe **RM**(*Random Multiple paths*).

RM-DSR, **RM-GF**, **RM-GBR** et **RM-BRWR** : nous proposons d'adapter la réplication des paquets en fonction de la distance en nombre de sauts. Un nœud lointain répliquera plus et inversement un nœud plus proche du puits répliquera moins. Parce que le paquet provenant d'un nœud loin du puits a plus de chance de rencontrer un attaquant en route, puisqu'il parcourt un chemin plus long. Ce mécanisme peut être introduit aux protocoles aléatoires, où chaque nœud dispose d'une information de la direction du puits (distance géographique, gradient, routes découvertes à l'avance, etc.). Chaque nœud a la possibilité donc de connaitre sa distance du puits (en nombre de sauts). Nous pouvons donc en bénéficier pour adapter la réplication. Par exemple, un nœud une distance d du puits réplique $d-1$ fois chaque paquet. Cependant, même si le taux de réplication dépend de la distance du puits, nous avons limité le nombre maximum de réplications par nœud à trois. Cette limitation est nécessaire, car les nœuds lointains peuvent engendrer un trafic important, surtout dans un réseau étendu avec un diamètre important.

RM-RWR : il n'est pas possible d'introduire ce mécanisme de réplications adaptatives à la marche aléatoire standard, parce que les sources ne disposent pas d'information de distance. Cependant, nous connaissons la distribution de distance (en nombre de sauts) des nœuds, grâce aux statistiques des expériences précédentes. Les nœuds sont répartis aléatoirement sur une zone $N \times N$ selon une distribution uniforme et avec un seul puits au centre. Nous

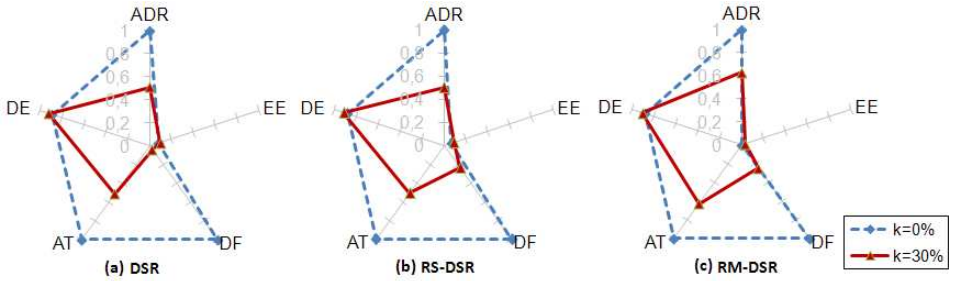


FIGURE 4.1: Surface de la résilience de DSR sans attaque et avec attaques ($k = 30\%$) : (a) Version classique DSR (b) Variante aléatoire RS-DSR (c) Variante aléatoire avec réplication RM-DSR.

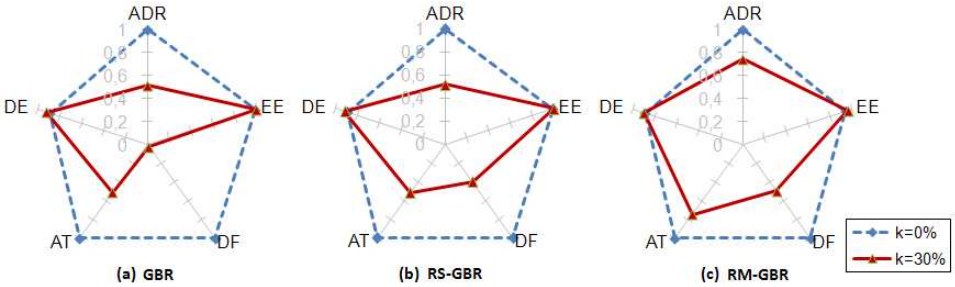


FIGURE 4.2: Surface de la résilience de GBR sans attaque et avec attaques ($k = 30\%$) : (a) Version classique GBR (b) Variante aléatoire RS-GBR (c) Variante aléatoire avec réplication RM-GBR.

connaissons la répartition des nœuds en fonction de leur distance du puits.

Nous prenons l'hypothèse de sources connaissant cette distribution $\sum_i^{d_{\max}} \frac{n_i}{n} = 1$, où n_i est le nombre de nœuds de niveau i (distance en nombre de sauts), n est le nombre total des nœuds et $d_{\max}$ est le nombre total des niveaux.

Quelle que soit sa distance, chaque nœud peut choisir aléatoirement un taux de réplication r selon cette distribution. Nous avons adopté ce mécanisme afin de comparer RM-RWR équitablement aux autres protocoles.

4.3.3 Évaluation des performances

Les simulations sont effectuées sur le simulateur WSNet [80] avec les mêmes hypothèses, le même modèle d'adversaire et les mêmes paramètres que dans les chapitres précédents.

Résultats des protocoles aléatoires sans réplication

Pour les variantes classiques des protocoles (DSR, GBR, GF, RWR et BRWR), les Figures 4.1 - 4.5 (a) illustrent la surface de résilience sans et avec attaques selon $m = 5$ paramètres de performances dans l'ordre suivant :

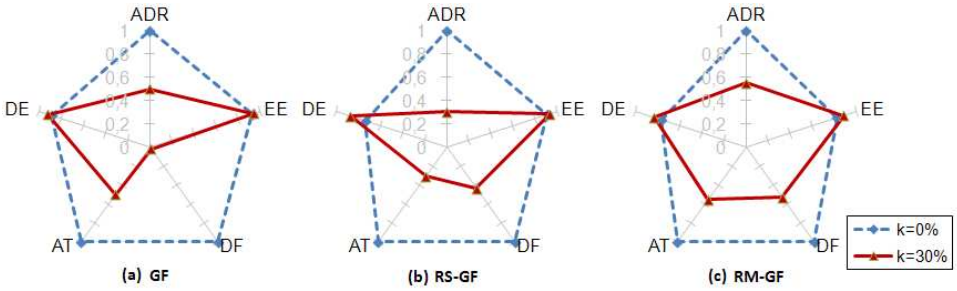


FIGURE 4.3: Surface de la résilience de GF sans attaque et avec attaques ($k = 30\%$) :(a) Version classique GF (b) Variante aléatoire RS-GF (c) Variante aléatoire avec réplication RM-GF.

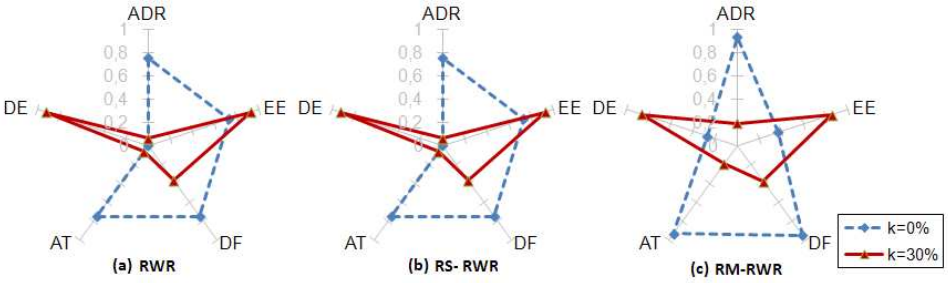


FIGURE 4.4: Surface de la résilience de RWR sans attaque et avec attaques ($k = 30\%$) : (a) Version classique RWR (b) Variante aléatoire avec réplication RM-RWR.

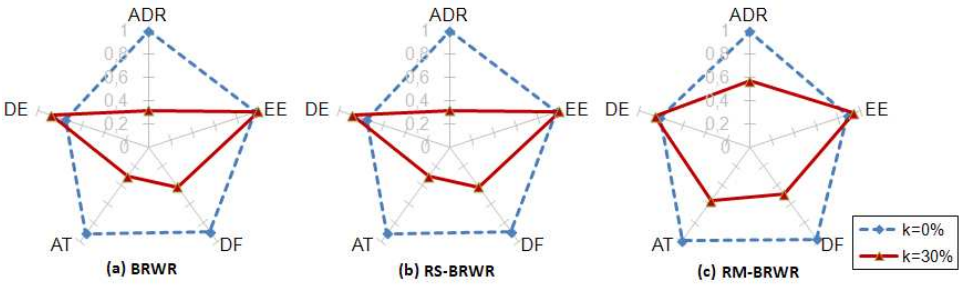


FIGURE 4.5: Surface de la résilience de BRWR sans attaque et avec attaques ($k = 30\%$) : (a) Version classique BRWR (b) Variante aléatoire avec réplication RM-BRWR.

1. ADR - Taux de livraison moyen
2. EE - Efficacité de la consommation énergétique
3. DF - Équité de livraison du réseau
4. AT - Débit moyen
5. DE - Efficacité du délai moyen

En cas d'attaques, comme constaté dans le chapitre 3, le taux de livraison moyen (ADR), le débit moyen (AT) et l'équité de livraison (DF) diminuent en raison de la perte des paquets causée par les attaques, mais aussi les deux phénomènes contre-intuitifs apparaissent concernant l'efficacité en délai (DE) et l'efficacité en énergie (EE) qui augmentent avec les attaques.

Comme expliqué précédemment dans le chapitre 3, l'attaque du réseau entraîne une diminution du trafic puisque moins de paquets circulent dans le réseau. Pour cette raison, le réseau dépense moins d'énergie, d'où une amélioration de l'efficacité énergétique (EE). Un autre phénomène rentre en jeu concernant l'amélioration du délai. En cas d'attaques, les paquets en provenance de nœuds éloignés du puits ont une chance plus grande de se perdre. Ainsi, les paquets effectivement reçus viennent statistiquement des nœuds les plus proches du puits. Ce phénomène est mesuré sur l'axe d'efficacité du délai (DE) car dans notre étude le délai est directement proportionnel à la longueur moyenne des routes.

Pour les variantes aléatoires des protocoles sans réplication (RS-GBR, RS-DSR, RS-GF, RWR et BRWR), la surface de résilience est représentée sur la Figure 4.1 - 4.5 (b). Sans attaque, les variantes aléatoires n'apportent pas de coût supplémentaire, sauf pour RS-GF, où nous constatons une diminution sur l'axe DE, l'efficacité en délai (Figure 4.3 (b)). Cela s'explique par l'allongement des chemins que le comportement aléatoire introduit. Dans RS-GF, un nœud choisit le prochain saut aléatoirement dans l'ensemble des voisins géographiquement plus proches du puits (au lieu de choisir le plus proche). Ce mécanisme autorise donc la construction de chemins plus longs que la version originale.

En cas d'attaques, en revanche, nous observons une amélioration sur l'axe DF, l'équité de livraison pour tous les protocoles. Le taux de livraison est réparti équitablement parmi les sources parce qu'un plus grand nombre de sources arrive à transmettre leurs données au puits grâce à la diversification des routes. Chaque paquet envoyé par une source peut prendre un chemin potentiellement différent grâce au choix aléatoire du prochain saut. La connectivité des routes entre les sources et le puits est ainsi améliorée, puisqu'un plus grand nombre de sources reste connecté au puits malgré l'augmentation d'intensité des attaques, même avec un taux de livraison moyen dégradé. Ces phénomènes sont mesurés sur l'axe d'équité de livraison (DF) qui représente l'écart type du taux de livraison des nœuds. C'est la distribution de succès de livraison parmi les sources.

Au vu de ces résultats, on observe que l'introduction d'un comportement aléatoire améliore significativement la résilience sans apporter d'inconvénient majeur pour tous des protocoles exceptés RS-GF où l'on constate un allongement des routes. L'allongement des chemins entraîne une diminution sur les axes ADR et EE : les paquets ont une plus grande probabilité de croiser un nœud attaquant (le taux de livraison moyen diminue) et les paquets étant retransmis un plus grand nombre de fois, la consommation d'énergie s'en trouve augmentée.

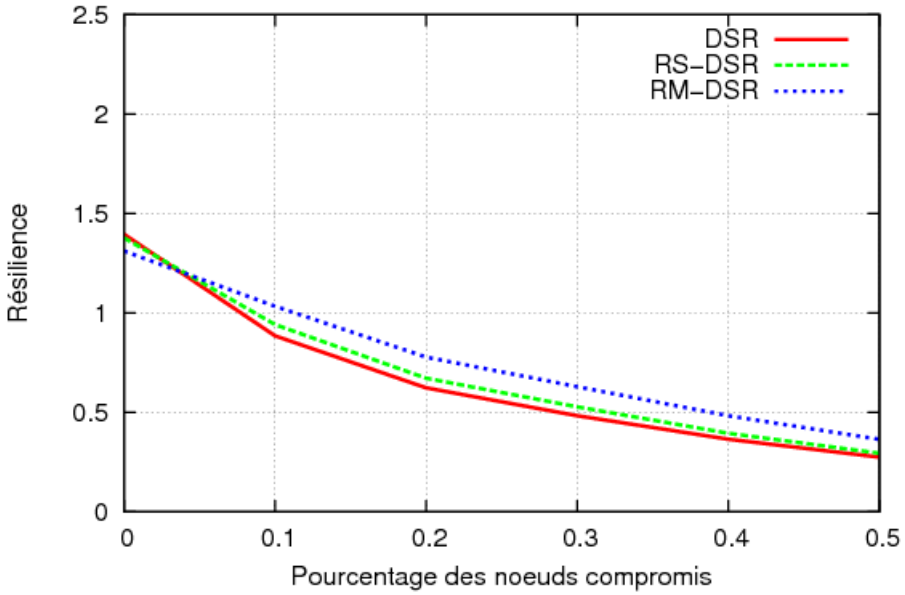


FIGURE 4.6: Évaluation quantitative de la résilience des variantes de DSR.

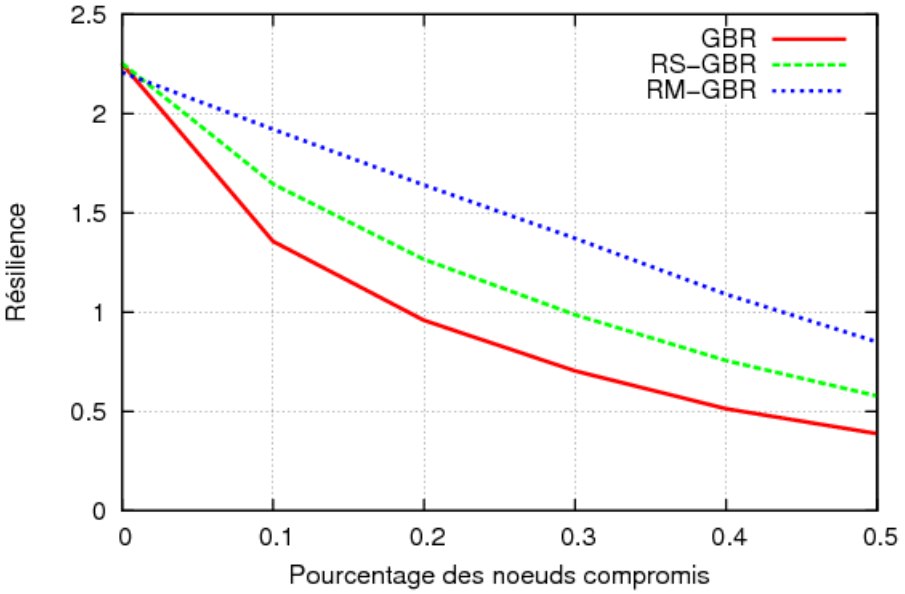


FIGURE 4.7: Évaluation quantitative de la résilience des variantes de GBR.

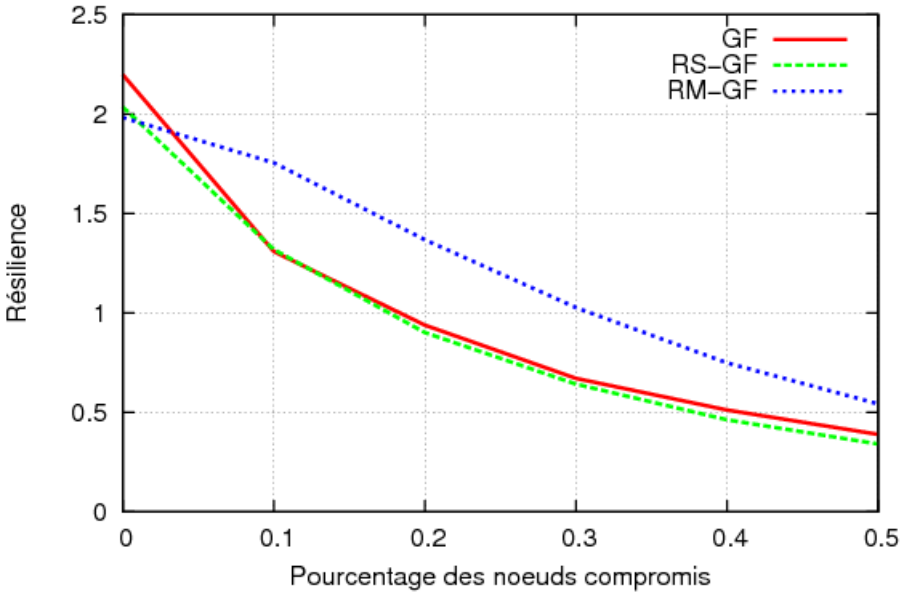


FIGURE 4.8: Évaluation quantitative de la résilience des variantes de GF.

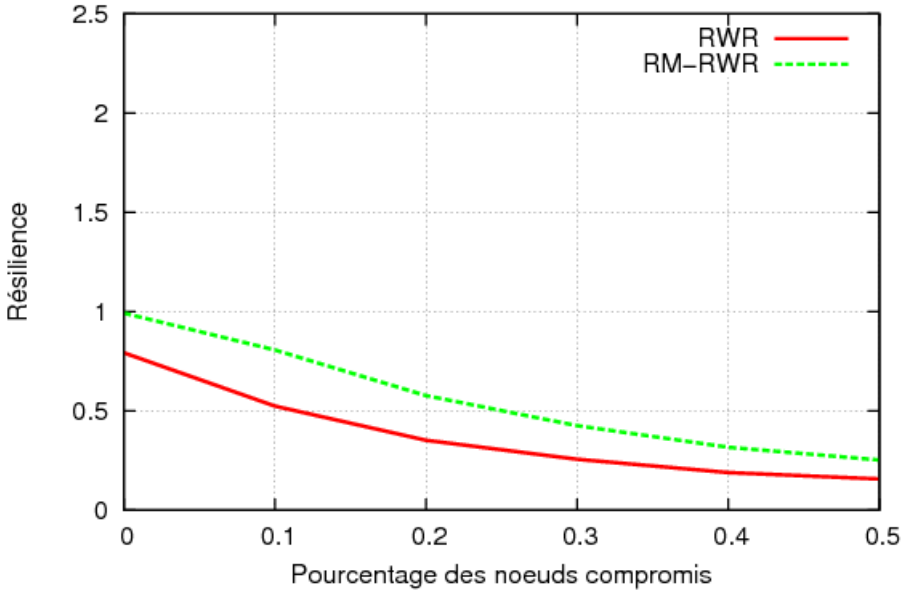


FIGURE 4.9: Évaluation quantitative de la résilience des variantes de RWR.

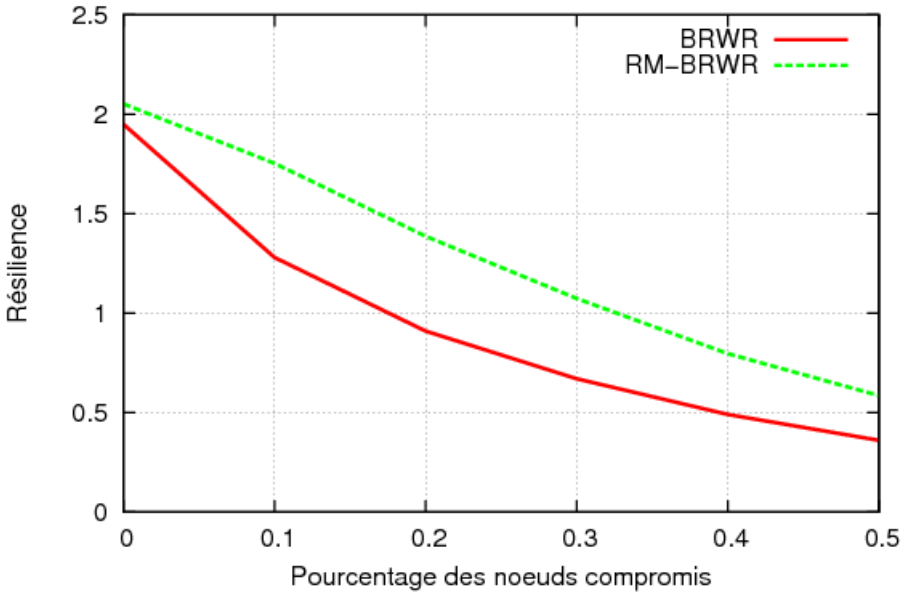


FIGURE 4.10: Évaluation quantitative de la résilience des variantes de BRWR.

L'évaluation quantitative de la résilience des variantes aléatoires des protocoles en calculant la grandeur de surface (Figure 4.6 - 4.8), nous confirme l'amélioration de la résilience par rapport à leur variante classique. Nous observons par exemple, une augmentation de résilience plus importante pour la variante aléatoire de RS-GBR (16%) par rapport sa version classique GBR en cas de 10% d'attaquants. Tandis que pour RS-DSR et RS-GF l'augmentation est plus modeste (6% et 1%, respectivement).

Maintenant, si nous comparons la résilience de tous les protocoles aléatoires (Figure 4.11), nous obtenons la classification suivante dans l'ordre décroissant de résilience :

1. RS-GBR
2. BRWR
3. RS-GF
4. RS-DSR
5. RWR

Même si l'introduction de l'aléatoire a amélioré la résilience de la majorité des protocoles classiques, elle ne permet pas de basculer la relation d'ordre. La raison de cet ordre est donc liée principalement aux stratégies initiales de construction des routes.

Résultats des protocoles aléatoires avec réplications

Nous introduisons de la réplication de paquets aux protocoles aléatoires pour augmenter le succès de livraison de chaque donnée. Comme nous l'avons mentionné préalablement, les protocoles aléatoires ont la possibilité de bénéficier de la redondance de trafic, puisque chaque envoi de paquet est une itération indépendante de la marche aléatoire.

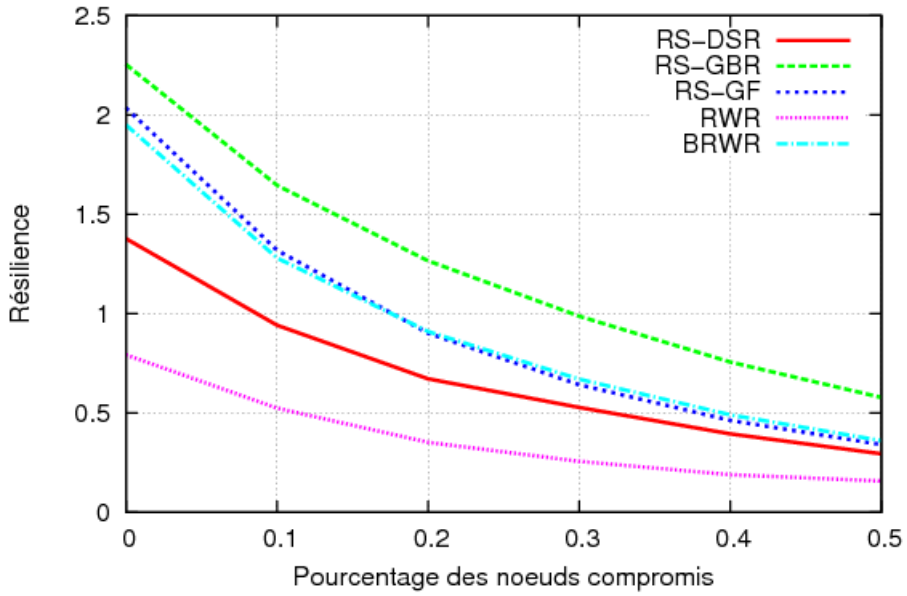


FIGURE 4.11: Évaluation quantitative de la résilience des protocoles de routage aléatoires en cas d'attaques.

Pour les protocoles aléatoires avec réplications (RM-GBR, RM-DSR, RM-GF, RM-RWR et RM-BRWR), la surface de résilience sans ($k = 0\%$) et avec attaques ($k = 30\%$) est présentée sur Figure 4.1 - 4.5 (c).

En cas d'attaques, nous observons une amélioration sur les axes ADR, AT et DF pour tous les protocoles aléatoires avec réplication. Ceci grâce aux multiples chemins, car chaque paquet est répliqué et envoyé sur un chemin construit de manière aléatoire.

Dans la section précédente, avec les variantes aléatoires sans réplication, nous avons déjà observé ce phénomène d'amélioration de DF, l'équité de livraison pour tous les protocoles aléatoires. Avec réplication, l'équité de livraison augmente encore, non seulement parce que les nœuds diversifient leurs routes, mais également parce que les nœuds lointains arrivent à mieux transmettre. Comme nous répliquons les paquets en fonction de la distance au puits, plus un nœud est loin du puits plus il réplique. C'est une façon de compenser la présence de longues routes par la diversité des routes et la redondance de données.

À cause d'une longue route, un paquet a moins de chance d'atteindre le puits. Cependant, un nœud loin du puits a plus de possibilités de varier les routes contrairement aux nœuds proches qui ont peu de choix. Pour ces raisons, l'équité de livraison est augmentée avec la réplication et les nœuds lointains sont moins pénalisés par leur longueur des routes.

Nous observons également une augmentation du taux de livraison moyen (ADR) et du débit moyen (AT) du réseau comparé au scénario sans réplication. La diminution de taux de livraison est donc plus progressive avec l'augmentation de l'intensité des attaques.

Cependant, la réplication des paquets a un coût. Elle augmente considérablement la consommation d'énergie et nous observons en conséquence une diminution notable de l'efficacité énergétique (EE). Les nœuds lointains ont, à l'origine, un coût important pour l'ensemble du réseau : leurs paquets parcourent un plus long chemin. Avec la réplication adaptative, ces

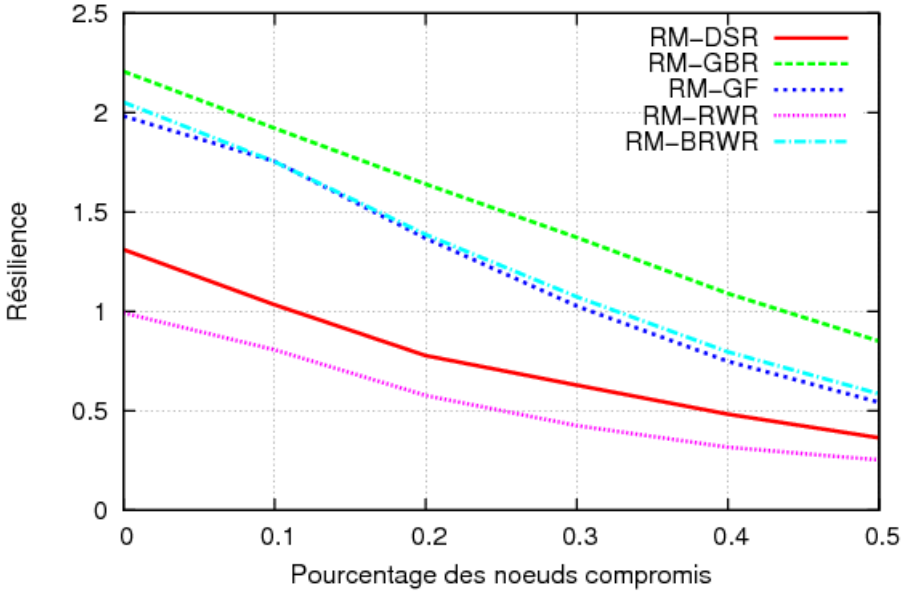


FIGURE 4.12: Évaluation quantitative de la résilience des protocoles aléatoires avec réplication.

nœuds éloignés sollicitent encore davantage le réseau.

Comme discuté précédemment sur le phénomène contre-intuitif concernant la consommation d'énergie : avec l'augmentation des attaques dans le réseau, la consommation d'énergie diminue puisqu'il y a moins de trafic qui circule dans le réseau. Avec la réplication des paquets, nous ajoutons du trafic en augmentant ainsi la consommation d'énergie. Ces deux phénomènes se compensent dans ce scénario. Cependant, nous observons au final une légère augmentation de la consommation d'énergie avec l'intensité des attaques.

Si nous étudions l'évaluation quantitative de la résilience des protocoles aléatoires avec réplication pour chaque protocole (Figure 4.6 - 4.10), nous avons la confirmation que la résilience est améliorée pour tous les protocoles par rapport à leur version classique. La forme de la courbe de résilience est moins convexe comparée aux cas sans réplication. Cela implique une diminution de performance plus progressive, lorsque l'intensité des attaques augmente.

Si nous regardons ensemble tous les protocoles aléatoires avec réplication, la Figure 4.12 confirme que l'introduction de la réplication ne change pas la relation d'ordre des protocoles :

1. RM-GBR
2. RM-BRWR
3. RM-GF
4. RM-DSR
5. RM-RWR

Cela nous montre que même si les mécanismes résilients aident les protocoles, un mauvais choix initial de stratégie de construction des routes ne peut être compensé.

En résumé, nous observons une amélioration de la résilience pour tous les protocoles aléatoires avec réplication comparés aux cas précédents sans réplication. La forme des courbes est moins convexe indiquant une diminution de la résilience moins brutale. Cependant, la résilience a un coût important en termes de consommation d'énergie.

Il est important de savoir combien d'énergie le réseau est capable de dépenser pour obtenir un succès de livraison souhaité pour une meilleure résilience aux attaques.

4.3.4 Taxonomie de la résilience des protocoles de routage

Cette étude par simulations nous montre que l'introduction des mécanismes résilients aux protocoles classiques améliore en effet leur résilience, mais également elle montre que la métrique de résilience que nous avons présentée dans le chapitre 3 permet de saisir la résilience.

Grâce à notre métrique quantitative de résilience, nous avons obtenu une classification des protocoles et nous proposons donc une nouvelle taxonomie de résilience des protocoles de routage (Figure 4.13).

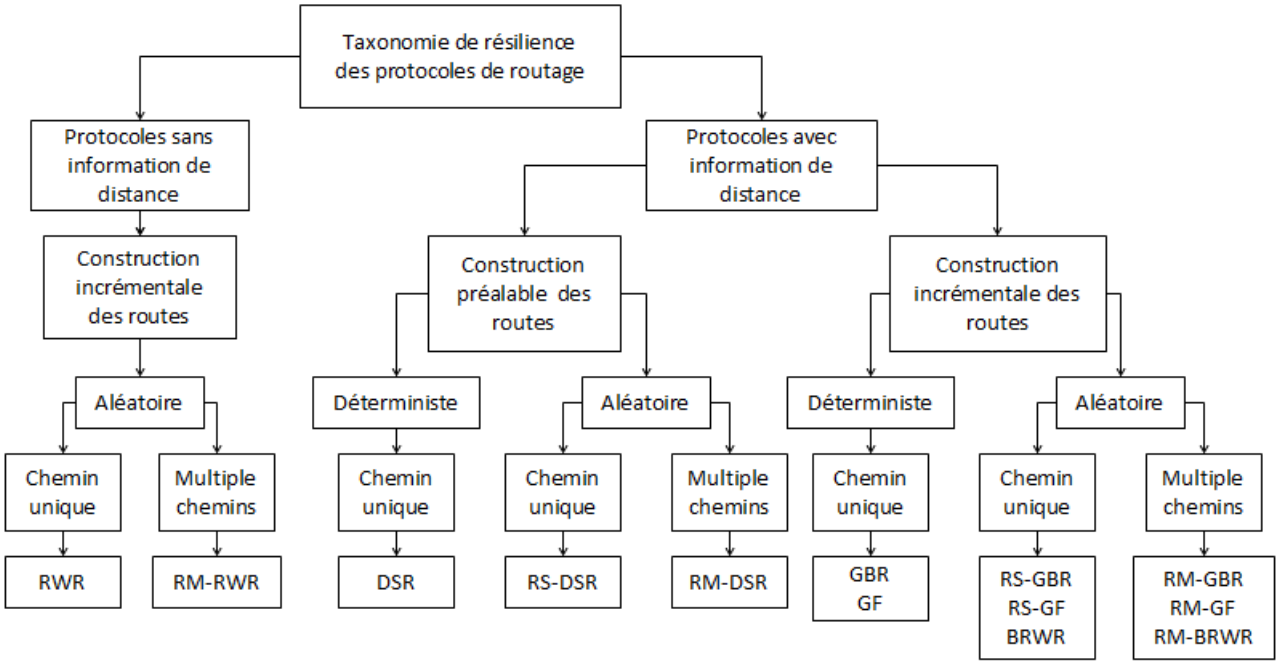


FIGURE 4.13: Taxonomie de la résilience des protocoles de routage.

Nous divisons les protocoles de routage en deux catégories principales : (i) sans information d'état et (ii) avec information d'état. Cette distinction est nécessaire pour comprendre pourquoi certains protocoles sont moins performants. Les protocoles de la catégorie (ii) tels que RWR ne permettent pas d'avoir l'information de la direction du puits au niveau de chaque nœud. Les paquets de données sont donc envoyés à un des voisins choisi aléatoirement sans savoir où se trouve le destinataire. RWR reste très pénalisé par la longueur de ses routes, même si le comportement aléatoire permet une diversification des routes et que la réplication

améliore le succès de livraison. Les protocoles de la catégorie (ii) peuvent envoyer des paquets en utilisant les chemins plus courts puisqu'ils connaissent l'emplacement du puits. Les protocoles de la catégorie (ii) sont divisés en deux sous-catégories (a) avec construction préalable des routes et (b) avec construction incrémentale des routes. Cette classification permet de distinguer les protocoles tels que DSR, utilisant des routes préétablies de bout-en-bout. Ces protocoles donnent la possibilité de construire de multiples chemins disjoints. Cependant, la diversification des routes par les nœuds sources est limitée au nombre de routes construites. La phase de construction des routes préalable est très coûteuse en termes de consommation énergétique. Par exemple, DSR est très gourmand en énergie à cause de l'inondation des paquets `RREQ` par chaque source et des paquets `RREP` envoyés sur le chemin inverse.

Les protocoles de la sous-catégorie (b) tels que GBR, GF et BRWR donnent la possibilité aux nœuds de décider quelles routes construire au moment de l'envoi des paquets de données. Ce mécanisme est plus dynamique parce que les nœuds peuvent s'adapter à leur environnement en ajustant la diversification des routes et la réplication de données.

Cette nouvelle taxonomie des protocoles nous montre que les différents mécanismes de construction des routes et l'information d'état utilisés jouent un rôle important pour la résilience des protocoles.

Les protocoles de routage les plus résilients sont donc les variantes aléatoires de GBR, BRWR et GF avec réplication des paquets, qui ont l'information d'état et qui construisent des routes incrémentalement.

4.4 Conclusion

Dans ce chapitre, nous avons présenté notre proposition des mécanismes résilients. Nous les avons appliqués aux protocoles classiques et nous les avons étudiés par simulations.

Selon notre métrique de résilience, la majorité des variantes aléatoires des protocoles ont une meilleure résilience par rapport aux variantes classiques sans causer de surcoût grâce à la diversité des routes. Les principaux mérites de notre proposition sont les suivants :

- Le taux de livraison moyen (ADR) et le délai moyen (AT) sont améliorés : la dégradation est moins brutale avec l'augmentation des attaques.
- Le taux de livraison est réparti équitablement (DF) parmi les sources : un plus grand nombre de sources arrive à transmettre leurs données au puits et avec un taux de livraison plus élevé. On note également que les nœuds éloignés ont un meilleur taux de livraison.
- La connectivité (DF) est améliorée : un plus grand nombre de sources reste connecté au puits avec l'augmentation des attaques.
- La redondance structurelle de la topologie physique est mieux exploitée et la consommation d'énergie globale (EE) est répartie équitablement parmi les nœuds : plus de nœuds participent au routage.

Cette étude montre donc que notre proposition appliquée aux protocoles classiques permet d'améliorer la résilience. Mais elle permet également de valider la métrique présentée dans le chapitre 3, en illustrant sa capacité à mesurer la résilience. La représentation graphique de la métrique permet de discerner instantanément l'amélioration que notre proposition

apporte sur chaque paramètre. De plus, le calcul quantitatif permet d'établir une relation d'ordre entre les protocoles. Elle nous a donc permis de proposer une nouvelle taxonomie de résilience des protocoles de routage.

La taxonomie nous confirme que le choix de la stratégie de construction des routes reste un aspect important pour la résilience. Certains protocoles tels que DSR ou RWR restent inefficaces même avec la réplication des paquets. DSR est trop pénalisé par la consommation d'énergie initiale à cause de la construction des routes par chaque source tandis que RWR est très pénalisé par la longueur de chemins. Quand à GF, la variante aléatoire présentée dans cette étude engendre également de longues routes. Les protocoles les plus intéressants suite à cette étude sont les variantes aléatoires de GBR et BRWR. Nous proposons donc, dans le chapitre suivant, une étude par simulations de GBR et ses variantes aléatoires contre plusieurs attaques combinées. Nous proposons également une étude analytique de BRWR dans le dernier chapitre de contribution.

Étude de la résilience du routage par gradient en cas de plusieurs attaques combinées

5

Sommaire

5.1 Introduction 75
5.2 Mécanismes résilients appliqués au routage par gradient (GBR) 76
5.2.1 GBR classique 76
5.2.2 GBR aléatoires sans réplication 76
5.2.3 GBR aléatoires avec réplications 77
5.3 Modèles d'adversaires 78
5.3.1 *Selective forwarding* basique 79
5.3.2 *Sinkhole* combinée 79
5.3.3 *Sybil* combinée 79
5.3.4 *Wormhole* combinée 81
5.4 Évaluation des performances 81
5.4.1 Résultats en cas d'attaque *Selective forwarding* basique 82
5.4.2 Résultats en cas d'attaque *Sybil* combinée 90
5.4.3 Résultats en cas d'attaque *Wormhole* combinée 94
5.4.4 Résultats en cas d'attaque *Sinkhole* combinée 98
5.5 Conclusion 100

5.1 Introduction

Dans le chapitre 4, nous avons évalué plusieurs protocoles (DSR [26], GF [28], GBR [27], RWR [29] et BRWR [2]) de différentes catégories selon notre métrique de résilience. Cette étude nous a permis de proposer une nouvelle taxonomie des protocoles. Selon cette taxonomie, les variantes de routage par gradient GBR ont donné les meilleurs résultats.

GBR [27] est un protocole adapté aux réseaux de capteurs, puisqu'il n'a besoin d'aucun équipement spécial tel que GPS. Il suffit d'une seule inondation d'un paquet de contrôle par le puits pour établir l'information de gradient. Grâce à l'indication de la direction du puits (distance en nombre de sauts), les données des sources convergent vers le puits. Certains protocoles de routage pour les réseaux de capteurs tels que Directed Diffusion [113] ou RPL [114] (normalisation IETF) considèrent ce type de mécanisme. Une fois le gradient établi, les sources ont l'information de distance au puits, mais également celles de leur voisinage. Cela permet à chaque nœud de classer ses voisins directs dans des sous ensembles en fonction de leur distance au puits. Chaque source a donc la possibilité de choisir son prochain saut dans un sous-ensemble, avec une probabilité associée, afin de diriger le paquet plus ou moins rapidement vers le puits. C'est une façon de calibrer le biais d'une marche aléatoire. Grâce à ce mécanisme, il est possible d'adapter la diversification des routes. Pour ces raisons, nous avons choisi GBR pour approfondir notre étude par simulations.

Notre but dans ce chapitre est d'étudier plusieurs valeurs de biais, où les paquets ont la possibilité d'être relayés par des nœuds de même niveau. Cela permet d'augmenter mécaniquement le nombre de candidats pour le prochain saut et de mieux exploiter les routes alternatives possibles.

Dans nos études précédentes, nous nous sommes concentrés principalement sur les attaques de non-retransmission des paquets de données (*Selective forwarding*). Dans ce chapitre, nous voulons étendre notre modèle d'adversaires à des attaques plus complexes, mais aussi la combinaison de plusieurs attaques par un nœud compromis. Le but d'un attaquant est non seulement de jeter des paquets de données passant par lui, mais également de tenter d'attirer plus de trafic vers lui pour amplifier son impact. Par exemple, dans l'attaque *Sybil*, un adversaire peut créer plusieurs fausses identités en mentant sur les paquets de contrôle. Les nœuds compromis peuvent s'entendre entre eux en créant des communications privées, produisant ainsi des attaques *Wormhole*. Un adversaire peut également compromettre les nœuds autour du puits en créant des attaques *Sinkhole*. En profitant du trafic convergeant vers le puits, cela permet d'attirer la majorité du trafic du réseau.

En résumé, nous proposons d'étendre notre étude par simulations sur les trois points suivants : (i) l'étude de l'impact de plusieurs attaques combinées (*Sybil*, *Wormhole* et *Sinkhole* avec *Selective forwarding*) (ii) l'étude de différentes valeurs de biais ($p = \{1, 0.8, 0.6, 0.5\}$) pour introduire de comportements aléatoires à GBR et (iii) l'étude de différentes façons de répliquer les paquets.

5.2 Mécanismes résilients appliqués au routage par gradient (GBR)

Nous introduisons les mécanismes résilients présentés dans le chapitre 4 au routage par gradient GBR. Contrairement au chapitre 4, nous proposons ici d'étudier plusieurs valeurs de biais introduits pour différents GBR aléatoires. Ensuite, nous introduisons deux types de réplications des paquets. Enfin, nous évaluons la résilience de ces protocoles en cas de plusieurs attaques combinées.

5.2.1 GBR classique

Dans la version classique du routage par gradient GBR-Dt (*Deterministic*), les routes sont construites en se basant sur le critère du plus court chemin pour une meilleure efficacité.

Un nœud source s envoie son paquet `DATA` à un voisin direct avec la "hauteur" minimale (distance minimum en nombre de sauts) afin de faire un maximum de progrès vers le puits. Le candidat du prochain saut v_i est choisi dans $V(s)$, $1 \leqslant i \leqslant m_s$ (Tableau 5.1). Si plusieurs voisins ont la même hauteur minimale, il choisit le premier de la table de voisinage et il le garde.

Nous ne considérons pas de réplication des paquets pour la version classique de GBR. Chaque copie empruntant le même chemin, la réplication n'a pas d'intérêt dans le cas de protocoles déterministes.

5.2.2 GBR aléatoires sans réplication

Les variantes aléatoires du routage par gradient GBR-Rd (*Random*) avec $p = \{1, 0.8, 0.6, 0.5\}$ sont des marches aléatoires biaisées par l'information de gradient indiquant la direction du puits.

Un nœud source s envoie donc son paquet `DATA` à un voisin choisi aléatoirement avec la probabilité p dans l'ensemble des voisins les plus proches du puits $v_i \in V(s)$, $1 \leqslant i \leqslant m_s$ (Table 5.1) et avec la probabilité $1-p$ dans l'ensemble des voisins de même niveau $w_j \in W(s)$, $1 \leqslant j \leqslant l_s$ (Table 5.1).

En d'autres mots, le paquet progresse vers le puits avec la probabilité p et reste au même niveau avec la probabilité $1 - p$. Un exemple est illustré sur la Figure 5.1 pour la variante GBR-Rd $p = 0.8$. Le nœud A va choisir aléatoirement le prochain saut parmi D et E avec la probabilité $p = 0.8$, et parmi B et C avec la probabilité $p = 0.2$. Notre but est de déterminer l'impact des différentes valeurs de p sur la résilience.

Nous avons choisi la valeur $p = 1$, car c'est la marche aléatoire la plus biaisée : à chaque saut, le paquet progresse vers le puits. Les chemins construits sont toujours les plus courts. GBR-Rd $p = 1$ est donc équivalent au protocole RS-GBR décrit précédemment dans le chapitre 4. Nous avons choisi la valeur $p = 0.5$, car le paquet progresse vers le puits seulement une fois sur deux en moyenne. Le paquet est envoyé à un voisin de même niveau le reste du temps. GBR-Rd $p = 0.5$ est donc équivalent au protocole BRWR étudié précédemment dans le chapitre 2. Nous avons évalué également les deux valeurs intermédiaires $p = 0.8$ et $p = 0.6$,

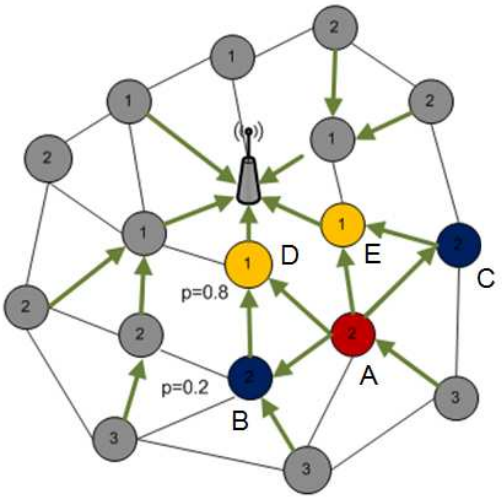


FIGURE 5.1: Exemple de fonctionnement de GBR-Rd $p = 0.8$.

TABLE 5.1: Notations

Notations	Descriptions
s	un nœud source du réseau
h_s	la hauteur de s
$U(s) = \{u_1, u_2, ..., u_{n_s}\}$	l'ensemble des voisins de s
$V(s) = \{v_1, v_2, ..., v_{m_s}\}$	l'ensemble des voisins de s avec la hauteur $< h_s$
$W(s) = \{w_1, w_2, ..., w_{l_s}\}$	l'ensemble des voisins de s avec la hauteur $= h_s$

car le paquet progresse vers le puits la majorité du temps et reste sur le même niveau la minorité du temps.

5.2.3 GBR aléatoires avec réplications

Nous considérons deux types de réplications : (i) réplications adaptatives et (ii) réplications uniformes (duplication des paquets à la source $x = 2$).

Dans le chapitre 4, nous avons proposé une réplication adaptative (i) en fonction de la distance au puits, en partant de l'observation qu'avec de longues routes, la probabilité de tomber sur un nœud attaquant augmente. Les nœuds lointains répliquent plus et les nœuds proches du puits répliquent moins. Cependant, même si le taux de réplication dépend de la distance au puits, nous avons limité le nombre maximum de réplications par nœud à trois. Cette limitation est nécessaire, car les nœuds lointains peuvent engendrer un trafic important, surtout dans un réseau étendu avec un diamètre important. De plus, pour un réseau avec un seul puits au milieu de la zone de déploiement, le nombre de nœuds par niveau (distance en nombre de sauts) n'est pas uniforme : nous avons un plus grand nombre de nœuds à 3 sauts du puits qu'à 2 sauts. Comme la réplication se fait en fonction de la distance au puits en nombre de sauts et comme il y a un plus grand nombre de nœuds au

niveau 3 qu'au niveau 2, l'augmentation du trafic est d'autant plus importante.

Enfin, sans cette limitation et étant donnée la dépendance du taux de réplication à l'information de distance, une attaque simple de déni de services serait de faire croire à un nœud qu'il est loin du puits pour qu'il réplique indéfiniment, augmentant le trafic du réseau et épuisant inutilement son énergie.

Nous proposons également d'étudier une réplication uniforme (ii), où chaque nœud réplique avec un taux fixe (x fois), prédéterminé quelle que soit sa distance au puits. Dans cette étude, nous considérons $x = 2$, où chaque paquet est dupliqué à la source. L'avantage de cette méthode est sa simplicité et son indépendance par rapport à l'information d'état. Ce mécanisme permet d'éviter des attaques de déni de service décrit précédemment, où un attaquant peut amener un nœud cible à croire qu'il est loin du puits pour qu'il réplique plus.

Quel que soit le mécanisme choisi, la réplication des paquets permet de profiter de la diversité des routes crées par des protocoles aléatoires. Chaque paquet peut prendre des routes potentiellement différentes et le choix des routes n'est pas prévisible.

5.3 Modèles d'adversaires

Nous considérons les hypothèses du réseau identiques à l'étude précédente 2.2.1. Pour les modèles d'adversaires, nous considérons les attaques plus complexes, parce que le but d'un adversaire est de perturber le bon déroulement des protocoles de routage pour causer le maximum de dégâts possible à la transmission des données entre les sources et le puits.

Nous avons identifié les mécanismes [1] qui peuvent être combinés avec la non retransmission des paquets de données (*Selective forwarding*) pour perturber efficacement l'acheminement des données.

Notre objectif est d'explorer les modèles d'adversaires plus complexes et plus réalistes. Après avoir compromis des nœuds, un adversaire a la possibilité de produire plusieurs attaques simultanément. Par exemple, l'attaque *Sybil* seule permet de créer de multiples identités non existantes ou de voler l'identité d'autres nœuds du réseau. Un nœud compromis peut ainsi générer de fausses informations de topologie et remplir la mémoire des ses voisins avec des informations inutiles. Cependant, en combinant l'attaque *Sybil* avec *Selective forwarding*, un adversaire peut causer des dégâts bien plus importants que si elle était produite séparément. Parce qu'en créant plusieurs identités, l'adversaire augmente la chance d'être choisi comme prochain saut. Il peut donc attirer plus de paquets de données vers lui avant de les éliminer.

Nous considérons donc les modèles d'adversaires, où les nœuds compromis tentent d'attirer le trafic en combinant les attaques *Sybil*, *Wormhole* et *Sinkhole* avec *Selective forwarding*.

1. (i) l'introduction de plusieurs identités (attaque *Sybil*), (ii) l'entente des attaquants via une connexion privée (attaque *Wormhole*), (iii) la compromission des nœuds proches du puits (attaque *Sinkhole*)

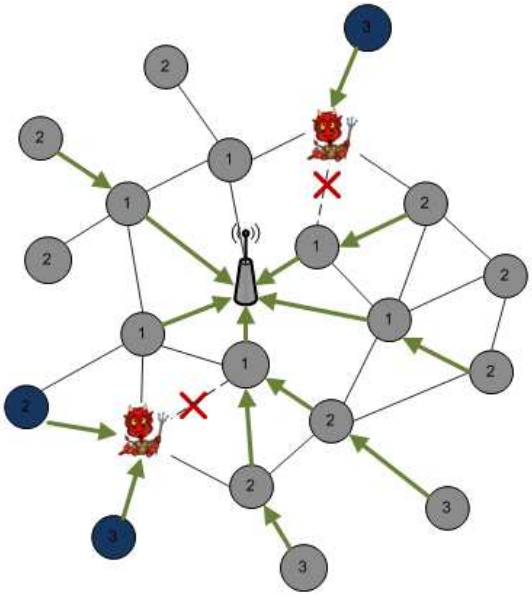


FIGURE 5.2: Attaque *Selective forwarding* basique.

5.3.1 *Selective forwarding* basique

Le *Selective Forwarding* (Figure 5.2) est une attaque où les nœuds malveillants ne retransmettent pas les paquets de données. Les hypothèses considérées ici sont identiques à celles de notre étude préliminaire de la section 2.2.1.

5.3.2 *Sinkhole* combinée

L'attaque *Sinkhole* combinée avec le *Selective forwarding* (Figure 5.3) consiste à attirer plus de trafic en se plaçant près du puits et jeter ensuite les paquets de données. Les hypothèses considérées ici sont également identiques à celles de notre étude préliminaire de la section 2.2.1.

5.3.3 *Sybil* combinée

L'attaque *Sybil* combinée avec le *Selective forwarding* (Figure 5.4) attire du trafic en introduisant plusieurs identités et en jetant ensuite les paquets reçus. Les nœuds compromis sont distribués de la même manière que dans l'attaque par *Selective forwarding* basique. Selon la taxonomie de l'attaque *Sybil* [43], notre modèle correspond à une "communication directe". Les nœuds *Sybil* communiquent directement avec les nœuds légitimes. De plus, ces premiers utilisent des "identités fabriquées", ils créent arbitrairement de nouvelles identités non existantes dans le réseau. Enfin, notre modèle est de type "simultané" : un attaquant peut utiliser simultanément toutes ses identités *Sybil*.

Avec ce modèle d'adversaire, les nœuds malveillants prennent deux identités. Un nœud compromis perturbe la phase d'établissement du gradient en dupliquant les paquets `INIT` et

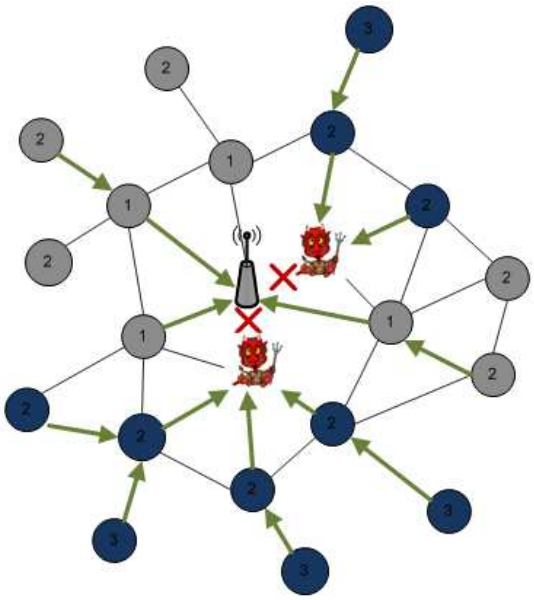


FIGURE 5.3: Attaque *Sinkhole* combinée.

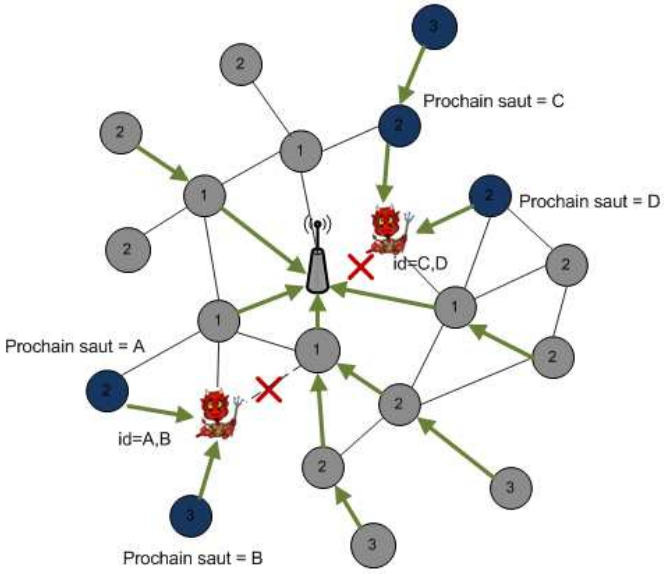


FIGURE 5.4: Attaque *Sybil* combinée.

en y mettant une fausse identité. L'objectif est de simuler la présence de plusieurs voisins là où il n'y a en réalité qu'un seul nœud physique. La probabilité d'être choisi pour prochain saut augmente donc pour un nœud malveillant, lui permettant ainsi d'attirer plus de trafic. Un nœud compromis ne fausse pas son gradient et les deux identités prennent le même vrai gradient. Nous avons choisi cette stratégie particulière pour séparer l'impact de l'attaque *Sinkhole* de l'attaque *Sybil*. La fausse identité est choisie aléatoirement dans un grand intervalle pour éviter les collisions entre identités. Une fois les deux identités créées, un nœud malveillant jette tous les paquets `DATA` en provenance de ses voisins.

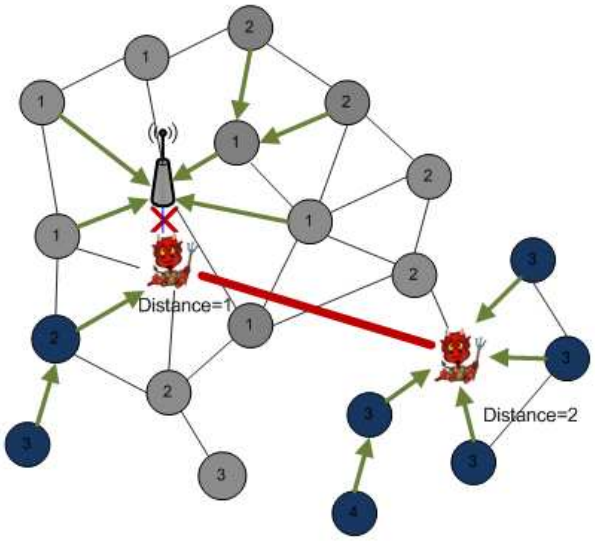


FIGURE 5.5: Attaque *Wormhole* combinée.

5.3.4 *Wormhole* combinée

L'attaque *Wormhole* combinée avec le *Selective forwarding* (Figure 5.5) définit l'entente entre deux nœuds malveillants pour attirer le trafic et jeter ensuite les paquets. Les deux nœuds malveillants se déclarent voisins, même s'ils sont éloignés physiquement. Ils créent un tunnel entre eux, via une connexion privée.

Chaque paire de nœuds malveillants $(w1, w2)$ avec une distance supérieure à deux sauts va créer un lien *Wormhole*. Un paquet `INIT` reçu par $w1$ est directement transmis à $w2$ via le tunnel. Les paquets `INIT` arrivent ainsi plus tôt que les autres paquets transmis sur une connexion normale. Si $w1$ est placé près du puits, $w2$ va obtenir un plus petit gradient que ses voisins et $w2$ pourra donc attirer du trafic de ses voisins. Le pourcentage k des nœuds malveillants est distribué aléatoirement sur l'ensemble du réseau. Le nombre total de liens *Wormhole* est $k/2$. Pour les simulations k varie entre 10% et 50% de la population de nœuds du réseau. Une fois le lien *Wormhole* créé entre deux nœuds malveillants $(w1, w2)$, ces derniers jettent tous les paquets `DATA` provenant de leurs voisins. Un nœud malveillant appartient à un seul lien *Wormhole* et le cas de plusieurs liens *Wormhole* provenant d'un seul nœud n'est pas considéré ici.

5.4 Évaluation des performances

Nous considérons des paramètres de simulations identiques aux chapitres précédents qui sont résumés sur le Tableau 2.1.

L'objectif de notre étude est de comparer plusieurs variantes de GBR (GBR-Dt, GBR-Rd $p = \{1, 0.8, 0.6, 0.5\}$) selon les trois scénarios suivants :

1. Sans réplication
2. Avec réplications uniformes

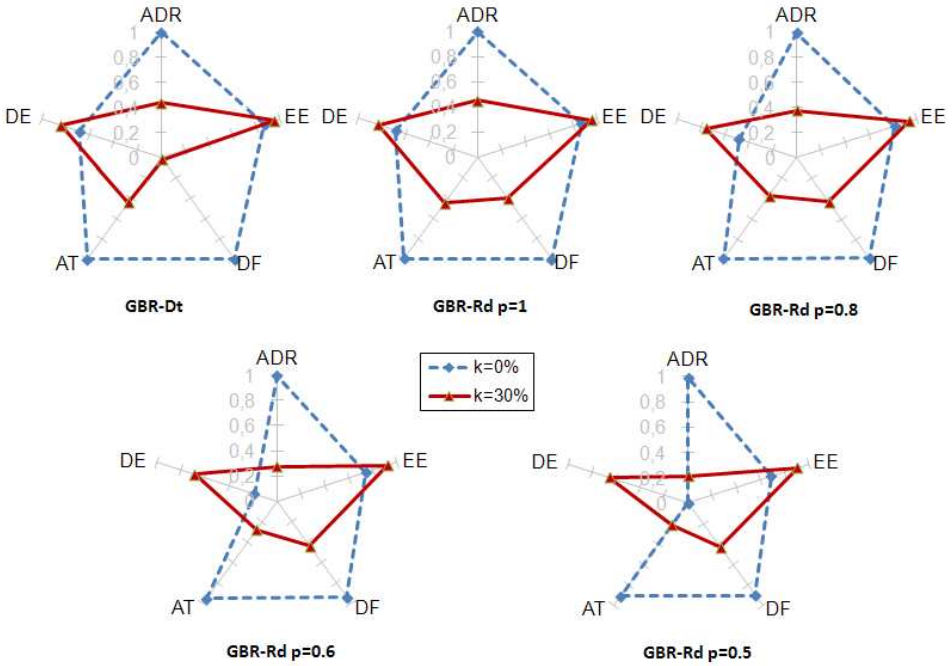


FIGURE 5.6: Surface de résilience en cas d'attaque *Selective forwarding* : sans réplication.

3. Avec réplications adaptatives

Nous évaluons leur résilience avec notre métrique pour quatre types d'attaques : *Selective forwarding* basique, *Sybil*, *Wormhole* et *Sinkhole* combinées.

5.4.1 Résultats en cas d'attaque *Selective forwarding* basique

Les résultats de simulations en cas d'attaque *Selective forwarding* pour le protocole GBR-Rd $p = 1$ sont identiques aux résultats du protocole RS-GBR du chapitre 4. Cependant, nous l'avons comparé précédemment aux protocoles de différentes catégories, tandis que dans ce chapitre, nous comparons plusieurs variantes de GBR entre elles. Le but de cette étude est d'approfondir les résultats de simulations en prenant plusieurs valeurs de biais, mais également l'impact de plusieurs attaques combinées.

La surface de résilience des cinq variantes de GBR en cas d'attaque *Selective forwarding* est présentée sur la Figure 5.6 - 5.8 selon les trois scénarios.

Pour l'évaluation de la résilience, nous rappelons les 5 paramètres de performances et leur ordre comme nous avons considéré au chapitre 3.

1. ADR - Taux de livraison moyen
2. EE - Efficacité de la consommation énergétique
3. DF - Équité de livraison du réseau
4. AT - Débit moyen
5. DE - Efficacité du délai moyen

5.4 Évaluation des performances

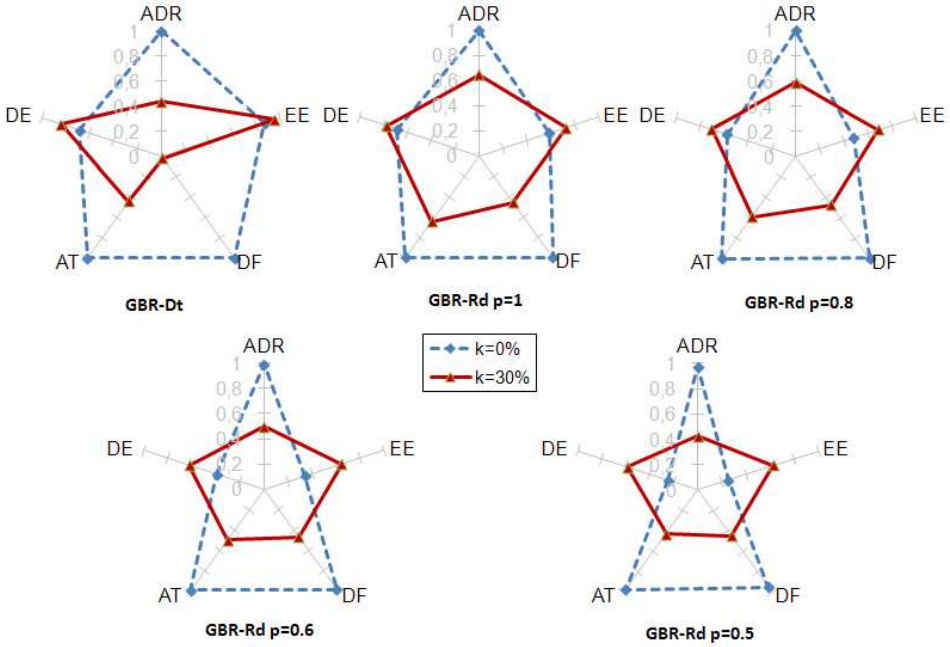


FIGURE 5.7: Surface de résilience en cas d'attaque *Selective forwarding* : avec réplications uniformes.

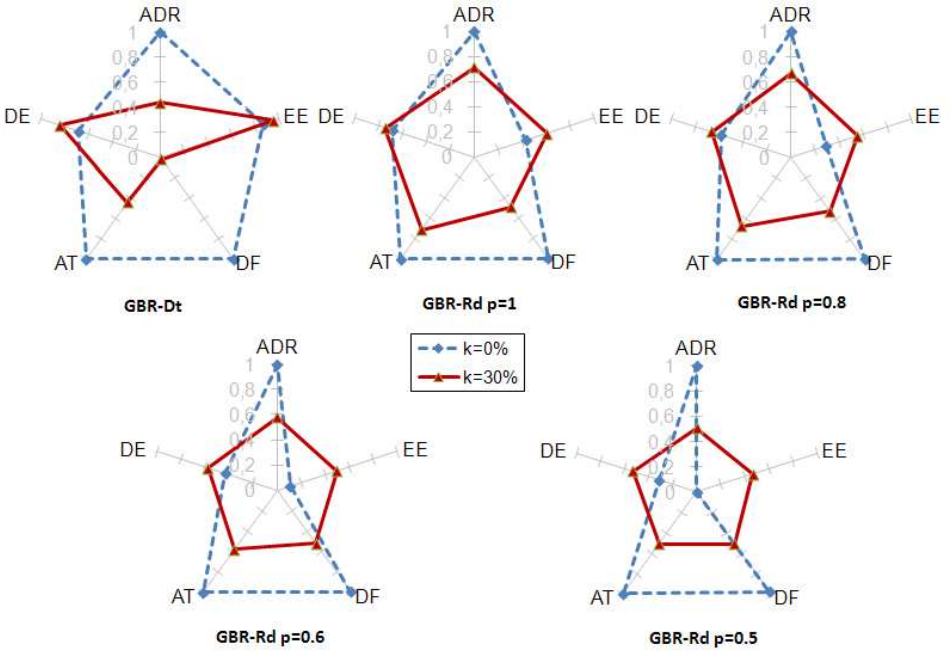


FIGURE 5.8: Surface de résilience en cas d'attaque *Selective forwarding* : avec réplications adaptatives.

Les surfaces de résilience du premier scénario sans réplication des paquets sont montrées sur la Figure 5.6. Sans attaque, nous observons une diminution de l'efficacité de délai (DE)

pour les variantes aléatoires avec la diminution du biais p. Dans nos simulations, le délai est directement proportionnel à la longueur moyenne de chemins. Avec la diminution du biais p, les chemins s'allongent et donc l'efficacité de délai est diminuée. Cela engendre également une légère diminution de l'efficacité d'énergie (EE), car le réseau dépense plus d'énergie pour transmettre un paquet parcourant de longues routes. Une valeur de p proche de 1, implique qu'un paquet se dirige plus vite vers le puits. La version classique GBR-Dt et la variante aléatoire GBR-Rd $p = 1$ se basent sur le critère du plus court chemin et ils ont donc exactement les mêmes valeurs sans attaque.

En cas d'attaque *Selective forwarding* ($k = 30\%$), nous constatons la dégradation du taux de livraison moyen (ADR), du débit moyen (AT) et de l'équité de livraison (DF). Cette dégradation est causée par la diminution du trafic due aux paquets jetés par les attaquants. Dans le même temps, l'efficacité en délai (DE) et en énergie (EE) s'améliore avec les attaques. Nous avons constaté le même phénomène dans le chapitre 3 : le réseau dépense moins d'énergie en cas d'attaque en raison de la diminution du trafic liée à la perte de paquets. Le délai moyen diminue, puisque les paquets reçus au puits viennent essentiellement des nœuds les plus proches du puits. En cas d'attaque de non-retransmission des paquets, les nœuds lointains n'arrivent plus à atteindre le puits. Ces résultats ne sont donc pas surprenants par rapport à l'étude précédente, car nous avons étudié les mêmes attaques.

Nous observons une différence importante entre GBR-Dt et les variantes aléatoires en termes d'équité de livraison (DF). Avec les protocoles déterministes, le succès de livraison n'est pas équitablement distribué parmi les nœuds : certains nœuds ont un très bon taux de livraison tandis que d'autres nœuds sont complètement déconnectés du puits, puisque chaque source utilise toujours la même route. La différence en équité de livraison (DF), entre la version classique GBR-Dt et les variantes aléatoires GBR-Rd est très visible. En revanche, les variantes aléatoires sont très proches en équité de livraison (DF). Nous illustrons donc sur la Figure 5.9, la différence de la distribution de taux de livraison en regroupant les nœuds en 5 classes ($c1$ à $c5$).

- Classe $c1$: nombre de nœuds avec $ADR = 100\%$;
  tous les paquets `DATA` envoyés par ces nœuds sont reçus au puits.
- Classe $c2$: nombre de nœuds avec $ADR \in [66\%; 100\%[$
- Classe $c3$: nombre de nœuds avec $ADR \in [33\%; 66\%[$
- Classe $c4$: nombre de nœuds avec $ADR \in]0\%; 33\%[$
- Classe $c5$: nombre de nœuds avec $ADR = 0\%$;
  aucun des paquets `DATA` n'est reçu au puits. Ces nœuds sont donc totalement déconnectés du puits.

Nous avons également regroupé les nœuds en fonction de leur distance au puits en nombre de sauts ($h1$ à $h4$). Pour le protocole GBR-Dt, nous avons la confirmation que la distribution de taux de livraison n'est pas équitable parce qu'il y a seulement deux classes ($c1$ et $c5$) qui apparaissent. Pour un nœud, soit tous ses paquets sont livrés avec succès et il n'y a pas de nœud malveillant sur sa route, soit ils sont tous perdus et il y a la déconnexion du nœud. La différence de l'équité de livraison entre les variantes aléatoires apparaît également sur la Figure 5.9. Avec le GBR-Rd $p = 1$, il y un faible nombre de nœuds déconnectés (de classe $c5$). Tandis que pour les autres variantes aléatoires, aucun nœud n'est déconnecté du puits

5.4 Évaluation des performances

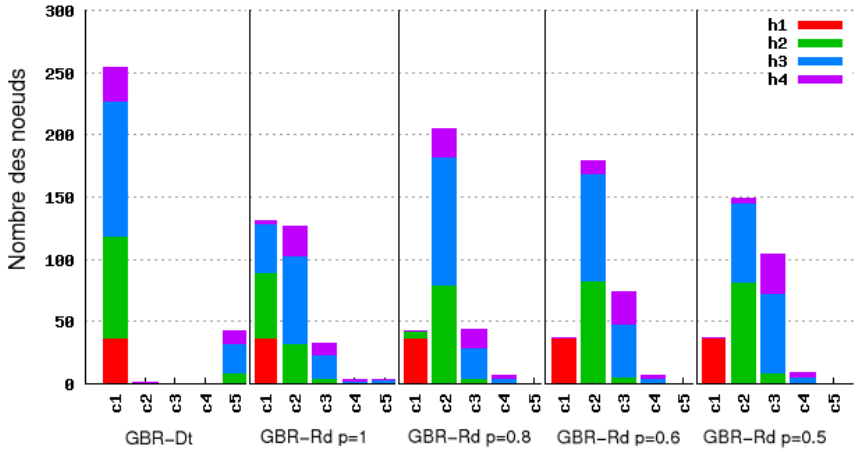


FIGURE 5.9: Distribution de taux de livraison en cas de $k = 10\%$ d'attaquants : Cinq classes $c1$ à $c5$ et quatre distances $h1$ à $h4$ en nombre de sauts.

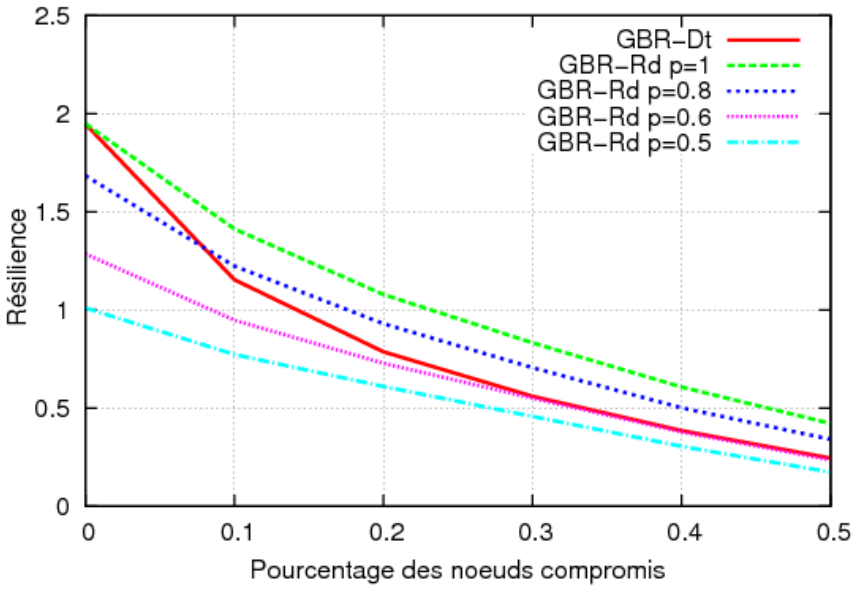


FIGURE 5.10: Résilience en cas d'attaque *Selective forwarding* : sans réplication.

(classe $c5$ n'est plus présente). La diminution du biais p permet une meilleure connectivité des sources même avec un taux de livraison dégradé.

Si nous examinons la résilience du scénario sans réplication sur la Figure 5.10, nous observons que la variante GBR-Rd $p = 1$ est plus résiliente que la version classique GBR-Dt sans coût supplémentaire (performance égale sans attaque).

Il est à noter également que la courbe de résilience de GBR-Dt se dégrade de façon brutale avec l'augmentation d'attaques (Figure 5.10). Cette dégradation brutale indique que GBR-Dt est particulièrement sensible aux attaques et que GBR-Dt absorbe mal l'impact des attaques.

La variante GBR-Rd $p = 0.8$ est également plus résiliente que GBR-Dt, mais avec un coût

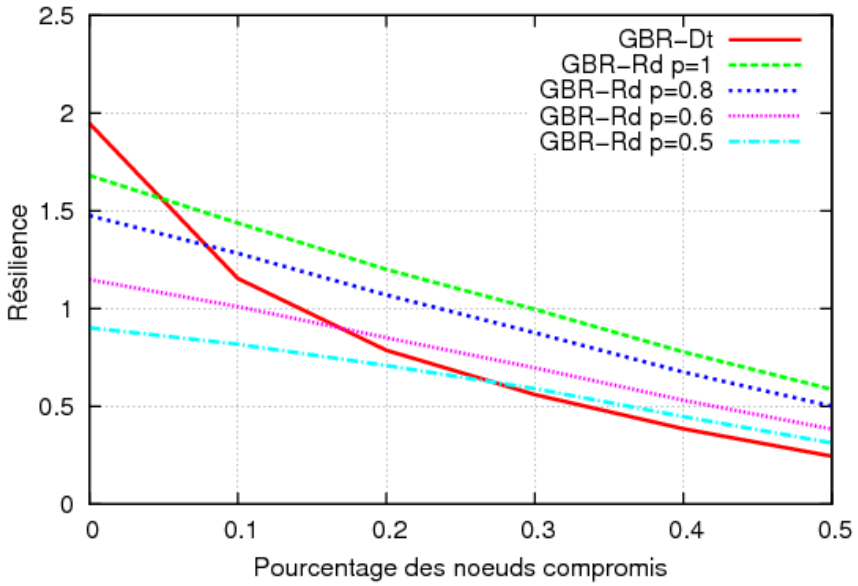


FIGURE 5.11: Résilience en cas d'attaque *Selective forwarding* : avec réplications uniformes.

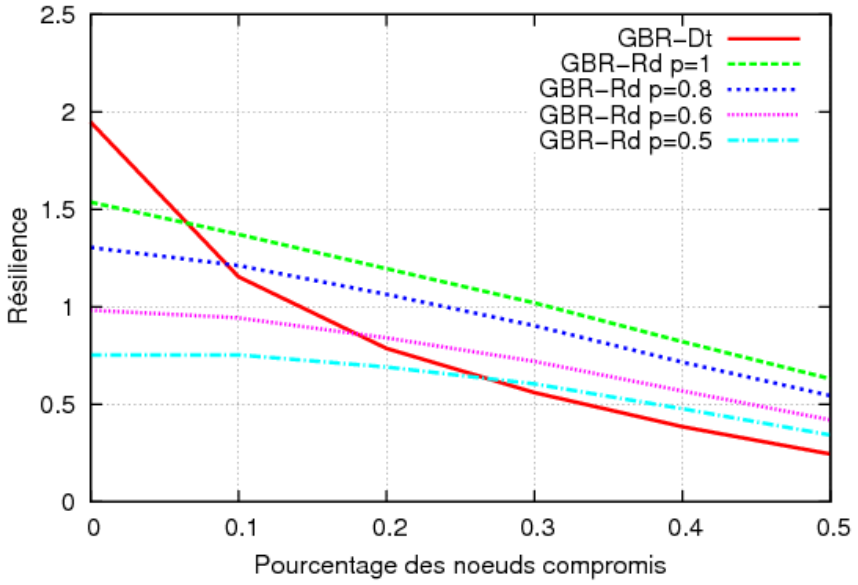


FIGURE 5.12: Résilience en cas d'attaque *Selective forwarding* : avec réplications adaptatives.

supplémentaire (moins performante sans attaque). Tandis que les variantes GBR-Rd $p = 0.6$ et $p = 0.5$ sont plus pénalisées par la longueur des routes.

Si nous regardons les résultats du scénario avec réplications (Figures 5.7 et 5.8), la surface de résilience permet de saisir instantanément l'amélioration du taux de livraison (ADR), du débit (AT) et de l'équité de livraison (DF) par rapport au scénario sans réplication. Nous remarquons une forme plus équilibrée de la surface de résilience ($k = 30\%$) pour les variantes aléatoires avec réplications. Elles ne sont pas pénalisées par un paramètre en particulier. Tous les paramètres sont donc équitablement impactés par les attaques. Toutes les variantes

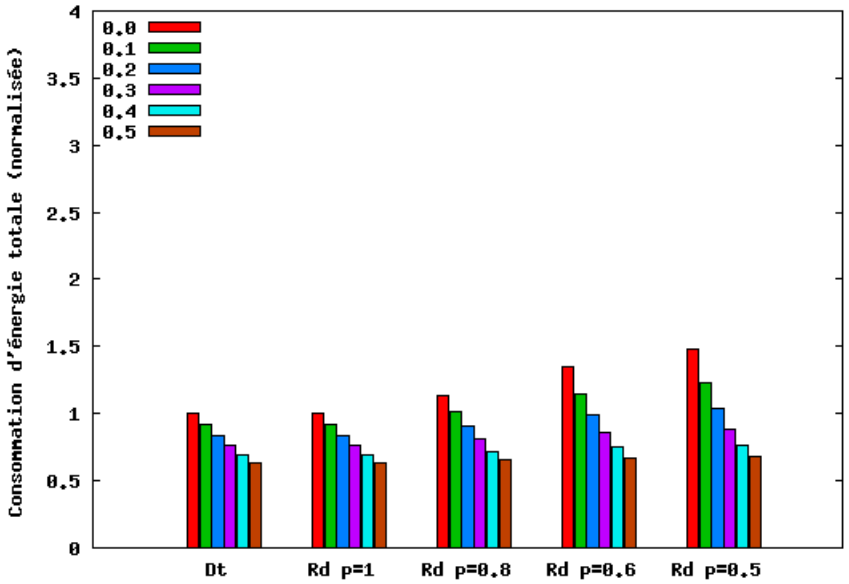


FIGURE 5.13: Consommation totale d'énergie (normalisée par GBR-Dt sans réplication) en cas d'attaque *Selective forwarding* : sans réplication.

aléatoires ont une meilleure résilience par rapport à la version classique, y compris GBR-Rd $p = 0.5$ au-delà de 30% d'attaquants (Figure 5.11). La faiblesse de la longueur des chemins est ainsi compensée par le bénéfice de la réplication des paquets.

La différence des réplications adaptatives (Figure 5.8) par rapport aux réplications uniformes (Figure 5.7) apparaît sur la consommation d'énergie (EE) plus élevée. Afin de percevoir en détail le surcoût énergétique des réplications, nous mesurons la consommation totale d'énergie de toutes les variantes, normalisée par le protocole GBR-Dt sans réplication (Figure 5.13), avec réplications uniformes (Figure 5.14) et adaptatives (Figure5.15).

Les réplications engendrent donc un surcoût énergétique non négligeable. Ce constat est particulièrement notable pour les réplications adaptatives. Les nœuds lointains sont plus nombreux que les nœuds proches et ils répliquent plus. Les nœuds dépensent donc plus d'énergie pour acheminer leurs paquets au puits (Figure 5.15). Cependant, ce surcoût permet d'améliorer le taux de livraison (ADR), le débit moyen (AT) et l'équité de livraison (DF) comparée aux réplications uniformes.

L'introduction des réplications de paquets améliore le taux de livraison (ADR) et nous observons la diminution de ce dernier avec l'augmentation d'attaques pour les cas avec réplications (Figures 5.17-5.18) et pour le scénario sans réplication (Figures 5.16).

Les courbes de taux de livraison du scénario sans réplication sont convexes et elles diminuent brutalement, quel que soit le biais (Figures 5.16). Tandis qu'avec les réplications (Figures 5.17-5.18), la forme des courbes devient plus concave, et le taux de livraison diminue de façon progressive. Cette tendance à la concavité des courbes est un signe de résilience que nous avons également observé sur les courbes de résilience précédemment. C'est ici particulièrement visible.

En résumé, sans attaque les variantes basées sur le plus court chemin GBR-Dt et GBR-Rd

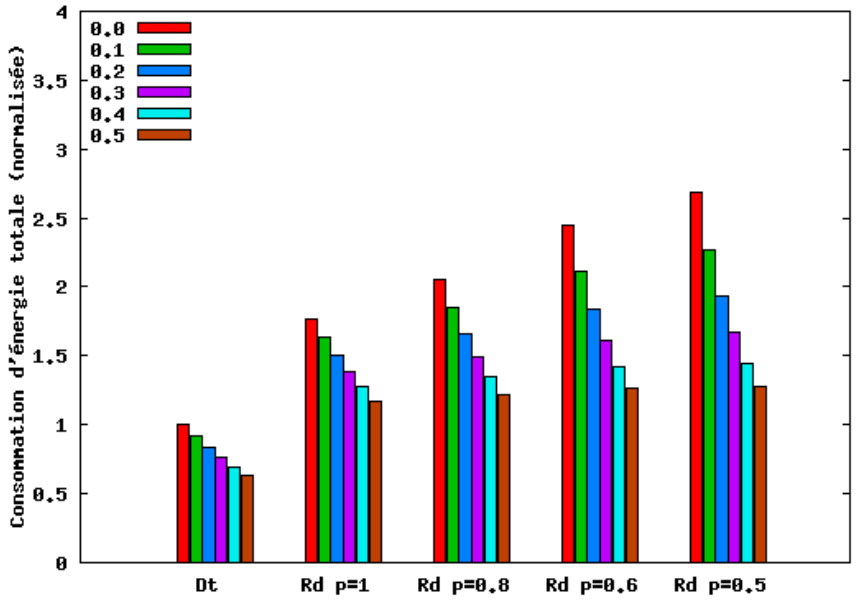


FIGURE 5.14: Consommation totale d'énergie (normalisée par GBR-Dt sans réplication) en cas d'attaque *Selective forwarding* : avec réplications uniformes.

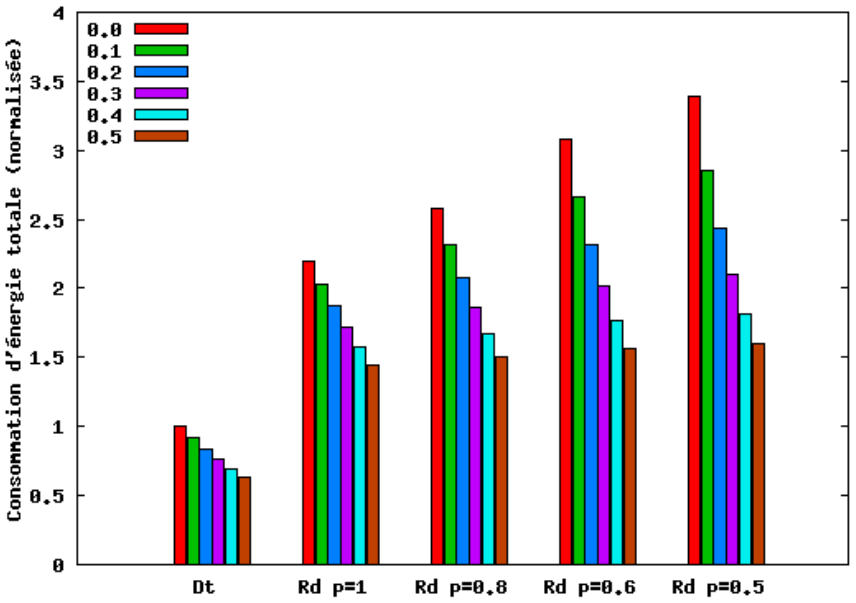


FIGURE 5.15: Consommation totale d'énergie (normalisée par GBR-Dt sans réplication) en cas d'attaque *Selective forwarding* : avec réplications adaptatives.

$p = 1$ sans réplication sont les plus efficaces. En cas d'attaque, GBR-Rd $p = 1$ est plus résiliente que GBR-Dt, mais GBR-Rd $p = \{0.8, 0.6, 0.5\}$ améliorent la connexité des sources. Avec peu d'attaques ($k \leq 20\%$), les réplications uniformes (Figure 5.11) sont plus efficaces. Avec plus d'attaques ($k > 20\%$), les réplications adaptatives (Figure 5.12) sont plus efficaces.

Les mécanismes permettant de détecter les attaques localement (du type IDS) peuvent

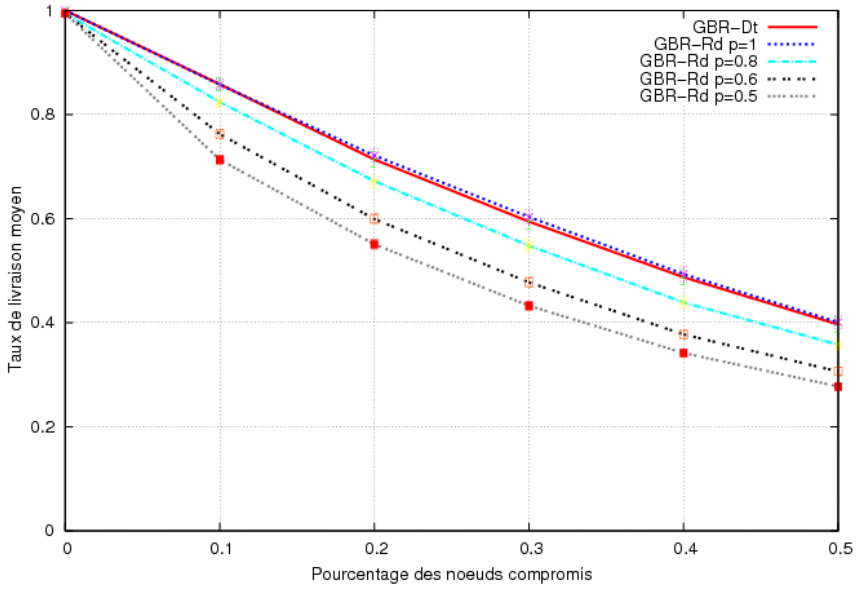


FIGURE 5.16: Taux de livraison moyen en cas d'attaque *Selective forwarding* : sans réplication.

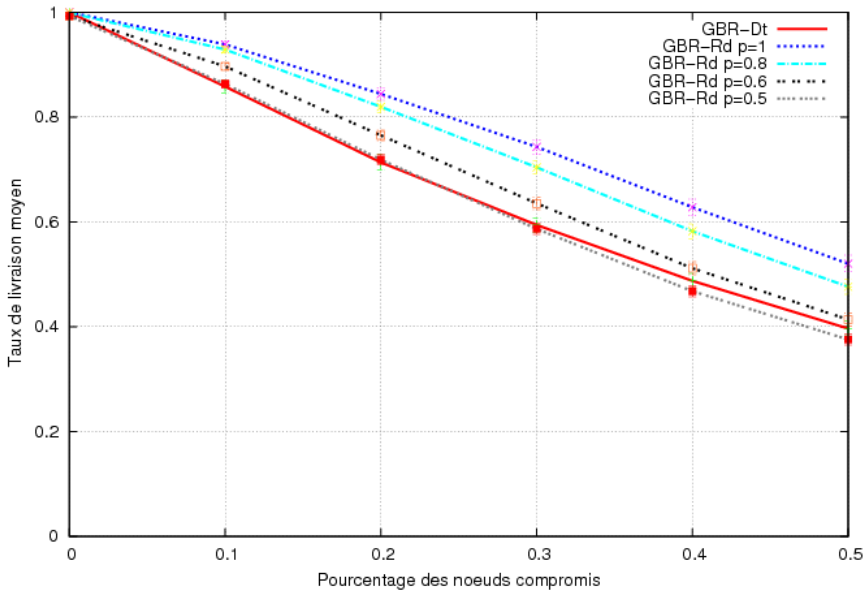


FIGURE 5.17: Taux de livraison moyen en cas d'attaque *Selective forwarding* : avec réplications uniformes.

aider à adapter les mécanismes de réplications et les biais en fonction de l'intensité des attaques afin d'éviter le surcoût de consommation d'énergie inutile. Par exemple, les mécanismes de couches MAC tel que la sur-écoute (*overhearing*, la réception de message non destiné à soi-même), ou l'analyse du trafic reçu au puits ou au niveau de chaque nœud peuvent permettre de détecter une anomalie et donc d'alerter les nœuds afin qu'ils ajustent leur taux de réplication et le biais. De cette façon, les nœuds ne répliquent pas de données en l'absence d'attaques mais seulement en cas de détection.

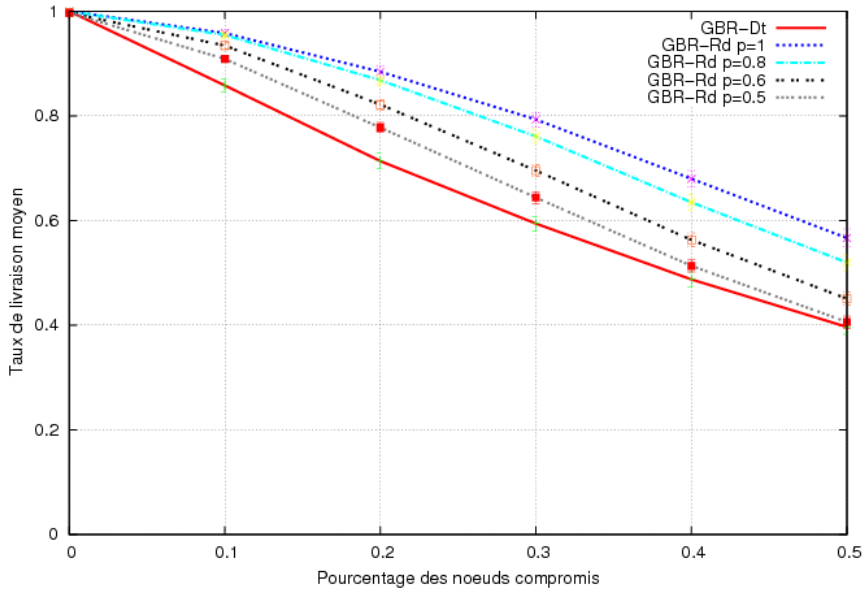


FIGURE 5.18: Taux de livraison moyen en cas d'attaque *Selective forwarding* : avec réplications adaptatives.

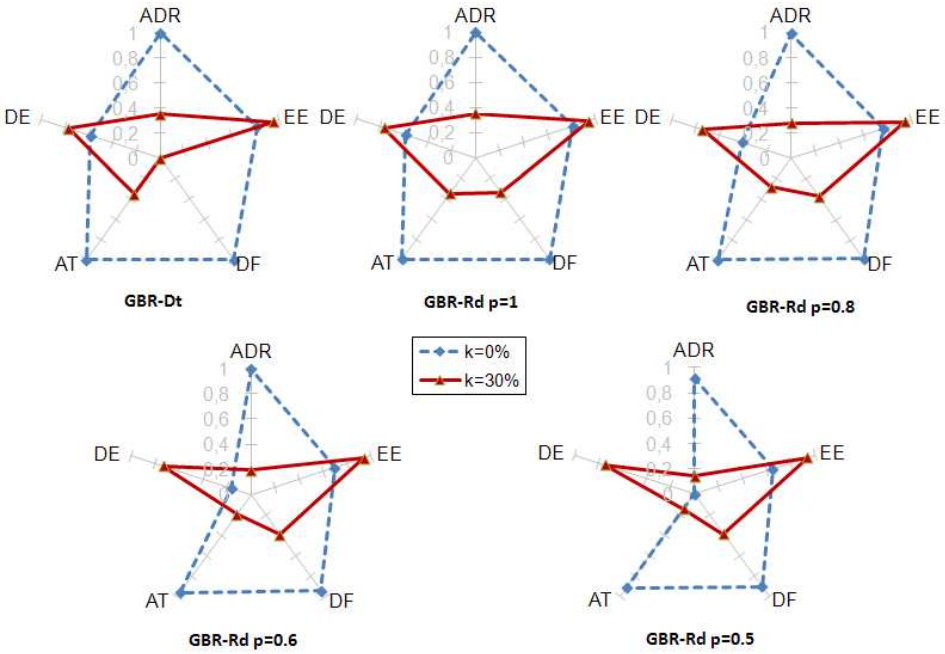


FIGURE 5.19: Surface de résilience en cas d'attaque *Sybil* : sans réplication.

5.4.2 Résultats en cas d'attaque *Sybil* combinée

L'étude d'attaque de non-retransmission des paquets, *Selective forwarding* basique, nous confirme les résultats du chapitre 4. Cependant, elle nous a permis d'observer que les variantes les moins biaisées permettent d'améliorer la connexité des sources. Mais également, les réplications uniformes sont plus efficaces que les réplications adaptatives quand il y a peu

5.4 Évaluation des performances

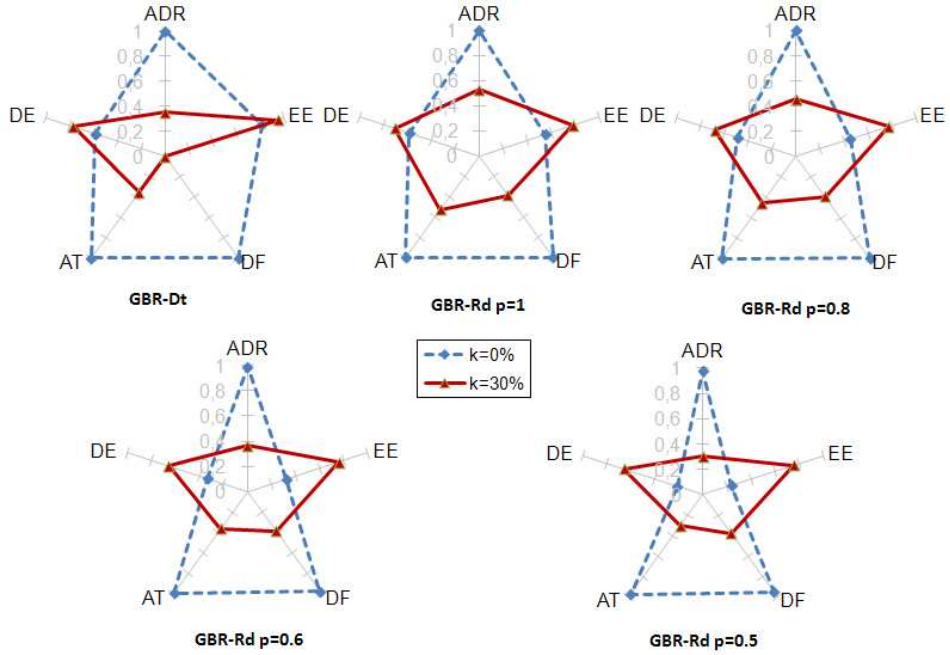


FIGURE 5.20: Surface de résilience en cas d'attaque *Sybil* : avec réplications uniformes.

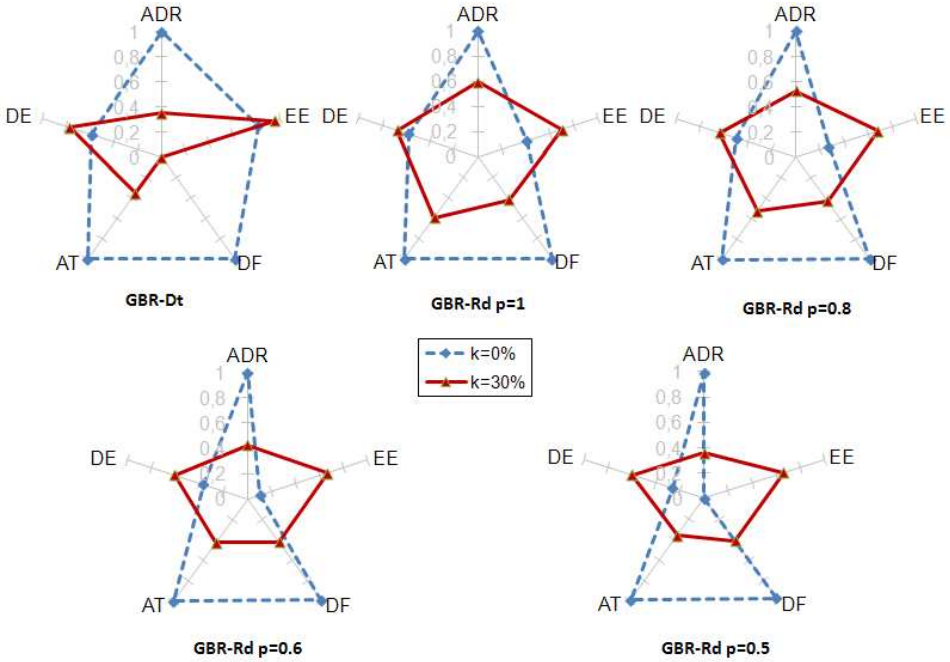


FIGURE 5.21: Surface de résilience en cas d'attaque *Sybil* : avec réplications adaptatives.

d'attaquants.

Nous combinons maintenant l'attaque *Selective forwarding* avec l'attaque *Sybil*.

Les surfaces de résilience selon les trois scénarios en cas d'attaque *Sybil* combinée (Figures

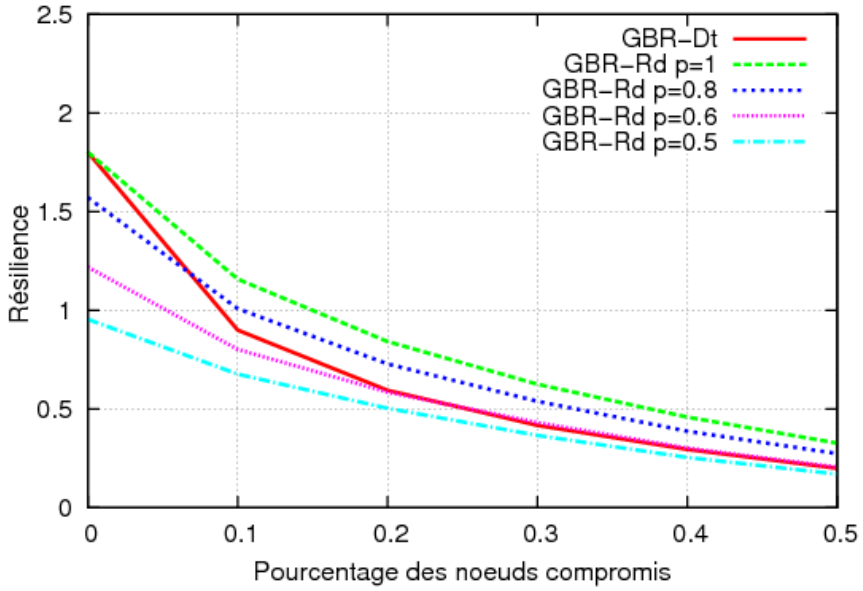


FIGURE 5.22: Résilience en cas d'attaque *Sybil* : sans réplication.

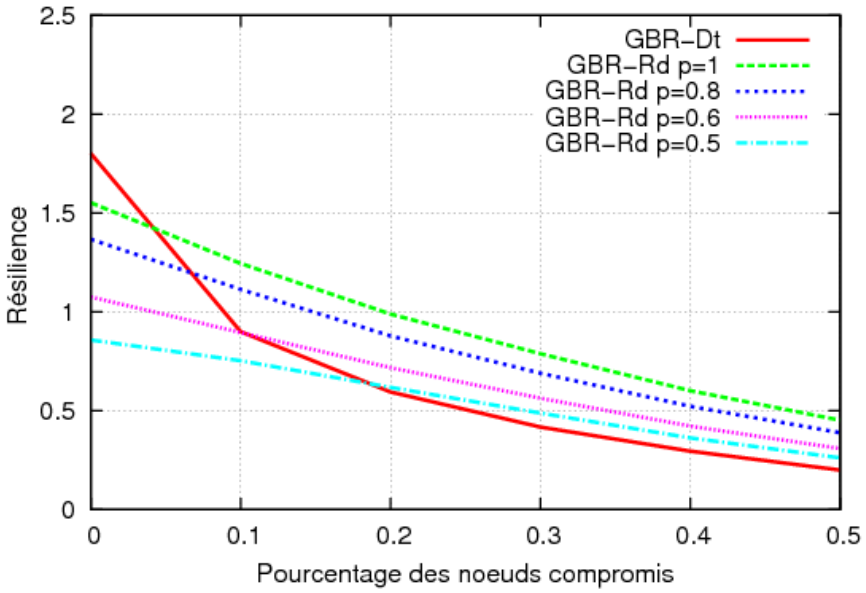


FIGURE 5.23: Résilience en cas d'attaque *Sybil* : avec réplications uniformes.

5.19- 5.21) permettent d'apercevoir immédiatement l'impact plus important sur le taux de livraison (ADR), le débit moyen (AT) et le l'équité de livraison (DF) comparé à l'attaque *Selective forwarding* basique. Cela s'explique par le fait qu'un nombre plus important de données sont attirées par des nœuds *Sybil*. Comme ces attaques ciblent particulièrement les paquets de données, ce sont les paramètres concernant la livraison de données (ADR, AT et DF) qui sont les plus affectés.

Il est à noter que, sans attaque, les surfaces de résilience n'apparaissent pas identiques entre *Sybil* combinée et *Selective forwarding*. Cela est dû à la normalisation effectuée en considérant

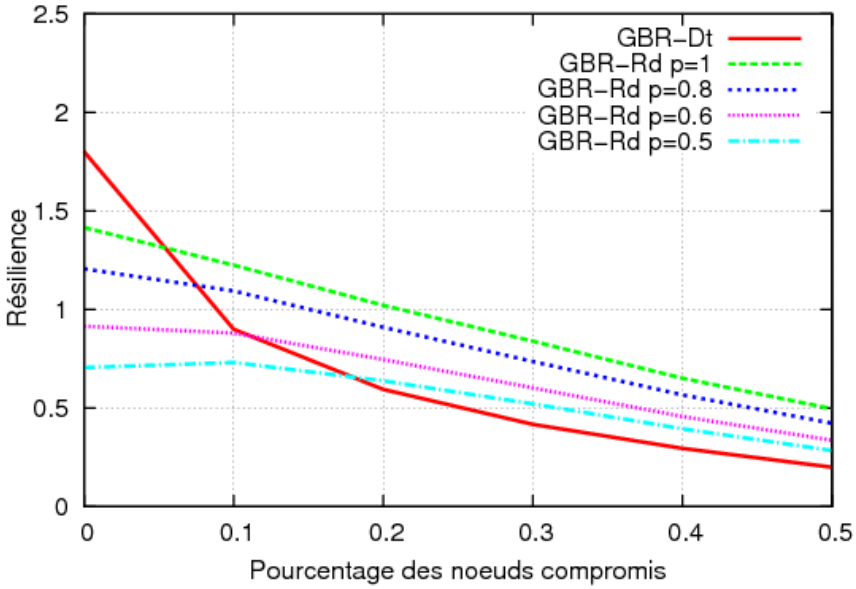


FIGURE 5.24: Résilience en cas d'attaque *Sybil* : avec réplications adaptatives.

les valeurs extrêmes (le minimum et le maximum) des protocoles que nous comparons selon chaque attaque. Comme expliqué précédemment, cette normalisation permet de faire un "zoom" sur les différences des protocoles et de pouvoir les comparer efficacement.

La diminution de la résilience est donc également plus importante avec l'attaque *Sybil* combinée (Figure 5.22) comparée à l'attaque *Selective forwarding* basique (Figure 5.10).

Cette étude nous affirme qu'en effet les mécanismes de l'attaque *Sybil* peuvent être combinés avec les attaques *Selective forwarding* permettant à un attaquant de causer des dégâts bien plus importants que si elle était produite séparément. Lorsque les nœuds malveillants créent deux identités, ils augmentent la probabilité d'être choisi comme prochain saut par leurs voisins (s'ils ont un gradient inférieur). Une fois choisis comme prochain saut, les nœuds *Sybil* reçoivent davantage de paquets à retransmettre. L'attaque *Sybil* joue ainsi le rôle d'amplificateur pour l'attaque de non retransmission des paquets.

En cas d'attaque *Sybil* combinée, les variantes du protocole GBR sont toutes plus impactées. Toutefois, la relation d'ordre des protocoles ne change pas par rapport aux résultats précédents. Les résultats de simulations confirment que les variantes aléatoires de GBR-Rd ont toutes une meilleure résilience comparée à la version classique GBR-Dt et ce, avec les deux scénarios de réplications (Figure 5.23 et 5.24), y compris GBR-Rd $p = 0.5$ au-delà de 20% d'attaquants.

Il est à noter que l'amélioration de la résilience apportée par les réplications uniformes est efficace pour $k <= 10\%$ (Figure 5.23). Avec plus d'attaques ($k > 10\%$), la réplication adaptative est plus efficace (Figure 5.24).

En résumé, l'effet de rallongement des routes est compensé par le bénéfice du trafic redondant. Notre proposition d'introduire de l'aléa et de la réplication améliore la résilience du routage par gradient en cas d'attaque *Sybil* combinée. Comme ces derniers jouent un rôle d'amplificateur, la réplication uniforme est efficace seulement quand le nombre d'attaquants

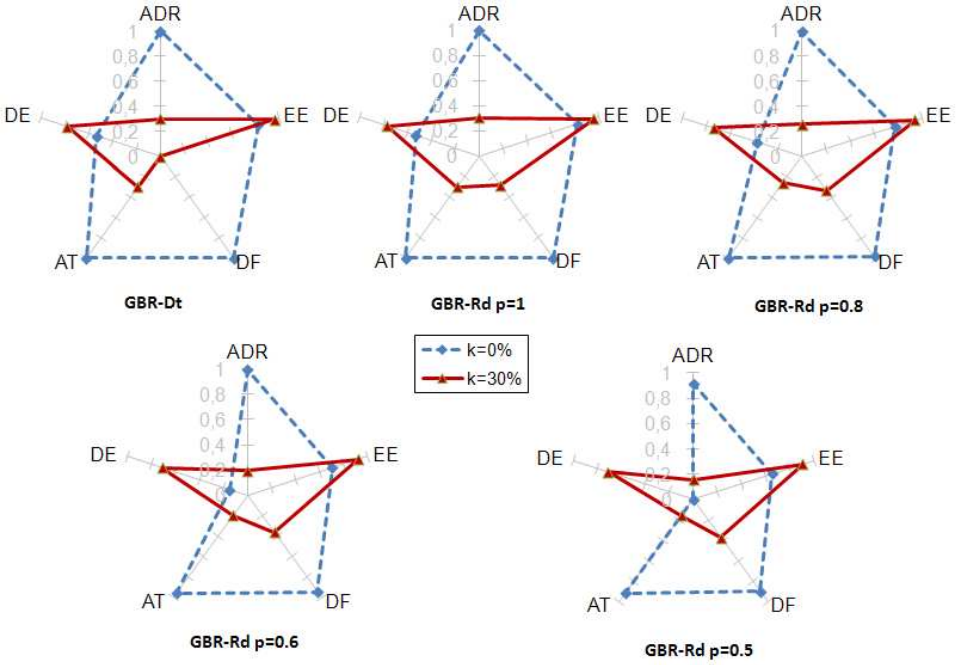


FIGURE 5.25: Surface de résilience en cas d'attaque *Wormhole* : sans réplication.

reste limité ($k <= 10\%$). La réplication adaptative, en revanche, devient efficace lorsque le nombre d'attaquants est plus important ($k > 10\%$).

5.4.3 Résultats en cas d'attaque *Wormhole* combinée

Les attaques *Selective forwarding* et *Sybil* peuvent être lancées par des nœuds compromis individuellement sans nécessairement d'entente entre eux. Pour l'attaque *Wormhole*, en revanche, les nœuds compromis établissent une connexion privée pour s'échanger des informations entre eux. C'est une attaque difficile à détecter, parce qu'il faut vérifier pour chaque nœud s'il est physiquement dans le voisinage des autres nœuds. L'attaque *Wormhole* seule peut engendrer une fausse topologie du réseau.

Nous combinons par la suite l'attaque *Wormhole* avec l'attaque *Selective forwarding*. Les surfaces de résilience des variantes de GBR selon les trois scénarios de réplications sont illustrées sur les Figures 5.25 - 5.27. Elles montrent que ce sont les mêmes paramètres que précédent (ADR, AT et DF) qui sont le plus impactés en cas d'attaque. En revanche, avec l'attaque *Wormhole* combinées elles le sont de façon plus importante.

Cela nous confirme encore l'amplification de l'effet de l'attaque *Selective forwarding*, puisque ces paramètres concernent directement la transmission des données.

La résilience de toutes les variantes du protocole GBR est donc plus impactée (Figure 5.28) en cas d'attaque *Wormhole* combinée que celle en cas d'attaque *Sybil* combinée (Figure 5.22) et *Selective forwarding* (Figure 5.10).

Soit une paire de nœuds *Wormhole* $(w1, w2)$. Une distance de deux sauts entre $w1$ et $w2$ permet déjà de créer une topologie déformée. Plus les deux nœuds $(w1, w2)$ sont loin l'un

5.4 Évaluation des performances

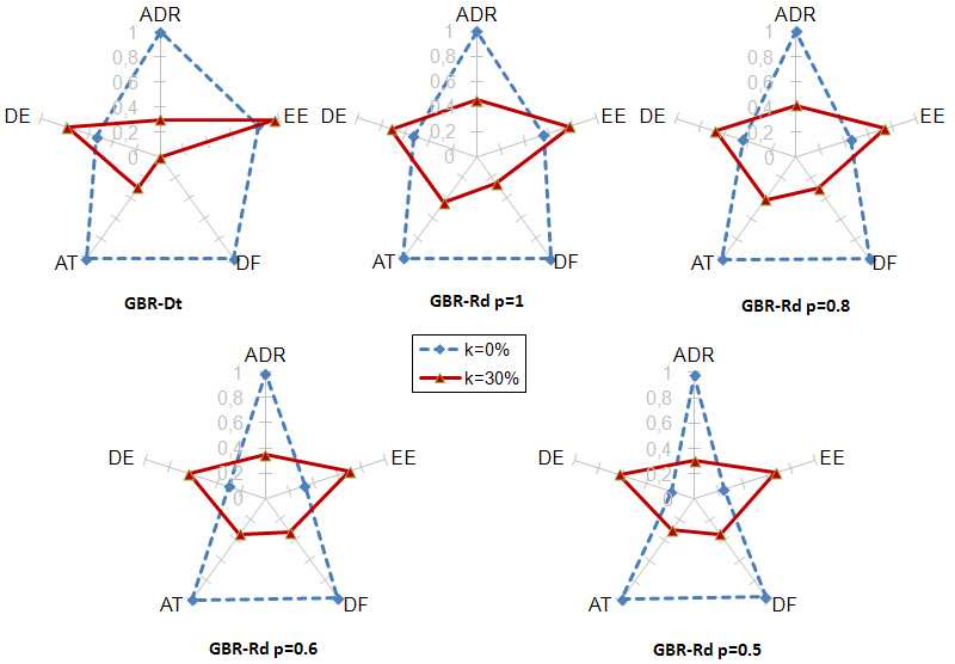


FIGURE 5.26: Surface de résilience en cas d'attaque *Wormhole* : avec réplications uniformes.

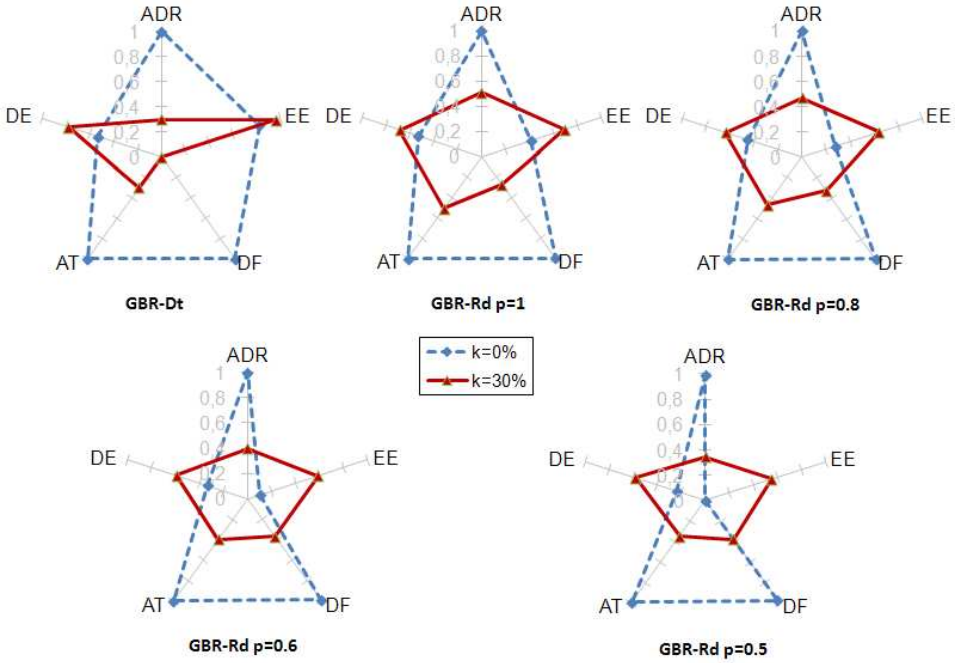


FIGURE 5.27: Surface de résilience en cas d'attaque *Wormhole* : avec réplications adaptatives.

de l'autre, plus les dégâts peuvent être importants. De plus, si l'un des nœuds *Wormhole* est proche du puits, l'autre nœud *Wormhole* peut attirer une grande partie du trafic de ces voisins grâce aux informations transmises par son allié. En effet, si $w1$ est voisin avec

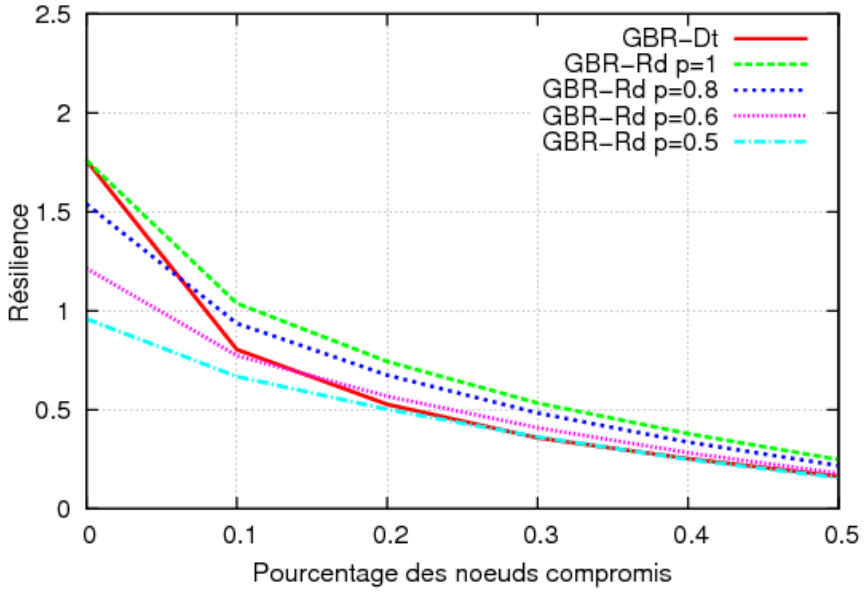


FIGURE 5.28: Résilience en cas d'attaque *Wormhole* : sans réplication.

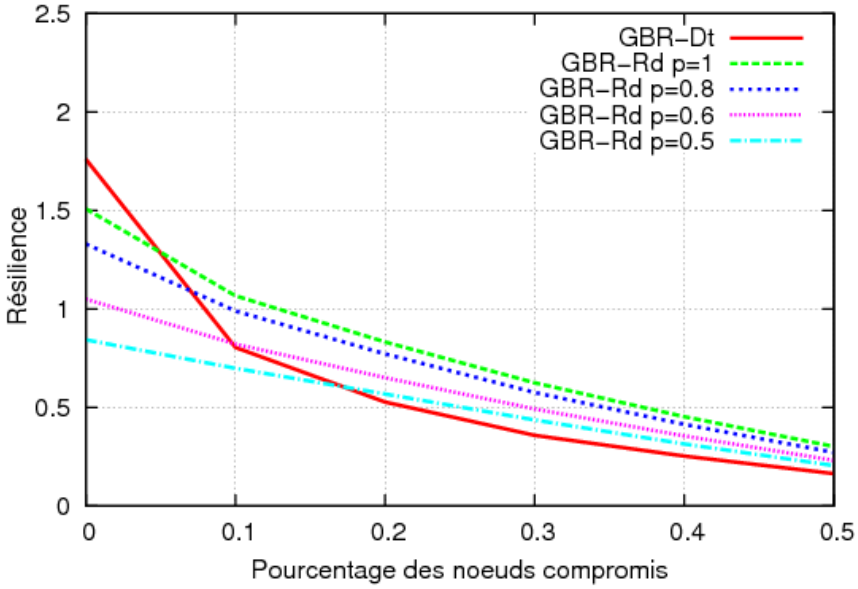


FIGURE 5.29: Résilience en cas d'attaque *Wormhole* : avec réplications uniformes.

le puits, $w2$ obtient un meilleur gradient comparé à ses voisins qui sont loin du puits. Les nœuds *Wormhole* permettent ainsi d'attirer efficacement le trafic vers $w2$. Les deux nœuds peuvent ensuite éliminer tous les paquets reçus.

Là encore, toutes les variantes aléatoires GBR-Rd ont une meilleure résilience que GBR-Dt en cas de réplications des paquets (Figure 5.29 et 5.30). Comme pour les attaques précédentes, notre proposition d'introduire de l'aléa et des réplications améliorent encore la résilience contre les attaques *Wormhole* combinées. De la même façon qu'en cas d'attaque *Sybil* combinée, l'amélioration de la résilience apportée par les réplications uniformes est efficace

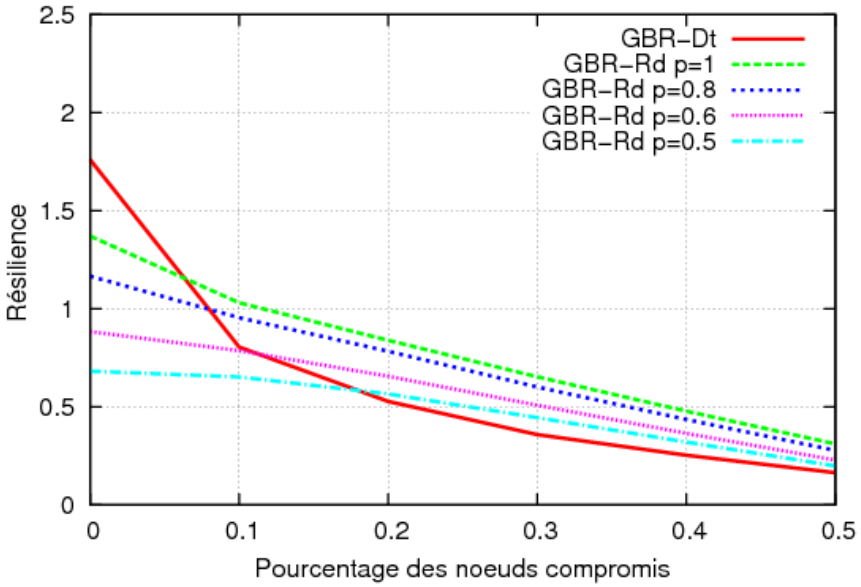


FIGURE 5.30: Résilience en cas d'attaque *Wormhole* : avec réplications adaptatives.

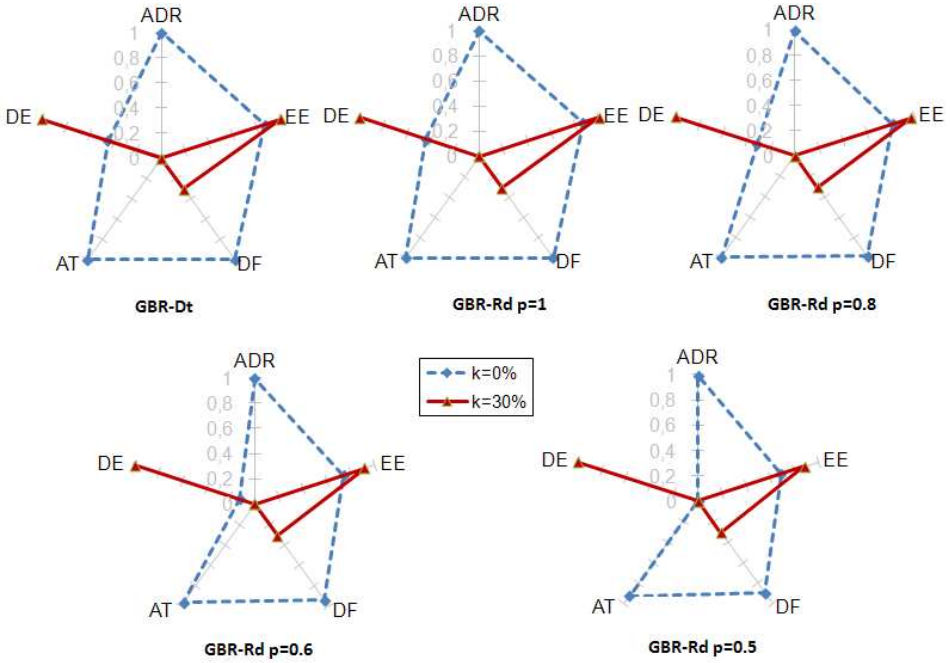


FIGURE 5.31: Surface de résilience en cas d'attaque *Sinkhole* : sans réplication.

pour $k <= 10\%$ (Figure 5.29). Avec plus d'attaques ($k > 10\%$), la réplication adaptative devient efficace (Figure 5.30).

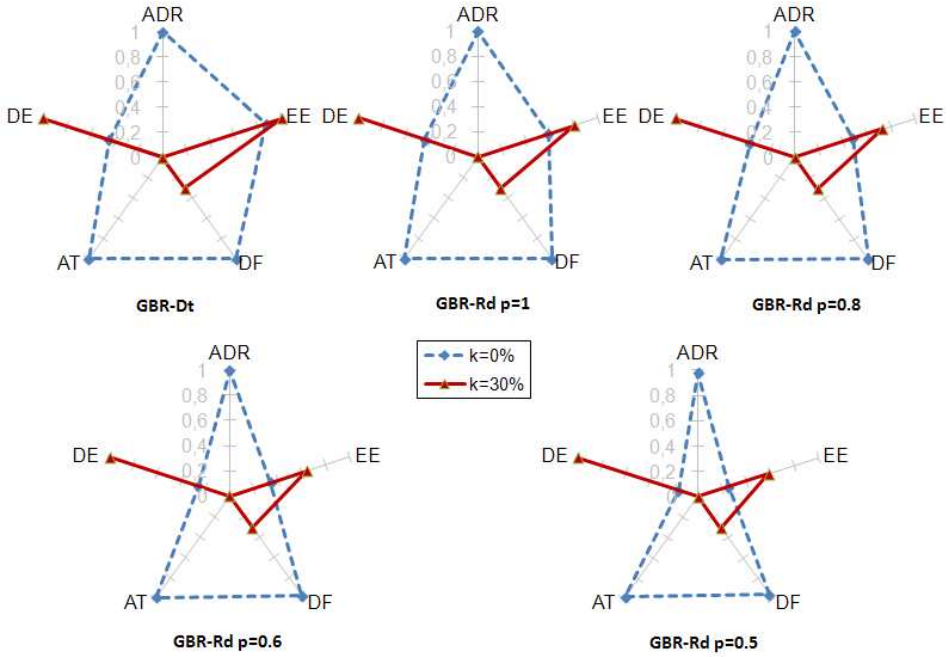


FIGURE 5.32: Surface de résilience en cas d'attaque *Sinkhole* : avec réplications uniformes.

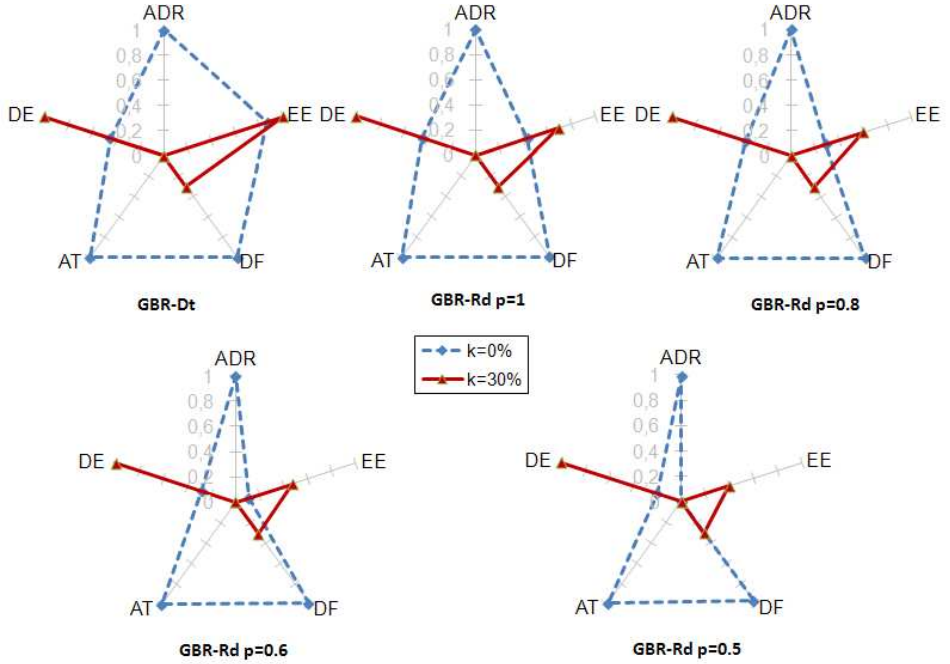


FIGURE 5.33: Surface de résilience en cas d'attaque *Sinkhole* : avec réplications adaptatives.

5.4.4 Résultats en cas d'attaque *Sinkhole* combinée

Les attaques précédentes sont toutes basées sur l'hypothèse que les nœuds compromis sont uniformément distribués dans le réseau. Cependant, un adversaire connaissant la position du

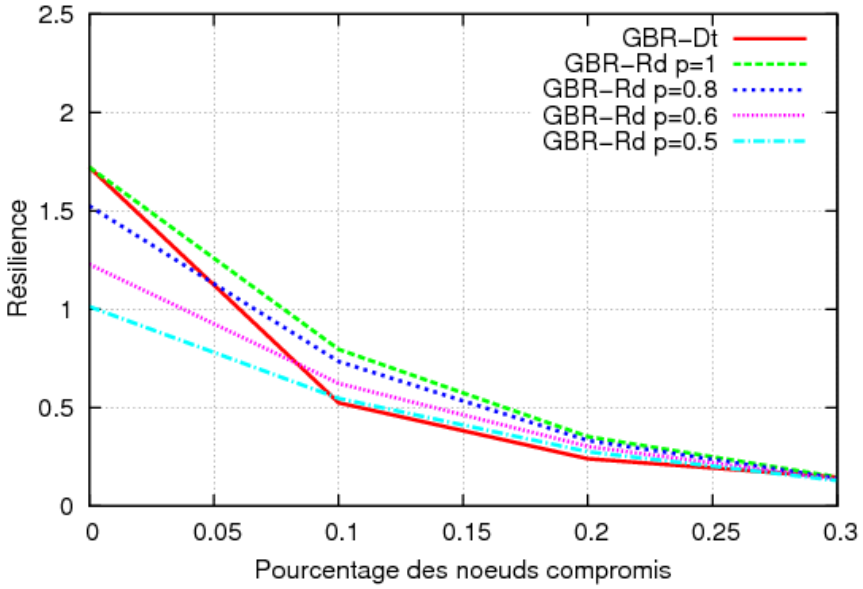


FIGURE 5.34: Résilience en cas d'attaque *Sinkhole* : sans réplication.

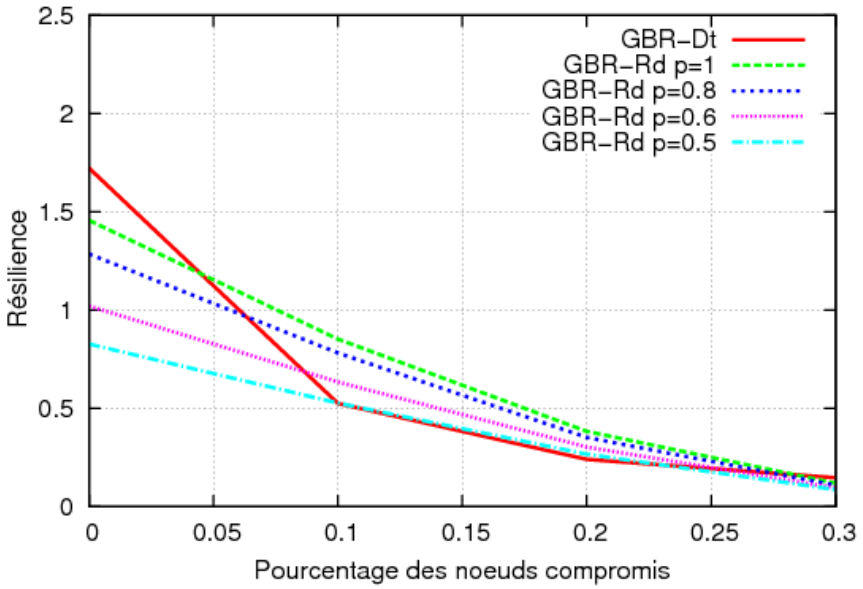


FIGURE 5.35: Résilience en cas d'attaque *Sinkhole* : avec réplications uniformes.

puits va tenter de compromettre les nœuds proches du puits pour une meilleure efficacité. Un nœud voisin du puits a plus de trafic à acheminer que les nœuds à la bordure du réseau. En profitant de cette propriété des réseaux *convergecast*, les adversaires peuvent produire des attaques *Sinkhole* bien plus efficaces contre le réseau.

Les surfaces de résilience ((Figure 5.31-5.33) et l'évaluation quantitative de la résilience ((Figure 5.35-5.36) des protocoles GBR selon les trois scénarios nous confirment que l'impact d'attaques *Sinkhole* combinées est le plus dévastateur comparé à toutes les autres attaques étudiées précédemment.

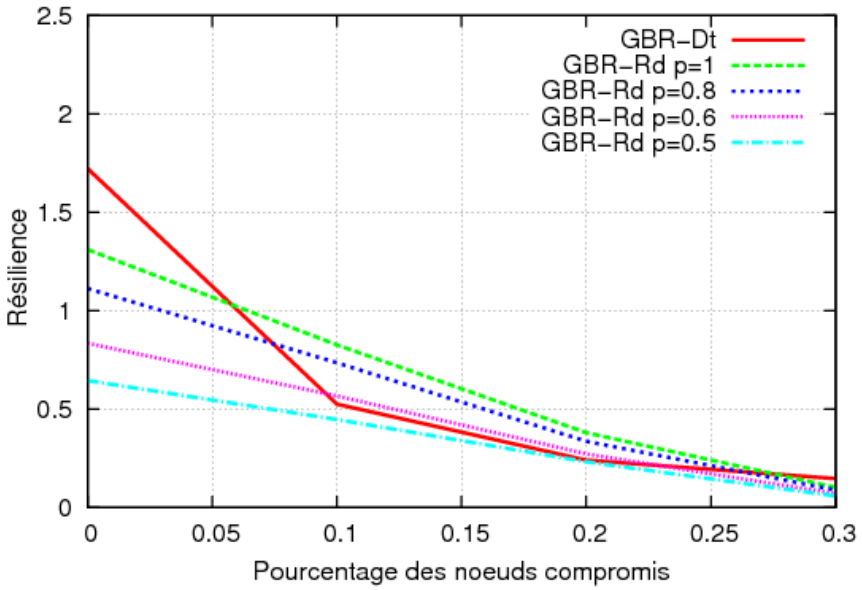


FIGURE 5.36: Résilience en cas d'attaque *Sinkhole* : avec réplications adaptatives.

Lorsqu'une grande partie des nœuds autour du puits est compromise, ils attirent naturellement la majorité du trafic du réseau. Nous varions le pourcentage d'attaquants jusqu'au $k = 30\%$ parce qu'au-delà de cette valeur la totalité des voisins du puits est compromise. En effet, nous observons sur les Figures 5.31 - 5.33 que le taux de livraison moyen (ADR) et le débit moyen (AT) sont proches de zéro. Dans ce cas, le puits reçoit des paquets uniquement de nœuds compromis et aucun paquet n'est reçu de nœuds légitimes.

Nous observons une convergence des courbes de résilience de tous les protocoles (avec ou sans réplication) avec l'augmentation des attaques. Cette convergence est liée au phénomène contre-intuitif que nous avons précisé dans le chapitre 3, où les deux paramètres (EE) et (DE) sur cinq s'améliorent avec les attaques. Comme nous l'avons expliqué précédemment, cette amélioration de l'efficacité énergétique (EE) est due à la diminution du trafic du réseau à cause des paquets perdus. L'amélioration de l'efficacité de délai (DE) est due aux paquets reçus au puits provenant principalement des nœuds sources proches. À cause de ce phénomène, la surface de résilience n'est pas nulle, même en cas d'attaques intenses.

Dans ce cas extrême, la réplication du trafic n'apporte aucun bénéfice, et elle engendre seulement du surcoût en consommation d'énergie (EE) que nous observons sur les Figures 5.31 - 5.33.

5.5 Conclusion

Dans ce chapitre, nous avons évalué les performances des cinq variantes de GBR (classique et aléatoires) avec différents biais p en présence de plusieurs attaques combinées. Nous les avons évalués selon trois scénarios (sans et avec deux types de réplications des paquets).

Les résultats de simulations nous confirment que, sans attaque, les variantes sans réplication et basées sur le plus court chemin GBR-Dt et GBR-Rd $p = 1$ sont les plus efficaces. En

cas d'attaque *Selective forwarding*, GBR-Rd $p = 1$ est plus résilient que GBR-Dt. Cependant GBR-Rd $p = \{0.8, 0.6, 0.5\}$ améliore la connectivité des sources (aucune source n'est déconnectée du puits). Avec peu d'attaques ($k <= 20\%$), les réplications uniformes sont plus efficaces. Tandis qu'en cas d'attaques plus intenses ($k > 20\%$), les réplications adaptatives sont plus efficaces.

Les attaques *Sybil* combinées avec *Selective forwarding* amplifient l'impact des attaques parce que les fausses identités permettent de donner l'illusion au voisinage qu'il y a plusieurs nœuds alors en réalité il n'y a qu'un nœud physique. L'effet de l'attaque *Wormhole* est encore plus important que *Sybil* quand elle est combinée avec *Selective forwarding*. Grâce à l'entente entre les deux nœuds *Wormhole*, quand ils sont bien placés (l'un des deux nœuds placé près du puits et l'autre physiquement loin) permettent d'attirer efficacement le trafic de leurs voisins. Pour ces attaques, les mécanismes aléatoires avec réplications des paquets permettent d'améliorer la résilience. Cependant, en cas d'attaque *Sinkhole*, les nœuds compromis autour du puits permettent d'absorber la majorité du trafic du réseau. Dans ce cas extrême, les mécanismes résilients ne permettent pas de lutter contre les nœuds compromis car le puits est complètement déconnecté du reste des sources.

En cas d'absence d'attaques, les réplications uniformes ou adaptatives n'apportent aucun gain. Elles ne sont que du gaspillage d'énergie, sans contrepartie. Avec peu d'attaques, les réplications uniformes sont plus efficaces. Dans le cas d'attaques intenses, le coût des réplications adaptatives peut être justifié en terme de gain en taux de livraison et d'équité mais ce coût est également fortement réduit en raison des pertes qui réduisent le trafic dans le réseau. Cela nous montre que les mécanismes de détection locale d'attaques (du type IDS) peuvent être utiles pour ajuster les réplications et les biais afin d'éviter de gaspiller l'énergie du réseau.

Les études menées jusqu'à ce chapitre sont empiriques, ce qui nous permet de justifier nos premières observations par des simulations. Cependant, une justification plus solide du point de vue théorique est nécessaire. Nous proposons dans le chapitre suivant une étude analytique de la marche aléatoire biaisée en cas d'attaque.

Étude théorique de la résilience des marches aléatoires biaisées 6

Sommaire

6.1 Introduction 103
6.2 Définitions, notations et hypothèses 103
6.2.1 Marche aléatoire biaisée 104
6.2.2 Hypothèses et paramètres de calcul 105
6.2.3 Longueur moyenne des chemins 106
6.3 Marche aléatoire biaisée sans réplication 107
6.3.1 Succès de livraison 108
6.3.2 Consommation d'énergie 110
6.4 Marche aléatoire biaisée avec réplications 113
6.4.1 Réplications uniformes 113
6.4.2 Réplications adaptatives 117
6.5 Évaluation de la résilience 119
6.5.1 Marche aléatoire sans réplication 119
6.5.2 Marche aléatoire avec réplications uniformes 120
6.5.3 Marche aléatoire avec réplications adaptatives 122
6.6 Conclusion 124

6.1 Introduction

Les mécanismes de routage bénéfiques à la résilience des protocoles destinés aux réseaux de capteurs ont été proposés dans le chapitre 4 et étudiés par simulations dans le chapitre 5. Ces mécanismes consistaient en trois éléments : (i) introduire de comportements aléatoires dans le routage (ii) limiter la longueur des routes (iii) répliquer les paquets de données. Les résultats de simulations nous ont confirmé que ces mécanismes améliorent la résilience des protocoles de routage classiques.

Introduire de l'aléa dans un protocole de routage, c'est ajouter de l'entropie [1] au système. Nous constatons deux intérêts majeurs pour la résilience : (a) la diversité des routes créées entre une source et le puits est améliorée (b) l'incertitude est augmentée pour les adversaires.

Grâce à (a), les sources peuvent envoyer chaque paquet sur une route construite aléatoirement, créant des routes potentiellement différentes à chaque envoi de paquets. Ceci améliore la connectivité des chemins entre les sources et le puits. Pour exploiter cette diversité des routes créées, nous répliquons x fois chaque paquet, afin d'améliorer le succès de transmission de chaque paquet de données. Tandis que (b) permet de rendre les protocoles non prévisibles grâce aux comportements aléatoires. Avec un choix aléatoire des routes, un attaquant ne peut plus agir sur le choix des routes. Cela peut donc le dissuader de tenter d'attirer plus de trafic en s'attribuant les meilleures valeurs (meilleur délai, meilleur débit, distance minimum, etc.).

L'introduction de comportements aléatoires améliore donc la résilience, mais jusqu'à quel seuil d'entropie est-elle bénéfique à la résilience ? Notre but est double : répondre à cette question et confirmer les résultats d'études par simulations du chapitre 5. Nous proposons dans ce chapitre une étude analytique en nous basant sur la théorie de la marche aléatoire. L'entropie maximale correspond donc à la marche aléatoire sans biais, puisqu'aucune information d'état (distance géographique, gradient, etc.) n'est disponible et la probabilité de choisir le prochain saut est uniforme. Au contraire, le biais permet de diriger la marche aléatoire pour arriver plus rapidement à la destination grâce à l'information d'état.

En résumé, nous évaluons tout d'abord, la marche aléatoire biaisée en cas d'attaques selon les deux métriques les plus importantes (succès de livraison et consommation d'énergie). Ensuite, nous y introduisons deux types de réplications (uniformes et adaptatives) que nous avons introduits au chapitre 5. Enfin, nous les évaluons selon notre métrique de résilience présentée au chapitre 3 pour avoir une vision d'ensemble.

6.2 Définitions, notations et hypothèses

Dans cette partie, nous introduisons formellement les notions de base de la marche aléatoire biaisée et les hypothèses que nous considérons.

1. L'entropie de Shannon correspond à la quantité d'information contenue ou délivrée par une source d'information (ou incertitude sur ce que la source émet) [25]

6.2.1 Marche aléatoire biaisée

Dans cette étude nous considérons un tore (*torus lattice*) $\Im$ (Fig. 6.1). Un tore est une grille carrée de taille $N \times N$, dont les extrémités opposées sont connectées et où chaque nœud a le même degré $d = 4$. Nous avons ainsi une structure régulière qui permet d'incorporer des éléments théoriques pour modéliser la marche aléatoire biaisée tout en s'adaptant aux propriétés de certains réseaux de capteurs connus [2].

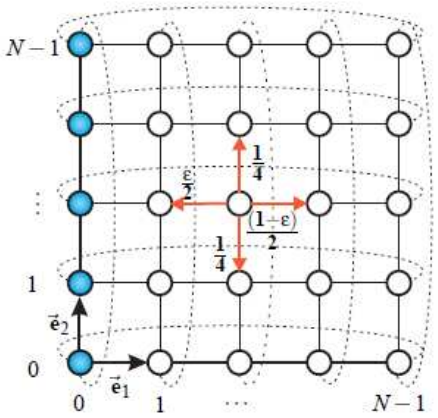


FIGURE 6.1: Tore $\Im$ [2]

Un nœud est représenté par un vecteur $\vec{r} \in \Im$ ayant les coordonnées (r_1, r_2), où r_1 et r_2 sont des entiers tels que $0 \leq r_1, r_2 \leq N-1$. Nous définissons la distribution de la probabilité de transition d'un nœud par $\sum_{\vec{s} \in \Im} \wp(\vec{r}, \vec{s}) = 1$, où $\wp(\vec{r}, \vec{s})$ est la probabilité de choisir le nœud $\vec{s}$ quand le paquet atteint le nœud $\vec{r}$. La séquence aléatoire des nœuds choisis de cette façon est une marche aléatoire sur $\Im$ et $\wp(\vec{r}, \vec{s})$ est la fonction de transition de la marche aléatoire [115]. Nous considérons deux types de nœuds dans le réseau : le puits pour collecter des données à un endroit donné du réseau et le nœud simple pour capturer les informations du monde physique et les transmettre au puits. Nous appelons le point de collecte C, l'ensemble des N puits à la bordure gauche de la grille $\vec{s_j} = (0, j)$, où $j = 0, 1, ..., N-1$. Les autres nœuds forment l'ensemble des nœuds capteurs S du réseau. Les données envoyées par l'ensemble des nœuds capteurs S sont collectées par l'ensemble des puits C sans faire correspondre un capteur à un puits spécifique. Ainsi, l'ensemble des puits C forme un seul point de collecte et c'est équivalent de considérer un seul puits.

Nous insistons sur cette hypothèse afin de correspondre le plus aux paramètres de nos simulations dans le chapitre 5. De plus, dans le contexte des réseaux de capteurs, les données des capteurs sont collectées souvent vers un ou peu de puits.

Dans une marche aléatoire standard sans biais, un nœud capteur $\vec{r}$ choisit un de ses voisins avec la probabilité $1/d = 1/4$ pour le prochain saut. Pour introduire le biais dans la marche aléatoire, un nœud capteur $\vec{r}$ choisit son voisin de haut et de bas avec la probabilité $1/4$, son voisin de gauche avec la probabilité $1/2 \times \varepsilon$ et son voisin de droite avec la probabilité $1/2 \times (1-\varepsilon)$, où $1/2 \leq \varepsilon \leq 1$. Ainsi, dépendant de la valeur de ε, la marche aléatoire est plus ou moins biaisée vers le point de collecte. Pour donner une interprétation au paramètre

ε, on introduit la notion d'*entropie* associée à la probabilité de transition d'un nœud. Plus il y a de l'information d'état moins il y a d'entropie. L'entropie est donnée par l'équation suivante [2] :

$$H_\varepsilon = \frac{3}{2} - \frac{1}{2}(\varepsilon log_2(\varepsilon) + (1-\varepsilon)log_2(1-\varepsilon)) \tag{6.1}$$

L'entropie permet de déterminer quelle quantité d'incertitude est introduite dans le choix du prochain saut. Pour $\varepsilon = 1/2$, H_ε est maximale, cela correspond donc à une marche aléatoire standard sans biais où l'incertitude est maximale : le nœud choisit un de ces voisins comme prochain saut avec une probabilité uniforme. Sinon, l'incertitude est moins importante et une direction est favorisée pour la marche aléatoire. La notion de *mean data gathering delay*, la longueur moyenne des chemins entre les nœuds capteurs et le point de collecte, est introduit par [2]. C'est le nombre de sauts moyen pris par un paquet de données initié au nœud capteur $\vec{r}$ pour atteindre le point de collecte C. Il s'agit de l'espérance de la longueur des chemins en fonction du biais de la marche aléatoire. La longueur moyenne des chemins est donnée par la formule suivante :

$$E(D_\varepsilon(\vec{r})) = \frac{2}{2\varepsilon - 1}(r_1 - N\frac{1-(\frac{\varepsilon}{1-\varepsilon})^{r_1}}{1-(\frac{\varepsilon}{1-\varepsilon})^N}), \text{ pour tout } \varepsilon \neq \frac{1}{2}, 1. \tag{6.2}$$

Pour $\varepsilon = \frac{1}{2}$, $E(D_{\frac{1}{2}}(\vec{r})) = 2r_1(N - r_1)$ et pour $\varepsilon = 1$, $E(D_1(\vec{r})) = 2r_1$.

6.2.2 Hypothèses et paramètres de calcul

Nos calculs se basent sur le modèle théorique présenté précédemment dans la section 6.2.1 en considérant les paramètres suivants. Nous prenons un tore en grille $\Im$ de taille $N \times N = 5 \times 5$, où $|S| = 20$ nœuds capteurs et $|C| = 5$ nœuds puits formant un seul point de collecte. Cette structure permet obtenir 4 niveaux, où les nœuds sont à une distance de 1 à 4 sauts du point de collecte. Ces dimensions sont celles qui correspondent le plus des paramètres des simulations effectuées précédemment dans le chapitre 5.

Dans les simulations, nous avions une zone de déploiement carrée $N \times N$, et un diamètre du réseau de 7 sauts. En plaçant un puits au centre, nous avions la distance maximum des nœuds égale à 4 sauts du puits. Cependant, nous avons une divergence sur deux points : (i) le nombre de puits (ii) le nombre de nœuds par niveau.

Concernant (i), nous avions un seul puits dans les simulations, tandis qu'ici nous avons $|C| = 5$ puits. Toutefois, les cinq puits forment un seul point de collecte, puisque les paquets ne sont pas adressés à un puits en particulier. Dès qu'ils arrivent à un des puits, nous considérons qu'il a atteint la destination avec succès. Comme les liens entre les cinq puits ne sont pas utilisés dans les calculs, c'est équivalent à considérer les cinq puits en un seul (voir la Figure 6.2).

Concernant (ii), nous avons dans ce modèle le même nombre de nœuds par niveau (5 nœuds par niveau, représentant 25% des nœuds du réseau). Cette configuration ne se retrouve pas dans les simulations. Dans les simulations du chapitre 5, le déploiement aléatoire des nœuds selon une distribution uniforme engendre naturellement un nombre de nœuds qui ne sera

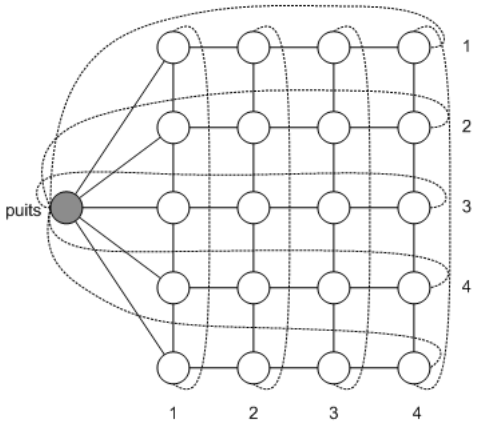


FIGURE 6.2: Tore avec un seul puits.

pas identique à chaque niveau. Le nombre de nœuds du niveau 1 est moins important (en moyenne $14,71\%$) que celui du niveau 2 ($30,77\%$). De la même façon, le nombre des nœuds de niveau 2 est moins important que celui du niveau 3 ($41,81\%$). En revanche, ce n'est pas le cas pour le dernier niveau ($12,71\%$) à cause de l'effet de bordure.

Les valeurs du biais $\varepsilon = \{1, 0.9, 0.8, 0.7, 0.6, 0.5\}$ sont identiques au modèle théorique [2]. $\varepsilon = 1$ correspond à l'entropie minimale et au biais maximal et $\varepsilon = 0.5$ correspond à l'entropie maximale et au biais minimal. Les auteurs de [2] ont choisi ces valeurs d'entropie à cause de la symétrie sur l'axe $X = N/2$. Il suffit donc de considérer les valeurs $1/2 \leq \varepsilon \leq 1$. Notez que nous avons dans ce modèle, le biais maximum qui permet de progresser vers le puits seulement une fois sur deux en moyenne. Cela ne permet pas de représenter les protocoles étudiés par simulations tels que GBR-Rd $p = 1$ du chapitre 5, où le routage se base sur le critère du plus court chemin. Dans ce modèle, la marche aléatoire la plus biaisée correspond au protocole le moins biaisé GBR-Rd $p = 0.5$ du chapitre 5 et BRWR du chapitre 2.

Nous considérons les valeurs $P_c = \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$ pour la probabilité de compromission des nœuds. Dans les simulations du chapitre 5, nous avions considéré des valeurs identiques, mais un pourcentage de compromission des nœuds plutôt qu'une probabilité de compromission. La probabilité de compromission d'un nœud est indépendante de celle des autres nœuds et permet d'éviter une explosion combinatoire sans impacter la cohérence de notre modèle théorique.

Enfin, nous considérons les attaques de non retransmission des paquets de données (*Selective forwarding*) équivalentes aux attaques présentées à la section 5.3.1 du chapitre 5.

6.2.3 Longueur moyenne des chemins

Les courbes de la Figure 6.3 sont obtenues grâce à l'équation (6.2) proposée dans [2] avec les paramètres de calcul présentés dans la section précédente 6.2.2. Ces courbes représentent la longueur moyenne des chemins (en nombre de sauts) en fonction de l'entropie sans attaque et elles correspondent aux résultats présentés dans [2]. Les calculs suivants effectués en cas d'attaques sont basés sur ces résultats.

Nous observons l'impact du biais sur la longueur moyenne des chemins. Quand la marche aléatoire est la plus dirigée vers la destination (le biais est maximum et $\varepsilon = 1$), la performance en matière de longueur moyenne des chemins est la meilleure. La marche est biaisée dans le sens de droite à gauche sur le tore. Nous avons donc une droite pour la marche aléatoire la plus biaisée.

Avec un biais minimum $\varepsilon = 0.5$ (le prochain saut est choisi avec une probabilité uniforme), la performance en termes de longueur moyenne des chemins est la plus médiocre : elle varie entre 8 à 12 sauts. Nous pouvons remarquer que la courbe présente une symétrie : la longueur des moyennes des chemins est égale à 8 sauts pour les nœuds à la position r_1 et r_4 et 12 sauts pour les nœuds aux positions r_2 et r_3.

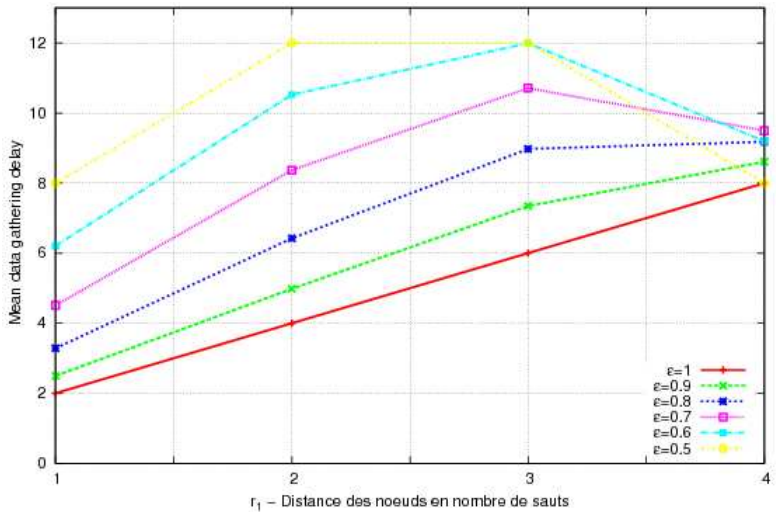


FIGURE 6.3: Longueur moyenne des chemins en fonction du biais ε pour chaque position r_1 des nœuds.

Ce résultat est particulièrement lié à la caractéristique du tore. La marche aléatoire sans biais a toute autant de chance d'atteindre le point de collecte dans le bon sens que dans le sens opposé. Cependant, elle reste la moins performante parce qu'elle a la possibilité de faire des boucles en repassant par des nœuds déjà visités. La marche aléatoire la plus biaisée est donc la plus performante parce qu'elle est dirigée vers le puits. Toutefois, avec la marche aléatoire la plus biaisée, un paquet se trouvant à un niveau 4 va faire un chemin long dans le bon sens, alors qu'il se trouve à un saut du puits dans l'autre sens. Ces phénomènes liés au tore influencent également les résultats en cas d'attaques.

6.3 Marche aléatoire biaisée sans réplication

Nous évaluons la marche aléatoire biaisée en cas d'attaques *Selective forwarding* en considérant l'envoi d'un paquet par nœud sans réplication. Nous étudions en détail le succès de livraison et la consommation d'énergie dans cette section avant d'évaluer notre métrique de résilience.

6.3.1 Succès de livraison

Pour la marche aléatoire biaisée, le taux de livraison moyen du réseau est présenté par la probabilité du succès de livraison moyen du réseau.

Nous calculons d'abord la probabilité de succès de livraison d'un nœud $\vec{r} \in S$ et ensuite nous calculons son espérance pour obtenir la valeur moyenne du réseau.

$$P_{\vec{r}}(P_c, \varepsilon) = (1 - P_c)^{E(D_\varepsilon(\vec{r}))-1} \tag{6.3}$$

où $E(D_\varepsilon(\vec{r}))$ (6.2) est la longueur moyenne des chemins pour la marche aléatoire biaisée [2].

Pour un nœud source $\vec{r}$, la probabilité de succès de livraison d'un paquet va dépendre du nombre de nœuds relais que le paquet rencontre avant d'atteindre le point de collecte. Si le paquet, partant du nœud $\vec{r}$, fait $E(D_\varepsilon(\vec{r}))$ sauts en moyenne, alors il traverse $E(D_\varepsilon(\vec{r})) - 1$ nœuds relais avant d'atteindre sa destination. Si $1-P_c$ est la probabilité qu'un nœud soit sain (non compromis), alors la probabilité qu'un paquet ne rencontre aucun nœud compromis en chemin est $1 - P_c$ à la puissance $E(D_\varepsilon(\vec{r})) - 1$.

Les courbes présentées sur les Figures 6.4-6.6 sont obtenues grâce à l'équation (6.3). Ces courbes représentent la probabilité de livraison d'un nœud r à la position r_1 en fonction de la probabilité de compromission des nœuds P_c.

La Figure 6.4(a) nous donne cette probabilité pour la marche aléatoire la plus biaisée et la Figure 6.6(b) nous donne cette probabilité pour la marche aléatoire la moins biaisée. Comme attendu, la probabilité de livraison diminue avec l'intensité des attaques, puisque les paquets de données sont perdus s'ils rencontrent un nœud compromis sur leur chemin. Nous observons, pour la plus biaisée (Figure 6.4(a)), une différence importante en fonction de la position des nœuds. Entre $r_1 = 1$ et $r_1 = 4$, avec 20% de nœuds attaquants, nous avons déjà un écart de 60% de succès de livraison. Ceci s'explique par le fait que plus le chemin est long, plus la probabilité qu'un paquet tombe sur un nœud compromis augmente. Et pour la marche aléatoire la plus biaisée, comme l'a montré la Figure 6.3, la distance au puits est directement proportionnelle avec la longueur moyenne du chemin.

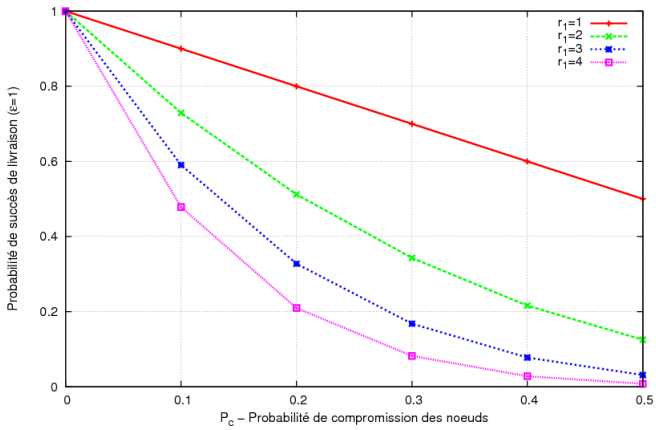
Pour la moins biaisée (Figure 6.6(b)), nous observons une faible différence en fonction de la position. Quelle que soit la position r_1, la diminution du succès de livraison est importante. Pour $r_1 = 1$, avec 20% de nœuds attaquants, le taux de livraison n'est que de 20%. De plus, par la symétrie observée sur la Figure 6.3, les performances en termes de succès de livraison sont identiques pour les nœuds de niveau 1 et 4 et les nœuds de niveau 2 et 3.

Nous calculons le succès de livraison moyen du réseau par la formule suivante :

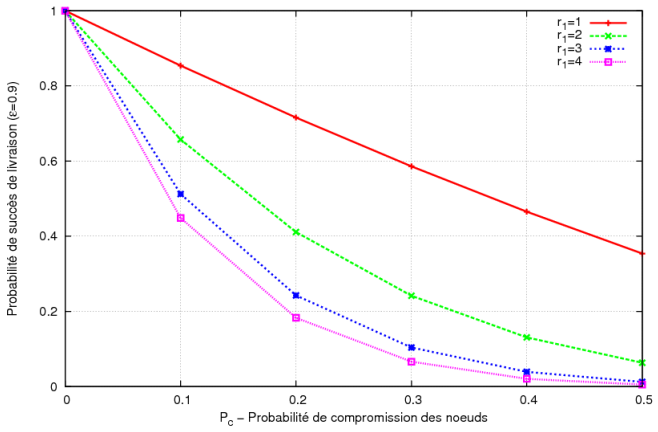
$$P_S(P_c, \varepsilon) = \sum_{\vec{r} \in S} \frac{P_{\vec{r}}(P_c, \varepsilon)}{|S|} \tag{6.4}$$

Les courbes de la Figure 6.7 illustrent la probabilité de livraison moyenne du réseau P_S en fonction de la probabilité de compromission des nœuds P_c et des biais ε. Elles sont obtenues grâce à l'équation (6.4). Comme précédemment, l'augmentation des attaques réduit de manière significative le succès de livraison moyen. Notez que la forme convexe des courbes nous rappelle les résultats présentés sur la Figure 5.16 du chapitre 5, plus particulièrement

6.3 Marche aléatoire biaisée sans réplication

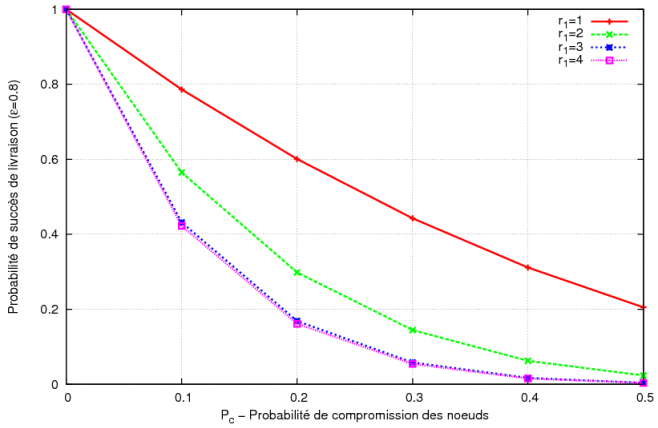


(a) $\varepsilon = 1$

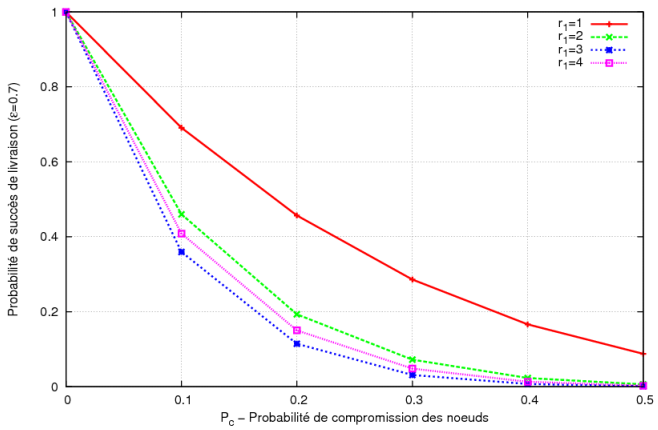


(b) $\varepsilon = 0.9$

FIGURE 6.4: Succès de livraison d'un nœud à la position r_1 en cas d'attaque (P_c).

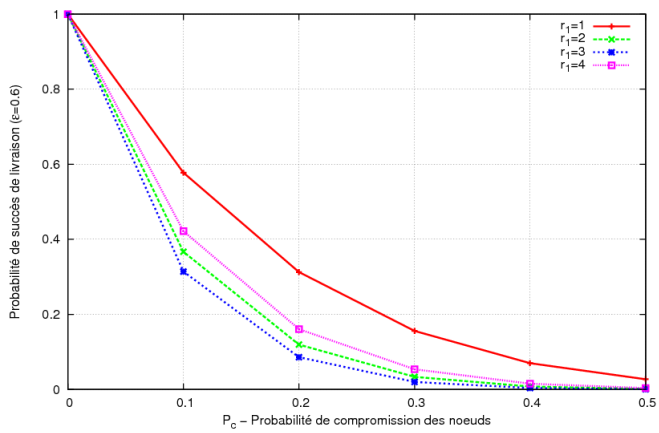


(a) $\varepsilon = 0.8$

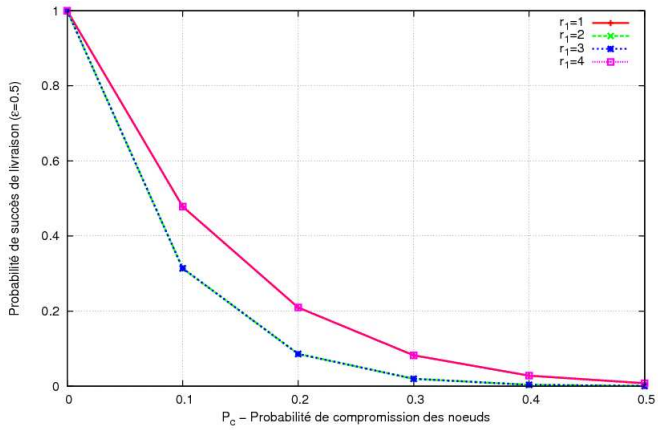


(b) $\varepsilon = 0.7$

FIGURE 6.5: Succès de livraison d'un nœud à la position r_1 en cas d'attaque (P_c).



(a) $\varepsilon = 0.6$



(b) $\varepsilon = 0.5$

FIGURE 6.6: Succès de livraison d'un nœud à la position r_1 en cas d'attaque (P_c).

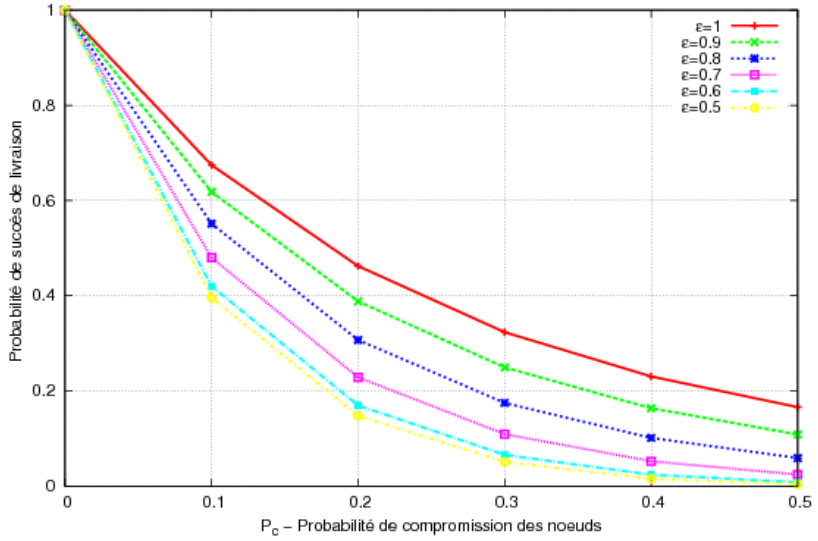


FIGURE 6.7: Succès de livraison moyen en cas d'attaque (P_c).

pour la marche aléatoire la plus biaisée $\varepsilon = 1$ correspondant au protocole le moins biaisé GBR-Rd $p = 0.5$ du chapitre 5.

Nous pouvons noter qu'avec un biais plus important, la diminution est moins importante, comme nous avons observé au chapitre 5. Par exemple, avec 20% d'attaques, la différence en termes de taux de livraison entre la marche aléatoire la moins et la plus biaisée est de l'ordre de 30%.

Là encore, le facteur qui explique ce phénomène est le même qu'au chapitre 5, la longueur moyenne des chemins. L'introduction du biais permet de diminuer considérablement cette longueur et diminue d'autant la probabilité de tomber sur un attaquant. Le biais permet donc d'obtenir un meilleur succès de livraison pour la marche aléatoire. Cependant, malgré ce gain en taux de livraison, la forme générale de ces courbes se ressemble : de forme convexe, elles traduisent une forte sensibilité aux attaques et une diminution brutale des performances même en présence d'un faible taux d'attaques.

Les résultats théoriques du succès de livraison nous confirment les résultats de taux de livraison obtenus par simulations du chapitre 5.

6.3.2 Consommation d'énergie

Pour calculer la consommation énergétique globale du réseau, nous comptons l'énergie que l'ensemble des nœuds du réseau dépense pour transmettre les paquets au puits. Nous considérons un modèle d'énergie simplifié identique aux simulations du 5 prenant en compte seulement le nombre des paquets envoyés et reçus à la couche réseau. Dans ce modèle, l'énergie dépensée par les mécanismes des couches MAC et physique tels que la sur-écoute (*overhearing*, la réception des paquets non destinés à soi-même), l'écoute inutile (*idle listening*), l'énergie dépensée pour l'envoi d'un paquet non reçu (collisions, interférences, etc.) ne sont pas pris en compte. Nous nous concentrons uniquement sur les phénomènes liés à la couche réseau.

6.3 Marche aléatoire biaisée sans réplication

Pour la marche aléatoire biaisée, la consommation d'énergie totale du réseau dépend également de la longueur moyenne des chemins. C'est le nombre total de retransmissions (l'envoi et la réception) effectuées par l'ensemble des nœuds du réseau pour acheminer les paquets au puits. Il s'agit donc de l'espérance de la longueur des chemins.

La consommation d'énergie totale $ET_{\vec{r}}$ d'un nœud $\vec{r}$ est le coût engendré (en nombre de retransmissions) au réseau pour acheminer tous ses paquets.

Elle est calculée par la formule suivante :

$$\begin{aligned} ET_{\vec{r}}(P_c,\varepsilon) &= \sum_{i=1}^{E(D_\varepsilon(\vec{r}))-1} iQ(X=i) \\ &= \sum_{i=1}^{E(D_\varepsilon(\vec{r}))-2} i(1-P_c)^{i-1}P_c + (E(D_\varepsilon(\vec{r}))-1)(1-P_c)^{(E(D_\varepsilon(\vec{r}))-1)} \end{aligned} \tag{6.5}$$

où, X est la variable aléatoire représentant le nombre de nœuds relais traversés par un paquet envoyé par un nœud $\vec{r}$. $Q(X=i)$ est la probabilité qu'un paquet rencontre i nœuds relais en chemin. La consommation d'énergie totale du réseau ET_S est calculée par la formule suivante :

$$ET_S(P_c,\varepsilon) = \sum_{\vec{r}\in S} ET_{\vec{r}}(P_c,\varepsilon) \tag{6.6}$$

La consommation d'énergie totale du réseau est divisée en deux parties : l'énergie utile EU_S, consommée par des paquets reçus avec succès et l'énergie gaspillée EG_S, consommée par des paquets perdus.

$$EU_S(P_c,\varepsilon) = \sum_{\vec{r}\in S}(E(D_\varepsilon(\vec{r}))-1)P_{\vec{r}}(P_c,\varepsilon) \tag{6.7}$$

$$EG_S(P_c,\varepsilon) = ET_S(P_c,\varepsilon) - EU_S(P_c,\varepsilon) \tag{6.8}$$

Les courbes de la Figure 6.8 sont obtenues grâce à l'équation (6.6). Elles illustrent la consommation d'énergie totale du réseau. La consommation d'énergie diminue avec l'augmentation des attaques. Ces résultats sont en accord avec les résultats obtenus par les simulations au chapitre 5 (baisse de la consommation d'énergie) avec l'augmentation d'attaques. Cette similitude confirme la pertinence du modèle théorique que nous utilisons.

L'interprétation de ces résultats est donc identique à celle donnée dans le chapitre 5 : avec les attaques, une partie des paquets est éliminée et le réseau dépense donc moins d'énergie en raison de la diminution du trafic. Là encore, le fait de biaiser la marche aléatoire fait économiser considérablement d'énergie au réseau. La baisse de la longueur moyenne des routes permet une dépense moindre pour l'acheminement des paquets.

La Figure 6.9 permet d'illustrer la part d'énergie gaspillée sur l'énergie totale en fonction d'attaques. Nous observons une augmentation importante de l'énergie gaspillée en fonction de l'intensité d'attaques. De manière assez intuitive, cela correspond à la simple augmentation du nombre de paquets perdus dû aux attaques. Cependant, nous pouvons remarquer

6.3 Marche aléatoire biaisée sans réplication

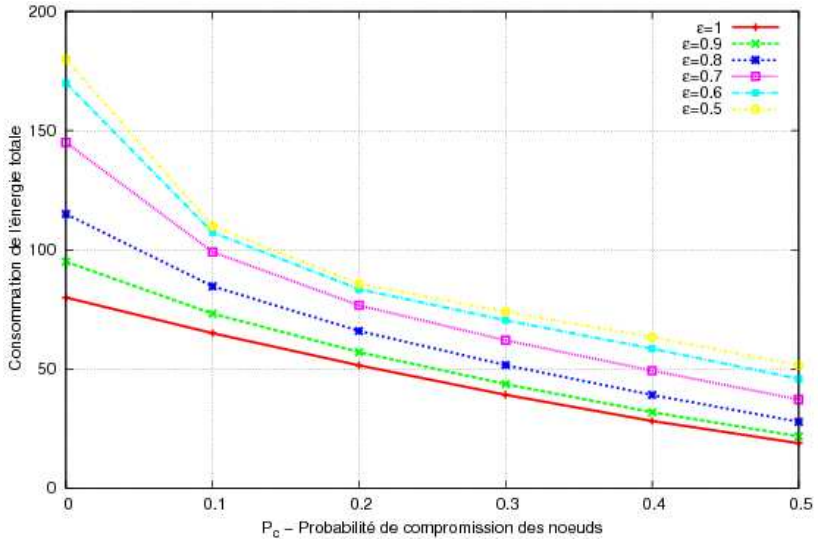


FIGURE 6.8: Consommation d'énergie totale en cas d'attaques (P_c).

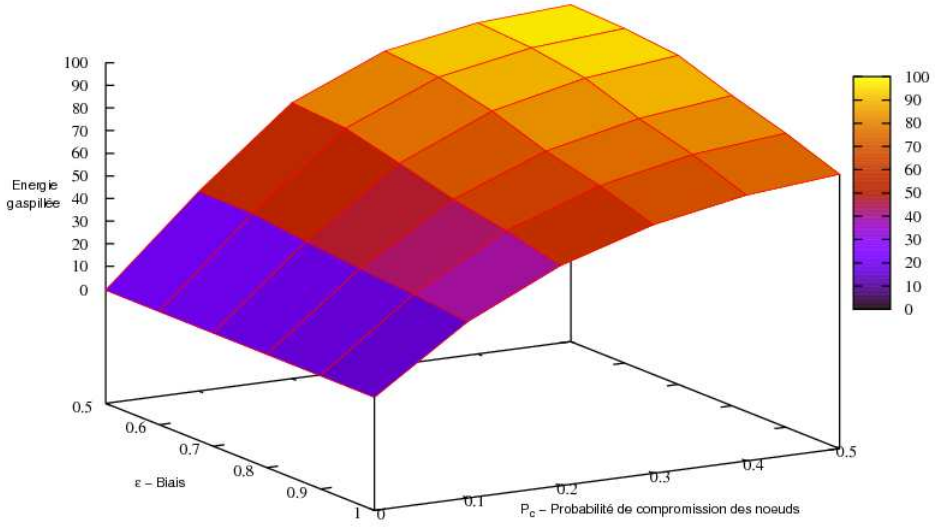


FIGURE 6.9: Part du gaspillage d'énergie à cause des paquets perdus sur l'énergie totale.

également que l'entropie accentue ce phénomène. Plus l'entropie est importante, plus l'attaque amplifie la part d'énergie gaspillée. Cela s'explique par deux éléments : d'une part, nous avons souligné que l'entropie pénalise le succès de livraison, et d'autre part, la taille des chemins est augmentée. Ces deux éléments ensemble ont pour conséquence d'augmenter la taille des chemins parcourus par les paquets perdus. Chaque paquet perdu dans le cas où le biais est minimal aura parcouru en moyenne un chemin significativement plus long que dans le cas où le biais est maximal. Un même paquet perdu dans les deux cas aura dépensé bien plus d'énergie dans le cas où le biais est minimal.

6.4 Marche aléatoire biaisée avec réplications

Nous introduisons dans cette partie les deux façons de répliquer : (i) réplications uniformes et (ii) réplications adaptatives que nous avons exposées à la section 5.2.3 du chapitre 5. Nous les introduisons ici dans la marche aléatoire biaisée et nous les évaluons en cas d'attaques *Selective forwarding*.

Comme nous l'avons souligné dans les hypothèses, une différence essentielle entre le modèle théorique utilisé ici et le graphe utilisé lors des simulations est le rapport entre le nombre de nœuds par niveau. Dans le tore, il y a le même nombre de nœuds par niveau. Dans les simulations du chapitre 5, le nombre de nœuds nœuds de niveau supérieur est plus grand que le nombre de nœuds proche du puits. Nous nous attendons à ce que cette divergence ait un impact dans le cas de réplications adaptatives.

6.4.1 Réplications uniformes

Dans cette section, le taux de réplication x est appliqué uniformément à tous les nœuds du réseau. Chaque paquet est donc répliqué x fois.

Succès de livraison

La probabilité de succès de livraison d'un nœud en fonction de x réplications de paquets est égale à la probabilité qu'au moins un paquet sur x copies soit livré avec succès. Elle est calculée par la formule suivante sous l'hypothèse d'indépendance de chaque itération.

$$\overline{P}_{\vec{r}}(P_c, \varepsilon, x) = 1 - (1 - P_{\vec{r}}(P_c, \varepsilon))^x \tag{6.9}$$

Pour la probabilité du succès de livraison moyenne du réseau en fonction de x réplications, nous calculons l'espérance de la probabilité du succès de livraison par la formule suivante :

$$\overline{P}_S(P_c, \varepsilon, x) = \sum_{\vec{r} \in S} \frac{\overline{P}_{\vec{r}}(P_c, \varepsilon)}{|S|} \tag{6.10}$$

Les courbes des Figures 6.10-6.11 représentent la probabilité de succès de livraison moyen des marches aléatoires en fonction des biais sans et avec $x = 2$ et $x = 3$ réplications calculée par les équations (6.4) et (6.10). Là encore, la forme des courbes plus concave correspond aux résultats présentés sur la Figure 5.17 du chapitre 5, plus particulièrement, pour la marche aléatoire la plus biaisée ($\varepsilon = 1$) en cas de réplications $x = 2$ (voir la Figure 6.10), correspondant au protocole le moins biaisé GBR-Rd $p = 0.5$ du 5.

La diminution du succès de livraison sous attaques est moins brutale comparée au cas sans réplication. Comme nous l'avons expliqué au chapitre 5, chaque copie répliquée est une itération indépendante de la marche aléatoire. Si le paquet original est perdu, la copie répliquée a la possibilité d'atteindre le puits, augmentant ainsi le succès de livraison de chaque paquet.

La forme générale des courbes en cas de réplications est donc très similaire aux résultats de

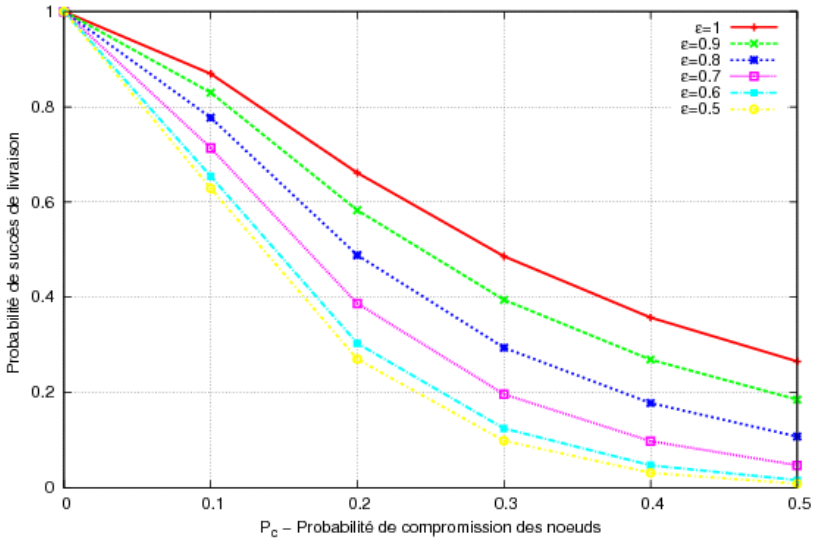


FIGURE 6.10: Succès de livraison moyen en cas d'attaque (P_c) : avec 2 réplications

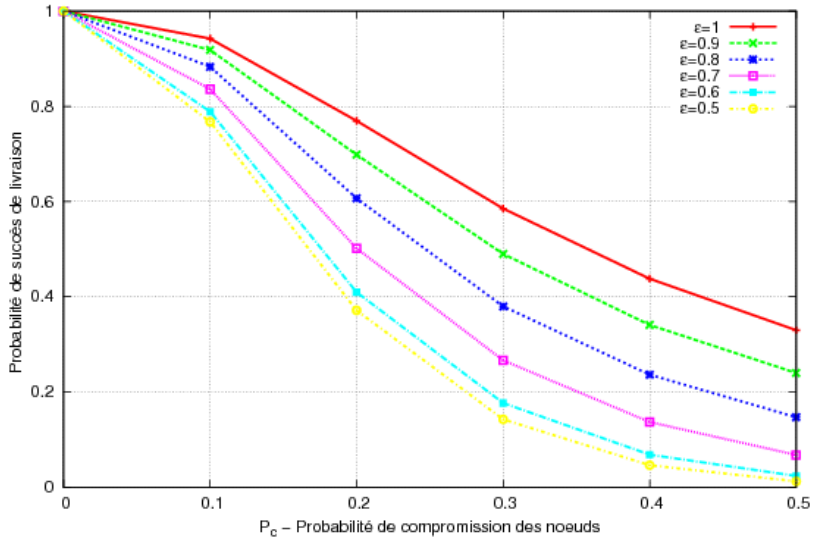


FIGURE 6.11: Succès de livraison moyen en cas d'attaque (P_c) : avec 3 réplications

nos simulations du chapitre 5. Cependant, les protocoles étudiés dans le chapitre 5 étaient plus performants en termes de taux de livraison grâce à une longueur de chemins moins importante. Comme nous l'avons souligné dans la section 6.2.2, nous avons, dans ce modèle, le biais maximum qui permet de progresser vers le puits seulement une fois sur deux en moyenne. Cela correspond donc au protocole GBR-Rd $p = 0.5$, le moins biaisé parmi les protocoles étudiés dans le chapitre 5.

Calculons maintenant le gain en termes de succès de livraison apporté par des réplications. Cela permet d'observer la différence de succès de livraison moyen entre les cas sans réplication et avec augmentation du taux de réplications.

Nous calculons par l'équation suivante, le gain en terme de succès de livraison apporté par

les réplications :

$$\overline{G}_S(P_c, \varepsilon, x) = \overline{P}_S(P_c, \varepsilon, x) - P_S(P_c, \varepsilon) \tag{6.11}$$

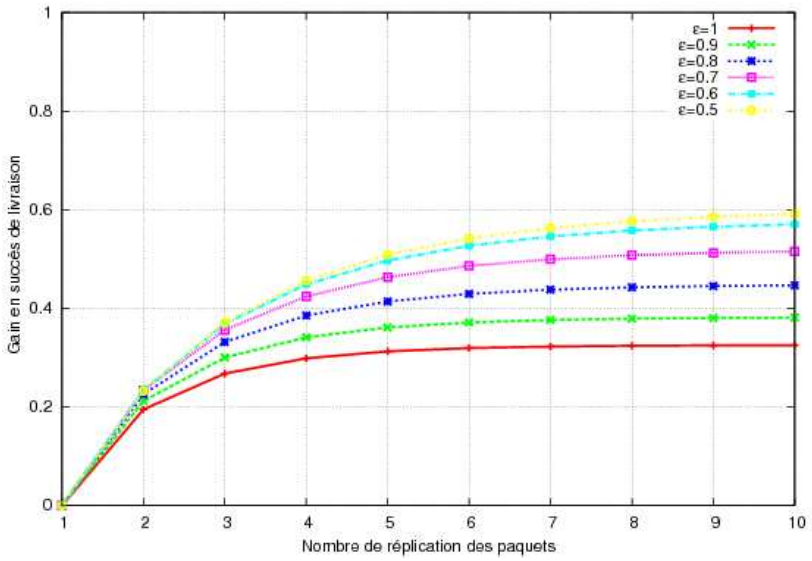


FIGURE 6.12: Gain en taux de livraison en fonction de taux de réplications x avec $P_c = 0.1$ d'attaques.

Les courbes de la Figure 6.12 illustrent le gain en taux de livraison calculé par l'équation (6.11) des marches aléatoires en fonction de l'intensité de taux de réplications des paquets.

Quel que soit le biais, nous observons un gain de 20% du succès de livraison avec $P_c = 0.1$ attaques en dupliquant des paquets (Figure 6.12). Pour la marche aléatoire la plus biaisée, nous constatons que le seuil de réplications est $x = 5$. Au-delà de cette valeur, le gain se stabilise. L'efficacité de la réplication pour cette marche et pour cette intensité d'attaque est maximale, il devient donc inutile de répliquer plus.

En résumé, les résultats théoriques sur le succès de livraison en cas de réplications uniformes correspondent aux résultats obtenus par simulations du chapitre 5. Quel que soit le biais, la marche aléatoire avec réplications de paquets permet d'améliorer le taux de livraison. Il est à noter que cette amélioration est plus conséquente pour les marches aléatoires les plus biaisées. En revanche, la réplication est très gourmande en termes de consommation énergétique comme nous l'avons observé par simulations, il convient d'en mesurer théoriquement le surcoût. Nous nous y attachons dans la section suivante.

Consommation d'énergie

Nous avons étudié la consommation d'énergie pour la marche aléatoire sans réplication. Dans cette section, nous étudions la consommation énergétique pour le cas avec réplications. Nous avons observé dans le chapitre 5 que la consommation énergétique des marches aléatoires avec réplications impactait de manière négative la résilience.

6.4 Marche aléatoire biaisée avec réplications

La consommation d'énergie totale $\overline{ET}_{\vec{r}}$ d'un nœud $\vec{r}$ est calculée par la formule suivante :

$$\overline{ET}_{\vec{r}}(P_c, \varepsilon, x) = xET_{\vec{r}}(P_c, \varepsilon) \tag{6.12}$$

La consommation d'énergie totale du réseau $\overline{ET}_S$ en fonction de x réplications de paquets par la source est donc calculée par la formule suivante :

$$\overline{ET}_S(P_c, \varepsilon, x) = \sum_{\vec{r} \in S} \overline{ET}_{\vec{r}}(P_c, \varepsilon, x) \tag{6.13}$$

Comme chaque itération de la marche aléatoire est indépendante, la consommation totale de l'énergie est égale à x fois la consommation d'énergie sans réplication.

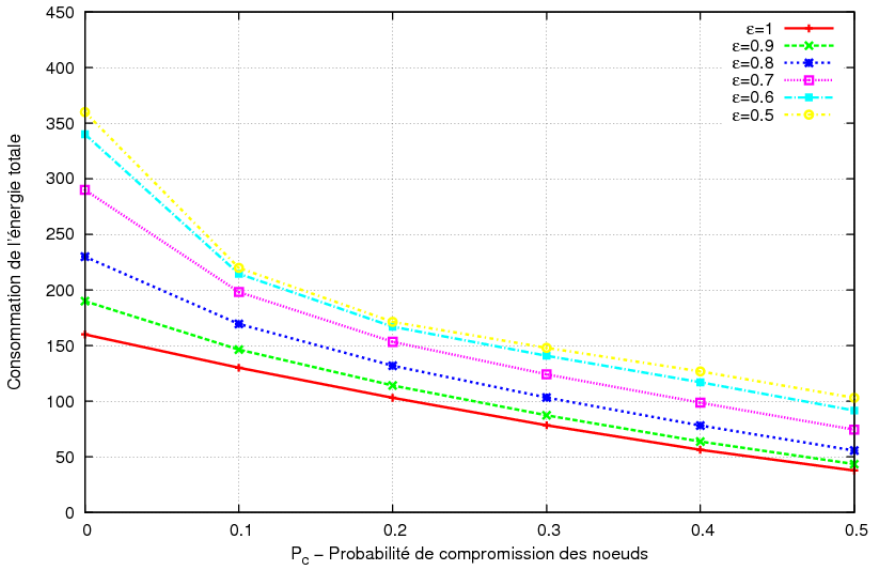
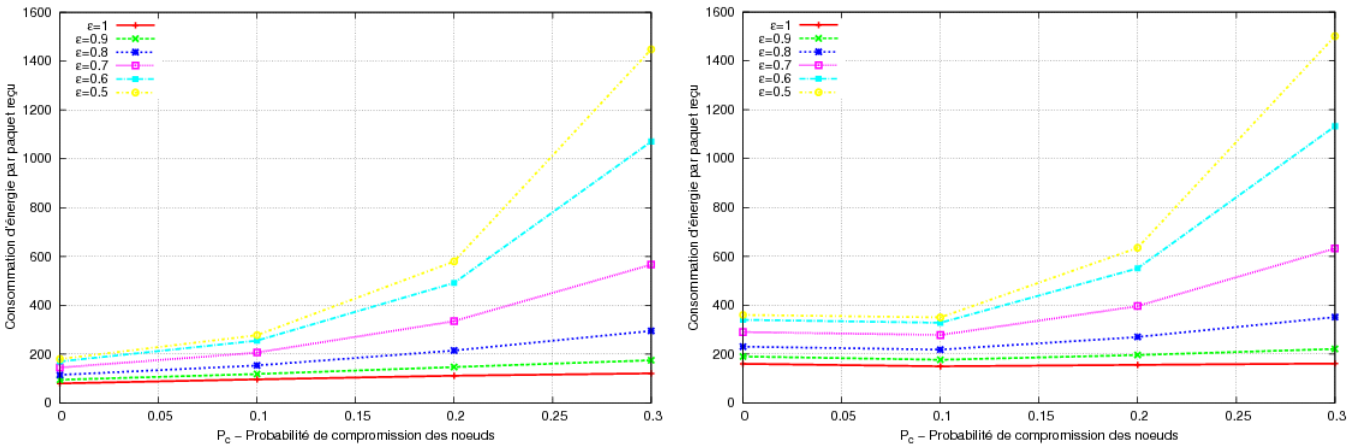


FIGURE 6.13: Consommation d'énergie totale en cas d'attaque : avec 2 réplications

Les courbes de la Figure 6.13 permettent de comparer la consommation d'énergie totale du réseau des marches aléatoires avec 2 réplications calculées avec les équations (6.6) et (6.13). Sans attaque, avec 2 réplications (Figure 6.13), la consommation d'énergie totale est 2 fois plus importante que dans le cas sans réplication (Figure 6.8). La réplication a donc un coût très important en termes d'énergie. Non seulement ces résultats théoriques sont adéquats avec les résultats par simulations du chapitre 5, mais aussi dans les proportions très proches.

Nous illustrons maintenant les courbes de la consommation d'énergie totale par paquets reçus sur la Figure 6.14 pour comparer l'efficacité en termes de consommation énergétique des marches aléatoires. Cela permet de différentier un protocole très consommateur d'énergie, mais dont le taux de livraison est haut et un protocole peu consommateur, mais dont le taux de livraison est très bas. Dis autrement, la meilleure façon de ne pas consommer d'énergie est de ne transmettre aucun paquet. Cette stratégie n'est évidemment pas satisfaisante pour un protocole routage.

Avec 2 réplications (Figure 6.14(b)) l'efficacité énergétique est moins bonne comparée au cas sans réplication (Figure 6.14(a)) : on dépense plus d'énergie pour transmettre un paquet.



(a) sans réplication

(b) avec 2 réplications

FIGURE 6.14: Consommation d'énergie par paquets reçus.

Face à un nombre important d'attaquants, la dépense d'énergie pour transmettre avec succès un paquet augmente exponentiellement pour la marche aléatoire sans biais.

6.4.2 Réplications adaptatives

Dans la marche aléatoire biaisée, nous faisons l'hypothèse que chaque nœud dispose d'une information concernant la direction du puits. Dans le cas où chaque nœud connaît sa distance au puits en nombre de sauts, nous pouvons en tirer profit pour adapter le taux de réplication pour chaque nœud.

Nous l'avons vu, un paquet provenant d'un nœud loin du puits a plus de chance de rencontrer un attaquant en route puisqu'il parcourt un chemin plus long. Partant de cette observation, nous adaptons la réplication des paquets en fonction de la distance en nombre de sauts. Un nœud lointain répliquera plus et inversement un nœud plus proche de puits répliquera moins. Chaque nœud réplique donc son paquet $\widetilde{x} = r_1$ fois, en fonction de sa distance au puits. Toutefois, nous avons fixé le taux de réplication maximum des nœuds à 3 pour diminuer la consommation énergétique. Cette stratégie de réplication est donc identique à celle du chapitre 5.

Les calculs sont effectués de manière identique au cas de réplications uniformes, mais en considérant le taux de réplications $\widetilde{x}$ pour chaque nœud en fonction de sa distance au puits.

Les courbes représentant la probabilité de succès de livraison des marches aléatoires biaisées avec les réplications adaptatives (Figure 6.15), nous montrent qu'elles permettent d'obtenir un meilleur succès de livraison que les $x = 2$ réplications uniformes pour peu d'attaques (Figure 6.10). Cependant, avec plus d'attaques, les réplications adaptatives deviennent moins efficaces.

Cela diverge des résultats obtenus par simulations au chapitre 5, où nous avons observé que les $x = 2$ réplications uniformes permettaient un meilleur taux de livraison avec peu

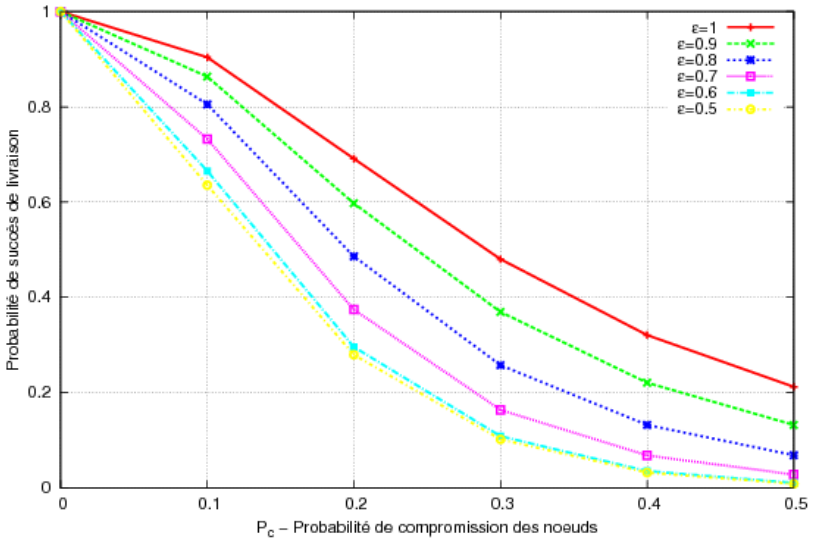


FIGURE 6.15: Succès de livraison moyen en cas d'attaques (P_c) : avec réplications adaptatives

d'attaques ($k \leq 20\%$) et les réplications adaptatives étaient plus efficaces avec plus d'attaques ($k > 20\%$) quel que soit le biais.

Cette divergence est due à la différence de support entre le modèle théorique étudié ici et le graphe utilisé par des simulations. En effet, comme nous l'avons souligné précédemment, le nombre de nœuds de différent niveau n'est pas identique. À cela s'ajoutent les caractéristiques particulières d'un tore, où les nœuds peuvent atteindre le point de collecte aussi bien dans le "bon" sens que dans le sens opposé.

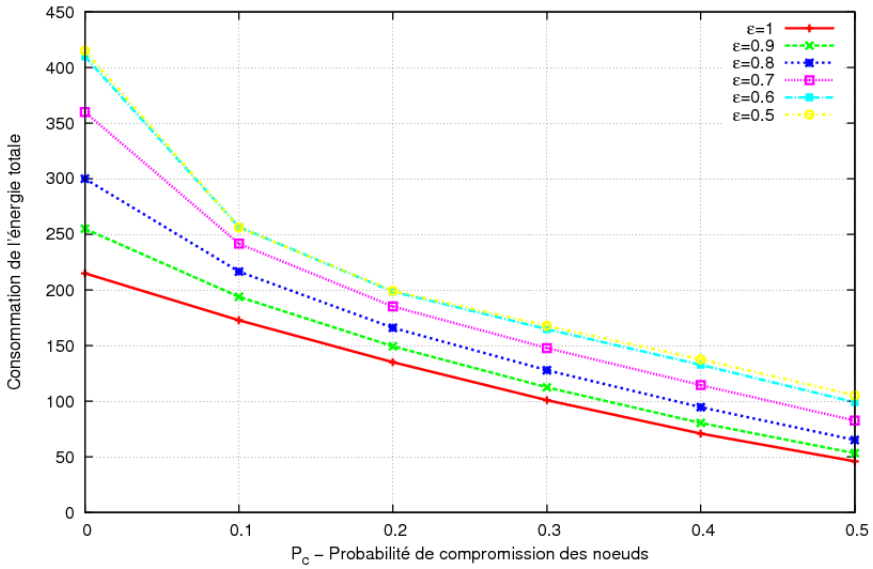


FIGURE 6.16: Consommation d'énergie totale en cas d'attaque : avec réplications adaptatives

Si nous observons maintenant la consommation d'énergie totale des marches aléatoires biaisées avec les réplications adaptatives sur la Figure 6.16, nous constatons qu'elles apportent

un surcoût d'énergie plus important que les $x = 2$ réplications uniformes. Ces résultats sont en accord avec les résultats de simulations du chapitre 5.

6.5 Évaluation de la résilience

La métrique de résilience regroupe les cinq paramètres de performance présentés dans le chapitre 3 dans l'ordre suivant :

1. **ADR - Taux de livraison moyen** : il est obtenu en calculant l'espérance de la probabilité de succès de livraison par l'équation (6.4).
2. **EE - Efficacité de la consommation énergétique** : elle est obtenue en calculant la consommation totale d'énergie du réseau par l'équation (6.6).
3. **DF - Équité de livraison du réseau** : elle est obtenue en calculant la distribution de taux de livraison parmi les nœuds. C'est l'écart type de la probabilité du succès de livraison des nœuds de l'équation (6.3).
4. **AT - Débit moyen** : c'est la quantité de données reçues au point de collecte par unité du temps. Dans notre contexte, elle est donc directement proportionnelle au succès de livraison de l'équation (6.4) puisque nous considérons seulement les phénomènes de la couche routage (bande passante non considérée).
5. **DE - Efficacité du délai moyen** : elle est directement proportionnelle à la longueur moyenne des chemins en cas d'attaques, puisque nous considérons seulement les phénomènes de la couche routage (délai de retransmissions, de propagations non considéré). Elle est obtenue en multipliant la longueur moyenne de chemins de l'équation (6.2) par la probabilité du succès de livraison de l'équation (6.3).

Tous ces paramètres sont normalisés par les valeurs extrêmes (le minimum et le maximum) des protocoles et des scénarios évalués pour obtenir la surface de la résilience, comme expliqué dans le chapitre 3.

6.5.1 Marche aléatoire sans réplication

La Figure 6.17 donne les surfaces de résilience des marches aléatoires en fonction du biais sans ($k = 0\%$) et avec attaques ($k = 30\%$) sans réplication. Sans attaque, le biais (ε) influence l'efficacité en délai (DE). Avec la diminution du biais, les paquets suivent des chemins plus longs. Cela influence légèrement la consommation d'énergie (EE), puisque les paquets voyagent plus longtemps dans le réseau. Ces phénomènes sont également observés dans les simulations du chapitre 5, et ces résultats théoriques les confirment.

En cas d'attaques ($k = 30\%$), nous observons également les mêmes phénomènes que dans les simulations au chapitre 5, le taux de livraison (ADR), le délai moyen (AT) et l'équité de livraison (DF) diminuent à cause de la perte des paquets. En revanche, l'efficacité en délai (DE) et de consommation d'énergie (EE) s'améliorent avec les attaques. Comme nous l'avons souligné dans nos résultats de simulations : la diminution du trafic dans le réseau (due aux paquets perdus) entraîne une réduction d'énergie. Les paquets reçus au puits viennent

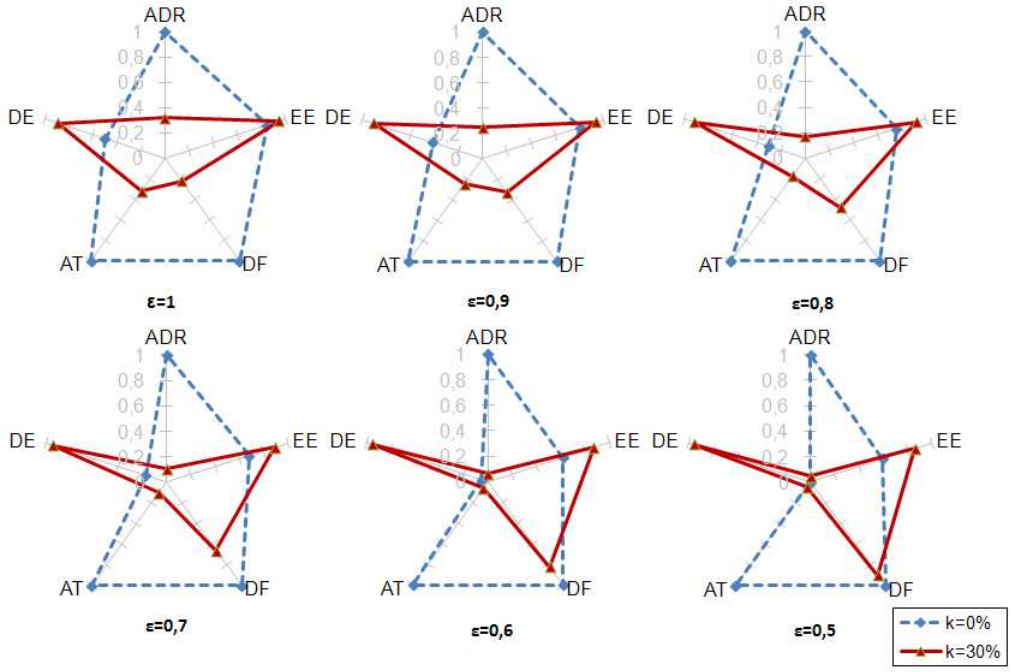


FIGURE 6.17: Surface de résilience des marches aléatoires biaisées sans réplication.

essentiellement des nœuds proches du puits, ce qui influence le délai moyen du réseau, ce qui confirme les résultats de simulations du chapitre 5.

Cependant, nous observons un phénomène contre-intuitif sur l'équité de livraison (DF) : en cas d'attaques, celui-ci s'améliore avec la diminution du biais (ε). Ceci est lié au succès de livraison extrêmement bas des marches aléatoires les moins biaisées. Tous les nœuds deviennent, en quelque sorte, équitablement mauvais en taux de livraison. Pour la marche aléatoire sans biais, nous avons constaté une symétrie en longueur de chemins en fonction de la position des nœuds sur la Figure 6.3. Cette symétrie de la longueur moyenne des chemins a influencé le succès de livraison des nœuds sur la Figure 6.6 (b). Nous constatons une faible différence en termes de succès de livraison entre des nœuds loin et des nœuds proches du puits. Si nous regardons la résilience des marches aléatoires biaisées sur la Figure 6.18, nous remarquons une grande différence entre la marche aléatoire la plus et la moins biaisée quand il y a peu d'attaques. Mais cette différence diminue avec l'intensité des attaques. Au-delà de 30% d'attaques et quel que soit le biais, la résilience chute. La forme convexe de la courbe de résilience est similaire aux résultats du chapitre précédents 5.

6.5.2 Marche aléatoire avec réplications uniformes

Les surfaces de la résilience des marches aléatoires biaisées avec 2 réplications uniformes sont présentées sur la Figure 6.19. La différence par rapport au scénario sans réplication s'exprime essentiellement sur la consommation d'énergie (EE) qui diminue. Cet aspect est également montré dans nos simulations au chapitre 5 : la réplication, si elle améliore le taux de livraison, a un coût important en énergie. Dans ce modèle théorique, la duplication uniforme

6.5 Évaluation de la résilience

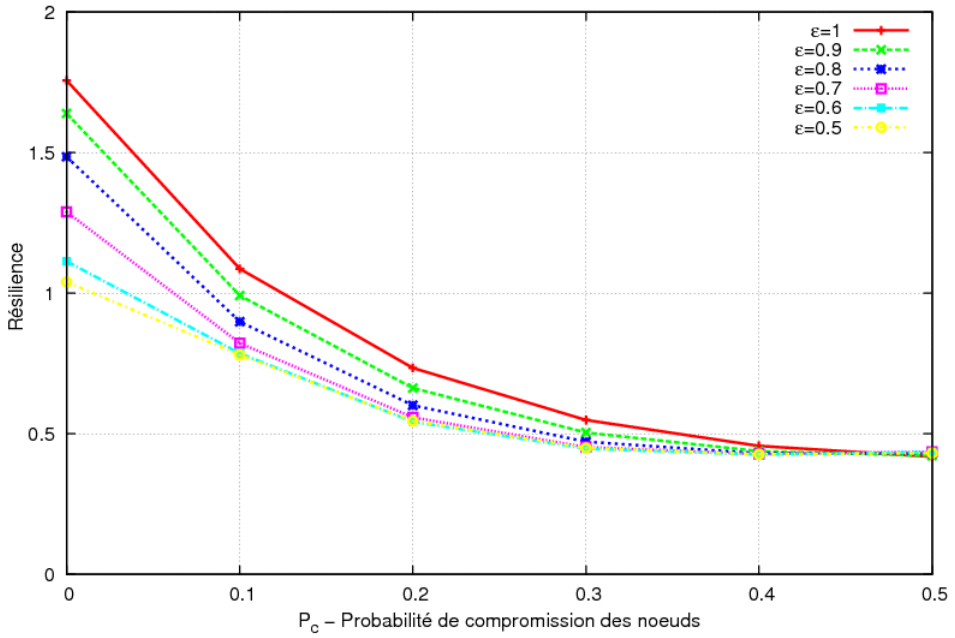


FIGURE 6.18: Résilience des marches aléatoires biaisées sans réplication.

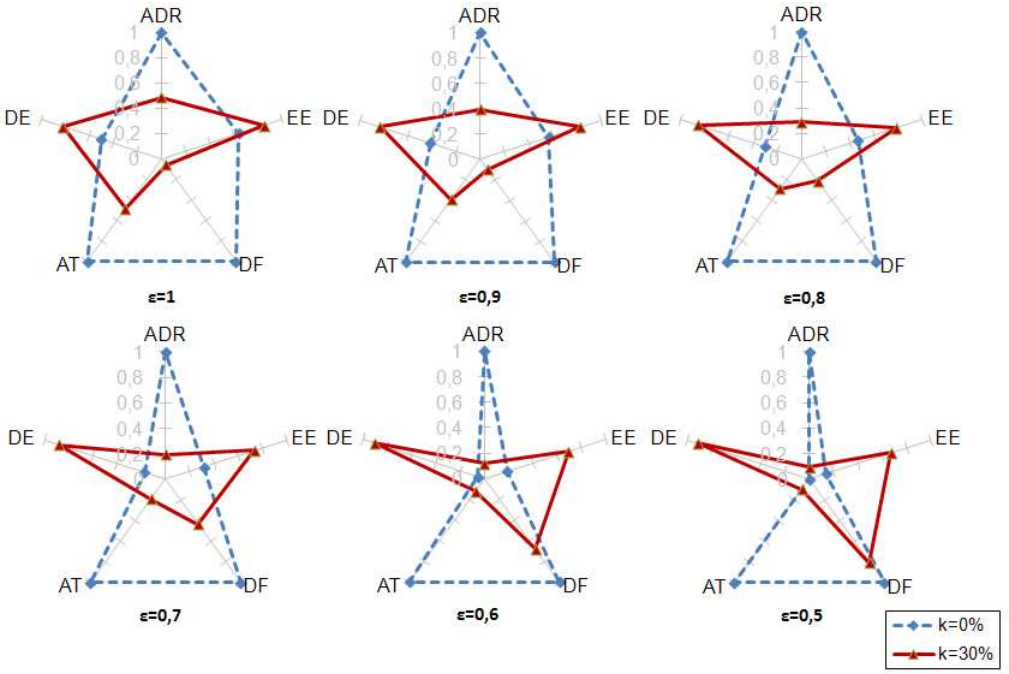


FIGURE 6.19: Surface de résilience des marches aléatoires avec réplications uniformes.

entraîne un doublement de la consommation d'énergie. L'effet du biais sur l'efficacité de délai (DE) est également conservé grâce à l'allongement des routes. Nous remarquons, de la même façon que dans nos simulations, que le taux de livraison (ADR) et le débit moyen (AT) sont améliorés par rapport au scénario sans réplication.

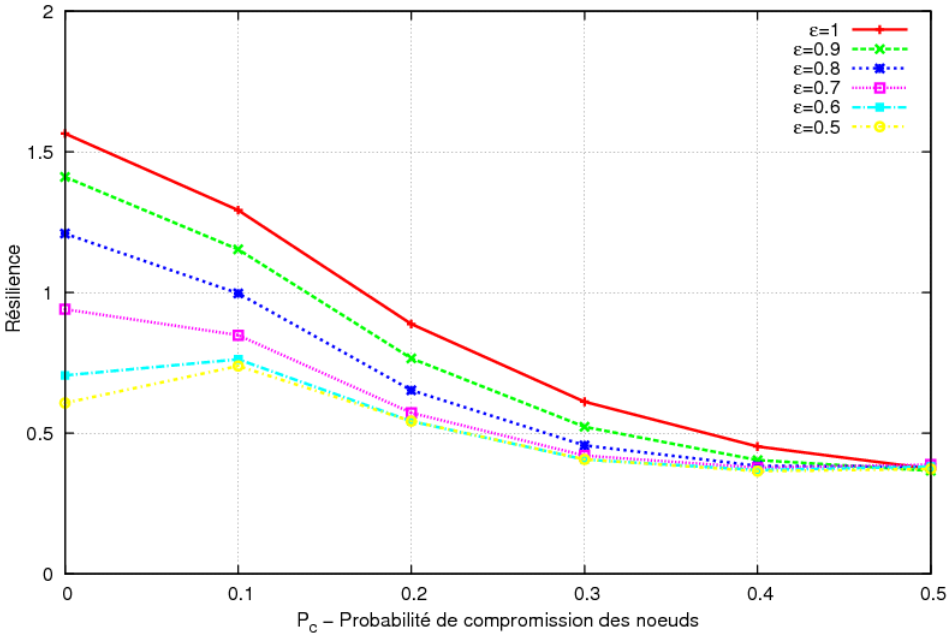


FIGURE 6.20: Résilience des marches aléatoires biaisées avec réplications uniformes.

Si nous regardons la courbe de la résilience des marches aléatoires avec 2 réplications uniformes sur la Figure 6.20, les formes des courbes des marches aléatoires les plus biaisées ($\varepsilon = \{1, 0.9, 0.8, 0.7\}$) sont très similaires aux résultats que nous avons obtenus dans nos simulations. Cependant, il s'est produit un phénomène surprenant. Pour les marches aléatoires les moins biaisées ($\varepsilon = \{0.6, 0.5\}$), leurs performances sans attaques sont plus médiocres comparées au cas d'attaques ($P_c = 0.1$). Pour les marches aléatoires les moins biaisées, la longueur de chemin est déjà pénalisante. De plus, avec les réplications de paquets nous augmentons considérablement la consommation d'énergie à cause des paquets répliques voyageant très longtemps dans le réseau. Le fait d'avoir peu d'attaques est donc bénéfique pour les marches aléatoires les moins biaisées parce que les paquets voyagent moins longtemps.

Dans notre étude par simulations du chapitre 5, nous n'avons pas vu ce phénomène parce les paquets avaient un délai d'expiration TTL (*time to live*). Au bout de TTL=32 sauts, les paquets sont détruits s'ils ne sont pas encore arrivés au puits. C'était une nécessité dans des calculs par simulations pour éviter des boucles trop longues. Grâce à cette étude, nous avons la limite théorique parce que les paquets n'expirent jamais. Les marches aléatoires les moins biaisées ($\varepsilon = \{0.6, 0.5\}$) sont donc non intéressantes dans notre contexte.

6.5.3 Marche aléatoire avec réplications adaptatives

La surface de résilience des marches aléatoires avec réplications adaptatives est présentée sur la Figure 6.21. Grâce à la représentation graphique de notre métrique, nous discernons instantanément une amélioration de la résilience pour les marches aléatoires les plus biaisées par rapport aux scénarios précédents. Cette amélioration est particulièrement prégnante

6.5 Évaluation de la résilience

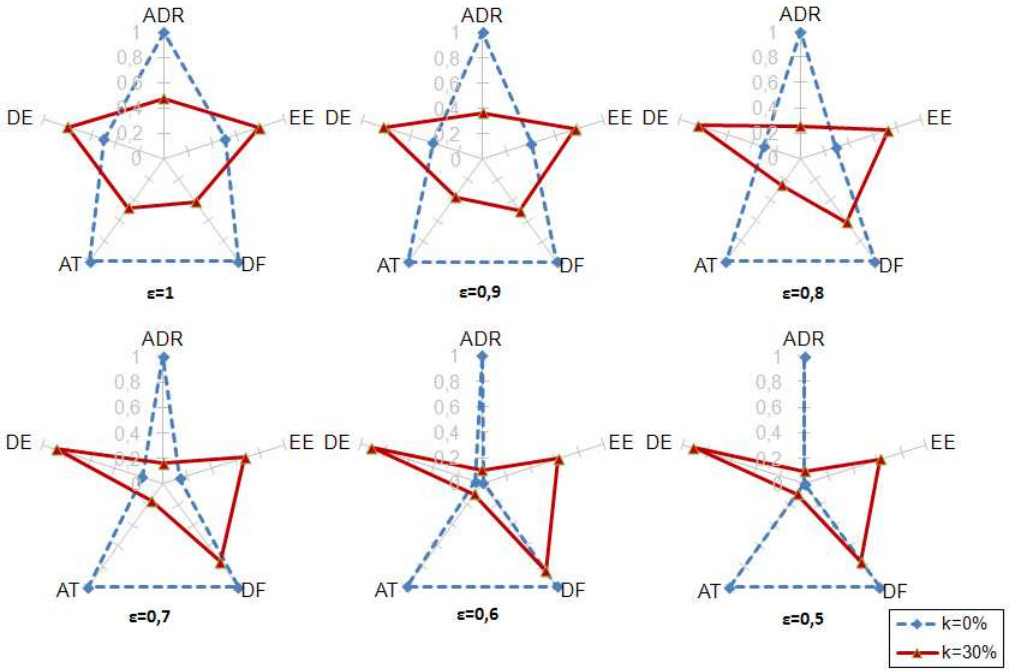


FIGURE 6.21: Surface de résilience des marches aléatoires avec réplications adaptatives.

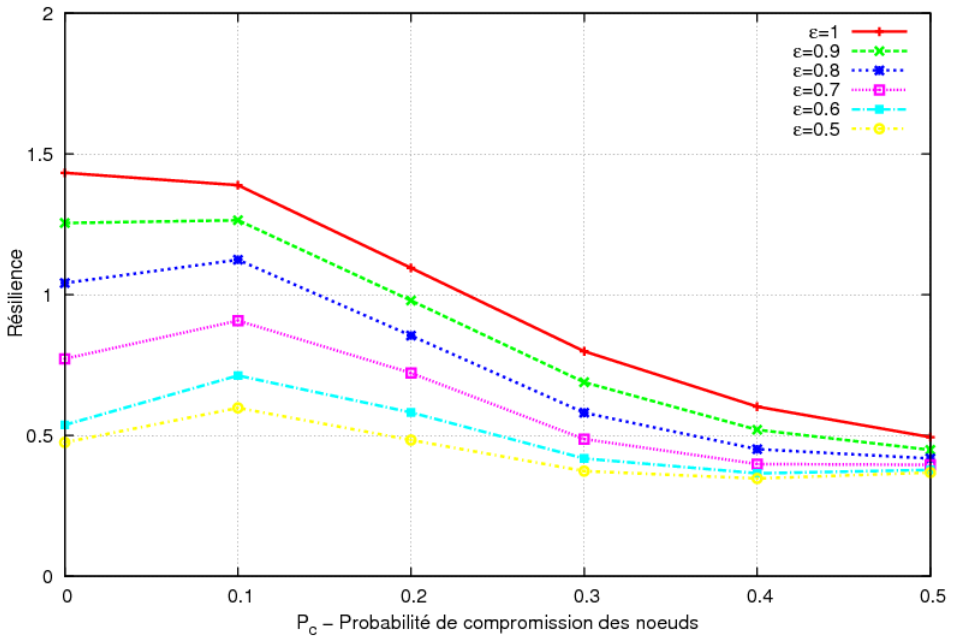


FIGURE 6.22: Résilience des marches aléatoires biaisées avec réplications adaptatives.

concernant l'équité de livraison (DF). Les paquets sont répliqués en fonction de la distance au puits ce qui améliore le taux de livraisons des nœuds lointains, augmentant ainsi l'équité de livraison. Ceci est d'autant plus visible pour la marche aléatoire la plus biaisée qu'elle avait le plus de différence en succès de livraison en fonction de la position des nœuds r_1

(Figure 6.4 (a)). C'est la marche aléatoire la plus biaisée qui profite le plus de la réplication adaptative.

L'évaluation de la résilience, montrée sur la Figure 6.22 nous confirme le résultat de simulations du chapitre 5, car nous observons une plus grande différence entre la plus et la moins biaisée en terme de résilience grâce aux réplications adaptatives.

Les marches aléatoires les moins biaisées ($\varepsilon \geq 0.8$) sont moins performantes sans attaque qu'en cas d'attaques ($P_c = 0.1$). Notez que pour les biais $\varepsilon = \{0.6, 0.5\}$, la résilience est moins importante avec les réplications adaptatives qu'avec les $x = 2$ réplications uniformes.

6.6 Conclusion

Dans ce chapitre, nous avons présenté une étude théorique de la résilience des marches aléatoires biaisées en cas d'attaques. Nous avons introduit les deux types de réplications (uniformes et adaptatives) proposées dans le chapitre 5.

Dans une première partie, nous nous sommes concentrés sur l'étude du succès de livraison et de la consommation énergétique, deux paramètres de notre métrique. Cette partie a permis de confirmer à la fois la pertinence du modèle théorique et les résultats de simulations obtenus au chapitre 5.

L'application de notre métrique aux marches aléatoires biaisées, dans une seconde partie, nous a prouvé encore son utilité. Elle nous a permis d'obtenir une vue synthétique de la résilience des marches aléatoires en fonction des valeurs de biais, selon les scénarios sans et avec les réplications.

Le seuil d'entropie bénéfique à la résilience est de $\varepsilon \geq 0.7$ quand nous les combinons avec les réplications uniformes. En dessous de cette valeur, la marche aléatoire est inefficace, à cause de la longueur des chemins trop importante, puisque sa performance est plus médiocre sans attaque qu'en cas d'attaques.

7 Conclusion

Sommaire

7.1 Bilan 126
7.2 Perspectives 129
7.2.1 Expérimentations 129
7.2.2 Mensonges sur les paquets de contrôle 129
7.2.3 Mécanismes de détection d'attaques 129
7.2.4 Codage réseau pour la résilience 130

7.1 Bilan

Notre objectif, à travers cette thèse, a été de définir et d'étudier la résilience des protocoles de routage pour les réseaux de capteurs et d'identifier des mécanismes résilients aux compromissions des nœuds. Les caractères spécifiques des réseaux de capteurs ne permettent pas d'appliquer directement les solutions de sécurité traditionnelles. De plus, les solutions cryptographiques deviennent inefficaces en cas de compromission des nœuds. Pour pallier à cela, nous avons étudié des solutions algorithmiques complémentaires à la cryptographie.

Durant la phase d'étude bibliographique, nous avons identifié les problématiques majeures de la sécurité des réseaux de capteurs, et en particulier celles touchant la couche réseau, car ses caractéristiques entraînent des vulnérabilités spécifiques. Nous avons identifié la nécessité de développer des solutions algorithmiques complémentaires aux méthodes cryptographiques traditionnelles. Ensuite, nous nous sommes intéressés aux comportements des protocoles de routage classiques en cas d'attaques. Nous avons évalué par simulations cinq protocoles de routage classiques, de différentes catégories et selon plusieurs métriques connues (le taux de livraison, la longueur moyenne des chemins, etc.). Cette étude a permis de constater que les protocoles aléatoires sont bénéfiques pour la sécurité puisqu'ils sont non prévisibles. Nous avons donc étudié l'état de l'art général des marches aléatoires et en particulier, des marches aléatoires biaisées. Ce travail préliminaire a mis en évidence le besoin de définir un cadre d'approche algorithmique pour la sécurité de réseaux de capteurs, mais aussi la nécessité de proposer des mécanismes rendant les protocoles plus résistants aux attaques internes.

Dans un premier temps, nous avons proposé le concept de résilience incluant une définition définissant clairement le terme et une métrique permettant de comparer efficacement les protocoles de routage en cas de compromission des nœuds. Notre approche se base sur une représentation graphique à deux dimensions permettant d'obtenir une vision d'ensemble de la résilience des protocoles. En d'autres mots, cette méthode permet de discerner les différents profils de protocoles (équitable en livraison, efficace en délai, gourmand en énergie, etc.) et de percevoir quels paramètres sont mis à mal en cas d'attaques. De cette façon, nous ne perdons pas d'information pour chaque paramètre de performance tout en ayant une vision synthétique. Nous avons lié cette représentation graphique (information qualitative) à une méthode d'agrégation de plusieurs paramètres (information quantitative) permettant d'obtenir une seule valeur de résilience. Grâce à cette métrique, nous avons pu établir une classification des protocoles selon leur résilience. Enfin, nous l'avons appliqué aux protocoles de routage classiques et nous les avons évalués par simulations pour montrer la pertinence de la métrique proposée.

Dans un deuxième temps, nous avons proposé les mécanismes de routage résilients que nous avons appliqués aux protocoles de routage classiques. Ces mécanismes consistent en l'introduction de comportements aléatoires, en la limitation de la longueur des chemins et en l'introduction de la réplication des données. Nous avons évalué leur résilience par simulations selon notre métrique. Cette étude a montré que non seulement ces mécanismes augmentent la résilience des protocoles classiques, mais également que notre métrique de résilience permet de la saisir. Les principaux mérites de notre proposition sont : une amélioration du taux de livraison moyen et du débit moyen, une distribution plus équitable de succès de livraison parmi les nœuds et une connectivité d'un plus grand nombre de sources au puits. Grâce

à la métrique de résilience, nous avons proposé une taxonomie permettant d'identifier les protocoles les plus résilients en cas d'attaques. Selon cette taxonomie, le routage par gradient et la marche aléatoire biaisée avec les mécanismes proposés sont les plus résilients.

Nous avons donc évalué par la suite le routage par gradient en cas d'attaques combinées pour approfondir notre étude, mais aussi pour savoir si les mécanismes proposés permettent d'augmenter la résilience même en cas d'attaques plus complexes, visant différents aspects du routage (construction des routes, paquets de contrôle, etc.).

Nous avons introduit plusieurs valeurs de biais aux variantes aléatoires du routage par gradient pour étudier l'influence de l'entropie et nous les avons comparées à sa version classique. Nous avons également évalué leur résilience en introduisant deux types de réplications (uniformes et adaptatives). Sans attaques, ce sont les variantes les plus biaisées sans réplications qui sont les plus performantes. En cas d'attaques peu importantes, les réplications uniformes sont plus efficaces, tandis qu'en cas d'attaques plus intenses, ce sont les réplications adaptatives qui se montrent les plus efficaces. Les études menées jusqu'ici étaient produites par des simulations. Nous avions donc besoin d'une justification théorique que nous avons proposée dans le chapitre 6.

Dans cette dernière partie, nous avons proposé une étude théorique en cas d'attaques non-retransmission des paquets de la marche aléatoire biaisée. Nous avons évalué l'influence du biais, mais aussi les deux types de réplications que nous avions évaluées précédemment par des simulations. En premier lieu, nous avons étudié le succès de livraison et la consommation d'énergie pour tous ces scénarios. Ensuite, nous les avons évalués selon notre métrique de résilience. Cette étude a montré que le biais est indispensable pour la résilience et le seuil d'entropie bénéfique à la résilience est $\varepsilon \geq 0.7$ quand la réplication de données est introduite. En dessous de cette valeur, la marche aléatoire est inefficace à cause de la longueur de chemins trop importante.

L'ensemble des travaux réalisés dans cette thèse se concentre autour de la résilience. Ce concept reste assez nouveau, en particulier dans le domaine des réseaux et télécommunications. Récemment, des travaux sur la résilience d'Internet [105, 106, 95] soulignent la lacune de la recherche autour de ce thème, le manque de vision systématique des problématiques concernant la résilience et d'une métrique quantitative pour la mesurer.

À travers cette thèse, nous avons voulu donner notre vision sur ce thème en nous concentrant sur les problématiques de sécurité des protocoles de routage dans le contexte des réseaux de capteurs. Nous avons ainsi proposé un cadre avec notre définition de la résilience et une métrique pour la mesurer. Nous sommes partis de la définition initiale de la résilience en mécanique, les capacités des matériaux à résister aux chocs, en faisant une analogie avec les réseaux résistants aux attaques. Nous avons voulu souligner un problème majeur de sécurité des réseaux de capteurs qu'est la compromission des nœuds. La compromission permet aux adversaires de pénétrer à l'intérieur du périmètre de sécurité et c'est la structure interne du réseau qui est ainsi modifiée. Ces nœuds compromis font partie du réseau et ont donc déjà connaissance des informations secrètes. Nous avons voulu aller au-delà des solutions cryptographiques, traditionnellement utilisées pour la protection de base, et apporter des solutions algorithmiques complémentaires. De notre point de vue, nous devons considérer simultanément plusieurs métriques essentielles (succès de livraison, équité de livraison, consommation d'énergie, délai et débit, etc.) pour exprimer la résilience. La métrique de résilience que nous

avons proposée dans cette thèse permet non seulement d'établir une relation d'ordre entre les protocoles, mais aussi d'obtenir une vision synthétique sur tous les aspects importants, trivialement parlant, en un coup d'œil. L'originalité de ce travail réside dans la liaison entre cette représentation graphique (qualitative) et la méthode d'agrégation de plusieurs paramètres (quantitative). En nous servant de cet outil, nous avons étudié plusieurs protocoles de routage classiques. Nous avons voulu comprendre les phénomènes liés à la résilience, mais aussi découvrir les mécanismes qui peuvent être bénéfiques à la résilience. Nous avons voulu apporter des mécanismes de routage résilients sans forcement détecter les attaques et les éliminer comme dans un IDS. Cependant, nous sommes conscients que certaines méthodes de détections d'intrusion (IDS) simples peuvent se montrer utiles pour permettre aux réseaux d'être plus dynamiques en s'adaptant à leur environnement. Par exemple, en observant le trafic dans son voisinage un nœud peut changer sa stratégie de routage en fonction des anomalies détectées. Grâce à la détection des attaques, les nœuds peuvent également ajuster leur taux de réplications de données en limitant le gaspillage d'énergie. La réplication du trafic est gourmande en consommation d'énergie et ce n'est pas souhaitable, en particulier dans les réseaux de capteurs avec des contraintes fortes en énergie. Il existe d'autres méthodes permettant d'ajouter de la redondance aux données tels que le codage réseau qui constitue une des perspectives de cette thèse. Au lieu de simplement relayer les paquets de données, les nœuds peuvent diviser des paquets, combiner des informations qu'ils contiennent et les transmettre.

Tout au long de cette thèse, nous nous sommes concentrés sur la sécurité de la couche réseau, afin d'isoler l'impact dû aux attaques et mieux comprendre les phénomènes liés à la résilience. Cependant, il est clair que les problématiques de sécurité ne peuvent se limiter uniquement à la couche réseau. Même si certaines attaques ciblent une couche en particulier, leurs impacts peuvent se percuter sur le déroulement des autres couches. Il est donc important d'étudier l'ensemble des couches pour apporter des solutions complètes. Toutefois, nous pensons que notre métrique de résilience permettra également de saisir les phénomènes des couches inférieures. Une autre perspective est d'étudier les mécanismes proposés par simulations en considérant des couches MAC/PHY non idéales (interférences, collisions, etc.) et de les étudier par des expérimentations.

Les mécanismes résilients aux attaques sont très liés à ceux de la tolérance aux fautes, la défaillance des nœuds et des liens. En ce sens, nous pouvons considérer les nœuds compromis qui ne retransmettent pas des paquets comme des défaillances non prévisibles. Les nœuds compromis peuvent donc être considérés comme non intelligents et agissants arbitrairement et individuellement comme dans le modèle de tolérance aux fautes byzantines (inspiré par le problème des généraux byzantins [116]). Toutefois, nous pensons qu'il faut aller au-delà de modèles simples, en considérant des modèles d'adversaires plus réalistes et plus complexes, où les nœuds compromis sont intelligents et peuvent communiquer entre eux en s'entendant sur des stratégies d'attaques communes, ou en mentant aux nœuds légitimes. Les travaux effectués au cours de la thèse offrent plusieurs perspectives que nous détaillons dans la section suivante.

7.2 Perspectives

Dans cette section, nous présentons les perspectives de cette thèse qui se situent dans l'extension des travaux réalisés.

7.2.1 Expérimentations

Dans cette thèse, nous avons étudié la résilience du routage des réseaux de capteurs en nous basant sur l'évaluation des protocoles par des simulations et par des calculs théoriques. Nous avons considéré des couches MAC/PHY idéales (sans interférence, sans perte due aux collisions) afin de se concentrer uniquement sur les aspects de routage et sur l'impact des attaques sur les performances des protocoles de routage. Une première perspective est de poser une hypothèse moins forte et d'étendre notre étude : (i) par des simulations en considérant des couches MAC/PHY non idéales (ii) par des expérimentations afin d'évaluer les performances dans un cas réel. Le simulateur WSNet [80] donne la possibilité de définir la propagation des signaux radio avec un grand réalisme que nous pouvons exploiter et la plateforme Senslab [11] donne la possibilité de réaliser des expérimentations à large échelle (256 nœuds).

7.2.2 Mensonges sur les paquets de contrôle

Une deuxième perspective est d'approfondir notre étude sur les attaques plus complexes telles que le mensonge dans les paquets de contrôle. Pour travailler dans cette direction, nous avons commencé une étude préliminaire et nous avons déposé un brevet [117]. Dans ce brevet, nous avons proposé des mécanismes résilients appliqués aux variantes aléatoires des routages géographiques pour sa construction des routes. Tout d'abord, nous avons proposé d'utiliser une fonction médiane qui est une fonction d'agrégation résiliente [118] à la place des fonctions minimum ou maximum pour la sélection du prochain saut pour se défendre contre les mensonges sur les paquets de contrôle. Ensuite, nous avons proposé d'utiliser plusieurs stratégies de sélection des routes et d'en choisir une aléatoirement afin de rendre le routage plus imprévisible. Nous allons donc continuer dans cette direction et évaluer les performances des différentes stratégies par des simulations ou par des expérimentations.

7.2.3 Mécanismes de détection d'attaques

Comme nous avons discuté précédemment, dans cette thèse, nous nous sommes concentrés plus particulièrement sur les mécanismes passifs quand le réseau subit des attaques sans forcement les détecter et les éliminer de manière réactive. Les mécanismes résilients que nous avons proposés ont l'avantage de rendre les réseaux imprévisibles grâce aux comportements aléatoires et à la réplication des paquets. Il serait intéressant d'étudier aussi les mécanismes de détection des attaques (une sorte d'IDS) que l'on peut mettre facilement en place en profitant par exemple de la sur-écoute (overhearing) [1] pour savoir si les voisins retransmettent

1. recevoir les paquets qui ne sont pas destinés à soi-même

les paquets, de l'analyse du trafic au niveau du puits ou au niveau de voisinage pour détecter des activités anormales. Nous pouvons donc adapter le taux de réplication ou l'entropie du réseau en fonction de l'intensité des attaques, puisque répliquer les paquets en absence d'attaques apporte un surcoût inutile.

7.2.4 Codage réseau pour la résilience

Une autre direction de perspective serait d'étudier le codage réseau (*network coding*) [119] à la place de la réplication des paquets pour diminuer le surcoût en consommation d'énergie. Concernant cette perspective, notre équipe de recherche au laboratoire CITI a commencé des études de codage réseau pour la sécurité [120]. Il serait intéressant d'identifier les vulnérabilités de cette méthode, mais aussi de comparer le surcoût énergétique avec notre approche. Notamment en évaluant la performance de cette méthode avec notre métrique de résilience.

Liste des publications

- **Journaux internationaux :**
  1. O.Erdene-Ochir, A. Kountouris, M.Minier and F.Valois. "A New Metric to Quantify Resiliency in Networking". In *IEEE Communication letters*, October 2012.
  2. O.Erdene-Ochir, A. Kountouris, M.Minier and F.Valois. "Enhancing Resiliency Against Routing Layer Attacks : Gradient based Routing in Focus". In *IARIA on-line journals, in International Journal on Advances in Networks and Services*, July 2011.

- **Conférences internationales :**
  1. O.Erdene-Ochir, M.Minier, F.Valois and A. Kountouris. "Resiliency Taxonomy of Routing Protocols in Wireless Sensor Networks". In *IEEE LCN 2012, the 37th International Conference on Local Computer Networks*, Florida, USA, October 2012.
  2. O.Erdene-Ochir, M.Minier, F.Valois and A. Kountouris. "Toward Resilient Routing for Wireless Sensor Networks : Gradient based Routing in Focus". In *Sensorsomm 2010, the 4th International Conference on Sensor Technologies and Applications*, Venice, Italy, July 2010. (**Best Paper Award**).
  3. O.Erdene-Ochir, M.Minier, F.Valois and A. Kountouris. "Resiliency of Wireless Sensor Networks : Definition and Analysis". In *IEEE ICT 2012, the 17th International Conference on Telecommunications*, Doha, Qatar, April 2010.

- **Brevet :**
  1. O. Erdene-Ochir, M.Minier, F.Valois and A. Kountouris, "Méthode résiliente à la présence des noeuds compromis pour la détermination des routes par un protocole de routage dans un réseau", N°FR-1158828, Sept.2011 (Int. : N°PCT/FR2012/052143, Sept. 2012).

- **Rapport de recherche :**
  1. O.Erdene-Ochir, M.Minier, F.Valois and A. Kountouris. *Resilient Networking for Wireless Sensor Networks*, in *INRIA Research Report*, RR-7230, France, March 2010.

Bibliographie

[1] O. Erdene-Ochir, M.Minier, F.Valois, and A.Kountouris, "Resiliency of wireless sensor networks : Definitions and analyses," in *IEEE International Conference on Telecommunications (ICT)*, Qatar, April 2010.

[2] I. Mabrouki, G. Froc, and X. Lagrange, "Biased random walk model to estimate routing performance in sensor networks," *9eme Rencontres Francophones sur les Aspects Algorithmiques des Telecommunications*, pp. 73–76, 2007.

[3] "http ://aresa2.minalogic.net/."

[4] Y. Ma, M. Richards, M. Ghanem, Y. Guo, and J. Hassard, "Air pollution monitoring and mining based on sensor grid in london," in *Sensors*, vol. 8, June 2008.

[5] R. Beckwith, D. Teibel, and P. Bowen, "Report from the field : Results from an agricultural wireless sensor network." in *In the 29th Annual IEEE International Conference on Local Computer Networks (LCN)*, Washington, DC, USA, 2004.

[6] K. Pister, T. Phinney, and P. Thubert, "Industrial routing requirements in low-power and lossy networks." *RFC5673 (Proposed Standard)*, October 2009.

[7] A. Brandt, J. Buron, and G. Porcu, "Home automation routing requirements in low-power and lossy networks." *RFC5826 (Proposed Standard)*, April 2010.

[8] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E.Cayirci, "Wireless sensor networks : a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, January 2002.

[9] D. Barakah, "A survey of challenges and applications of wireless body area network (wban) and role of a virtual doctor server in existing architecture," in *Third International Conference on Intelligent Systems, Modelling and Simulation (ISMS)*, Febrary 2012.

[10] "http ://bullseye.xbow.com :81/products/productdetails.aspx ?sid=164."

[11] "http ://www.senslab.info/."

[12] "http ://www.shimmer-research.com/."

[13] E. D. L. Andersson and L. Zhang, "Report from the IAB workshop on Unwanted Traffic March 9-10, 2006," RFC 4948 (Informational), August 2007. [Online]. Available : http://www.ietf.org/rfc/rfc4948.txt

[14] J. P. Walters, Z. Liang, W. Shi, and V. Chaudhary, "Wireless sensor network security : A survey," in book chapter of security," in *Security in Distributed, Grid, Mobile, and Pervasive Computing*. CRC Press, 2007.

[15] X. Chen, K. Makki, K. Yen, and N. Pissinou, "Sensor network security : a survey," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 2, pp. 52–73, 2009. [Online]. Available : http://dx.doi.org/10.1109/SURV.2009.090205

[16] A. Wood and J. Stankovic, "Denial of service in sensor networks," *Computer*, vol. 35, no. 10, pp. 54–62, October 2002.

[17] C. Karlof and D. Wagner, "Secure routing in wireless sensor networks : attacks and countermeasures," *Ad Hoc Networks*, vol. 1, no. 2-3, pp. 293–315, August 2003.

[18] A. Becher, Z. Benenson, and M. Dornseif, "Tampering with motes : Real-world attacks on wireless sensor networks," in *Sicherheit - Schutz und Zuverlässigkeit, Beiträge der 3. Jahrestagung des Fachbereichs Sicherheit der Gesellschaft für Informatik e.v. (GI)*, Magdeburg, Germany, February 2006.

[19] "http ://www-coronis-com.dyn.elster.com/."

[20] R. J. Ellison, R. C. Linger, T. Longstaff, and N. R. Mead, "Survivable network system analysis : A case study," *IEEE Software*, vol. 16, no. 4, pp. 70–77, July 1999.

[21] J. P. G. Sterbenz, R. Krishnan, R. Hain, A. Jackson, D. Levin, R. Ramanathan, and J. Zao, "Survivable mobile wireless networks : issues, challenges, and research directions," in *Proceedings of the 2002 ACM Workshop on Wireless Security (WiSe)*, Atlanta, USA, September 2002.

[22] J. Anderson, "Computer security threat monitoring and surveillance," *James P. Anderson Co*, 1980.

[23] D. Denning, "An intrusion detection model," 1986.

[24] K. Scarfone and P. Mell, "Guide to intrusion detection and prevention systems (idps)," *Computer Security Resource Center*, 2010.

[25] C. Shannon, "The mathematcal theory of communication," *Urbana : University of Illinois Press*, 1949.

[26] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing*, S. US, Ed., vol. 353, 1996.

[27] C. Schurgers and M. B. Srivastava, "Energy efficient routing in wireless sensor networks," in *Military Communications Conference Proceedings on Communications for Network-Centric Operations : Creating the Information Force*, vol. 1, McLean, USA, October 2001.

[28] B. Karp and H. T. Kung, "Gpsr : greedy perimeter stateless routing for wireless networks," in *Proceedings of the 6th annual international conference on Mobile computing and networking*, Boston, USA, August 2000.

[29] S. D. Servetto and G. Barrenechea, "Constrained random walks on random graphs : routing algorithms for large scale wireless sensor networks," in *Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications*, Atlanta, USA, September 2002.

[30] B. Parno, A. Perrig, and V. Gligor, "Distributed detection of node replication attacks in sensor networks," in *IEEE Symposium on Security and Privacy (S&P)*, Oakland, USA, May 2005.

[31] A. Perrig, J. Stankovic, D. Wagner, and C. Rosenblatt, "Security in wireless sensor networks," *Communications of the ACM*, vol. 47, no. 6, pp. 53–57, June 2004.

[32] W. Znaidi and M. Minier, "Proposition de gestion des clés et de contrôle d'accès dans un réseau de capteurs," in *10ème Journées Doctorales Informatique et Réseau*, Belfort, France, February 2009. [Online]. Available : http://hal.archives-ouvertes.fr/hal-00402758/en/

[33] X. Wenyuan, M. Ke, W. Trappe, and Z. Yanyong, "Jamming sensor networks : attack and defense strategies," *IEEE Network*, vol. 20, no. 3, pp. 41–47, May 2006.

[34] J. T. Chiang and Y.-C. Hu, "Dynamic jamming mitigation for wireless broadcast networks," in *IEEE INFOCOM*, 2008.

[35] ——, "Cross-layer jamming detection and mitigation in wireless broadcast networks," in *Proceedings of the 13th Annual International Conference on Mobile Computing and Networking (MOBICOM)*, Montréal, Canada, September 2007.

[36] A. Chan, X. Liu, G. Noubir, and B. Thapa, "Broadcast control channel jamming : Resilience and identification of traitors," in *IEEE International Symposium on Information Theory (ISIT)*, Nice, France, June 2007.

[37] P. Tague, M. Li, and R. Poovendran, "Mitigation of control channel jamming under node capture attacks," *IEEE Transactions on Mobile Computing*, vol. 8, no. 9, pp. 1221–1234, 2009.

[38] A. D. Wood, J. A. Stankovic, and S. H. Son, "Jam : A jammed-area mapping service for sensor networks," in *Proceedings of the 24th IEEE Real-Time Systems Symposium (RTSS)*, Cancun, Mexico, December 2003.

[39] W. Znaidi, M. Minier, and J.-P. Babau, "An ontology for attacks in wireless sensor networks," INRIA, Research Report RR-6704, 2008.

[40] A. L.Toledo and X. Wang, "Robust detection of mac layer denial-of-service attacks in csma/ca wireless networks," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 3, pp. 347–358, September 2008.

[41] A. Rachedi, "Contributions a la securite dans les reseaux mobiles ad-hoc," Ph.D. dissertation, University of Avignon, France, November 2008.

[42] Y. W. Law, M. Palaniswami, L. V. Hoesel, J. Doumen, P. Hartel, and P. Havinga, "Energy-efficient link-layer jamming attacks against wireless sensor network mac protocols," *ACM Transactions on Sensor Networks*, vol. 5, no. 1, pp. 1–38, 2009.

[43] J. Newsome, E. Shi, D. Song, and A. Perrig, "The sybil attack in sensor networks : analysis & defenses," in *Proceedings of the Third International Symposium on Information Processing in Sensor Networks (IPSN)*, Berkeley, USA, April 2004.

[44] C. A. Melchor, B. Ait-Salem, P. Gaborit, and K. Tamine, "Active detection of node replication attacks," *International Journal of Computer Science and Network Security (IJCSNS)*, vol. 9, pp. 13–21, February 2009.

[45] S. Marti, T. J. Giuli, K. Lai, and M. Baker, "Mitigating routing misbehavior in mobile ad hoc networks," in *6th Annual International Conference on Mobile Computing and Networking (MobiCom)*, Boston, USA, August 2000.

[46] P. Michiardi and R. Molva, "Core : a collaborative reputation mechanism to enforce node cooperation in mobile ad hoc networks," in *Communications and Multimedia*

Security, ser. IFIP Conference Proceedings, B. Jerman-Blazic and T. Klobucar, Eds., vol. 228. Portoroz, Slovenia : Kluwer, September 2002.

[47] D. Ganesan, R. Govindan, S. Shenker, and D. Estrin, "Highly-resilient, energy-efficient multipath routing in wireless sensor networks," in *2nd ACM international symposium on Mobile ad hoc networking & computing*. Long Beach, USA : ACM, October 2001.

[48] N. Hanusse, D. Ilcinkas, A. Kosowski, and N. Nisse, "Locating a target with an agent guided by unreliable local advice : how to beat the random walk when you have a clock ?" in *In the 29th Annual ACM Symposium on Principles of Distributed Computing (PODC)*, Zurich, Switzerland, July 2010.

[49] Y.-C. Hu, A. Perrig, and D. B. Johnson, "Packet leashes : A defense against wormhole attacks in wireless networks," in *22nd Annual Joint Conference of the IEEE Computer and Communications Societies*, San Fransisco, USA, April 2003.

[50] L. Hu and D. Evans, "Using directional antennas to prevent wormhole attacks," in *Network and Distributed System Security Symposium*. San Diego, USA : The Internet Society, February 2004.

[51] W. Znaidi, M. Minier, and J.-P. Babau, "Detecting wormhole attacks in wireless networks using local neighborhood information," in *Proceedings of the IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. Cannes, France : IEEE, September 2008.

[52] K. Heurtefeux and F. Valois, "Distributed localization protocol for routing in a noisy wireless sensor network," in *In the 5th International Conference on Mobile Ad-hoc and Sensor Networks*, Wu Yi Shan, China, December 2009.

[53] D. Wagner, "Resilient aggregation in sensor networks," in *Proceedings of the 2nd ACM Workshop on Security of ad hoc and Sensor Networks (SASN)*, Washington, USA, October 2004.

[54] H. Alzaid, E. Foo, and J. G. Nieto, "Secure data aggregation in wireless sensor network : a survey," in *Proceedings of the 6th Australasian Information Security Conference (AISC)*, Darlinghurst, Australia, January 2008.

[55] L. Hu and D. Evans, "Secure aggregation for wireless network," in *Symposium on Applications and the Internet Workshops (SAINT)*, Orlando, USA, January 2003.

[56] P. Jadia and A. Mathuria, "Efficient secure aggregation in sensor networks," in *11th International Conference on High Performance Computing (HiPC)*, Bangalore, India, December 2004.

[57] A. Perrig, R. Szewczyk, V. Wen, D. E. Culler, and J. D. Tygar, "Spins : security protocols for sensor netowrks," in *7th Annual International Conference on Mobile Computing and Networks*, Rome, Italy, July 2001.

[58] Y. Yang, X. Wang, S. Zhu, and G. Cao, "Sdap : a secure hop-by-hop data aggregation protocol for sensor networks," in *Proceedings of the 7th ACM Interational Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, Florence, Italy, May 2006.

[59] H. Chan, A. Perrig, B. Przydatek, and D. Song, "SIA : secure information aggregation in sensor networks," *Journal of Computer Security*, vol. 15, no. 1, pp. 69–102, 2007.

[60] H. Chan, A. Perrig, and D. X. Song, "Secure hierarchical in-network aggregation in sensor networks," in *Proceedings of the 13th ACM Conference on Computer and Communications Security (CCS)*, Alexandria, USA, October 2006.

[61] W. Du, J. Deng, Y. S. Han, and P. Varshney, "A witness-based approach for data fusion assurance in wireless sensor networks," in *IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 3, San Francisco, USA, December 2003.

[62] A. Mahimkar and T. Rappaport, "Securedav : a secure data aggregation and verification protocol for sensor networks," in *IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 4, Dallas, USA, November 2004.

[63] Z. Yu and Y. Guan, "A dynamic en-route scheme for filtering false data injection in wireless sensor networks," in *25th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, Barcelona, Spain, April 2006.

[64] D. Westhoff, J. Girão, and M. Acharya, "Concealed data aggregation for reverse multicast traffic in sensor networks : Encryption, key distribution, and routing adaptation," *IEEE Transactions on Mobile Computing*, vol. 5, no. 10, pp. 1417–1431, 2006, member-Westhoff, Dirk and Student Member-Girao, Joao and Student Member-Acharya, Mithun.

[65] J. Domingo-Ferrer, "A provably secure additive and multiplicative privacy homomorphism," in *5th International Conference on Information Security (ISC)*, Sao Paulo, Brazil, September 2002.

[66] C. Castelluccia, A. C.-F. Chan, E. Mykletun, and G. Tsudik, "Efficient and provably secure aggregation of encrypted data in wireless sensor networks," *ACM Transactions Sensor Networks*, vol. 5, no. 3, pp. 1–36, 2009.

[67] P. Papadimitratos and Z. Haas, "Secure rotuing for mobile ad hoc networks," in *Communication Networks and Distributed Systems Modeling and Simulation Conference*, San Antonio, Texas, 2002.

[68] Y. C. Hu, A. Perrig, and D. B. Johnson, "Ariadne : a secure on-demand routing protocol for ad hoc networks," *Wireless Networks*, vol. 11, no. 1-2, pp. 21–38, January 2005.

[69] K. Sanzgiri, B.Dahill, B. N. Levine, C. Shields, and E. M. Belding-Royer, "A secure routing protocol for ad hoc networks," in *IEEE International Conference on Network Protocols*. Paris, France : IEEE Computer Society, November 2002.

[70] A. Perrig, R. Canetti, J. D. Tygar, and D. Song., "The tesla broadcast authentication protocol," *RSA CryptoBytes*, 2002.

[71] C. E. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (dsdv) for mobile computers," *ACM SIGCOMM Computer Communication Review*, vol. 24, no. 4, pp. 234–244, October 1994.

[72] ——, "Highly dynamic destination-sequenced distance-vector routing (dsdv) for mobile computers," *Comp. Commun. Rev.*, pp. 234–244, October 1994.

[73] M. Tubaishat, J. Yin, B. Panja, and S. Madria, "A secure hierarchical model for sensor network," *SIGMOD*, vol. 33, no. 1, pp. 7–13, 2004.

[74] L. Buttyán and J.-P. Hubaux, "Nuglets : a virtual currency to stimulate cooperation in self-organized mobile ad hoc networks," Ecole Polytechnique Fédérale de Lausanne, Tech. Rep., 2001.

[75] M. Menth, R. Martin, M. Hartmann, and U. Spörlein, "Efficiency of routing and resilience mechanisms in packet-switched communication networks," *European Transactions on Telecommunications*, vol. 21, pp. 108–120, March 2010.

[76] K.-H. Kim, W.-D. Jung, J.-S. Park, H.-G. Seo, S.-H. Jo, C.-M. Shin, S.-M. Park, and H.-N. Kim, "A resilient multipath routing protocol for wireless sensor networks," in *International Conference on Networking*, ser. Lecture Notes in Computer Science, P. Lorenz and P. Dini, Eds., vol. 3421. Ile de la Réunion : Springer, April 2005.

[77] T. Roosta, S. Pai, P. Chen, S. Sastry, and S. Wicker, "Inherent security of routing protocols in ad-hoc and sensor networks," in *IEEE Global Telecommunications Conference*. Washington, USA : IEEE, November 2007.

[78] A. D. Wood, L. Fang, J. A. Stankovic, and T. He, "Sigf : a family of configurable, secure routing protocols for wireless sensor networks," in *ACM Workshop on Security of ad hoc and Sensor Networks (SASN)*, ACM, Ed., VA, USA, October 2006.

[79] B. Blum, T. He, S. Son, and J. Stankovic, "Igf : A state-free robust communication protocol for wireless sensor networks," Technical report, Univ. of Virginia, Charlottesville, VA, USA, Tech. Rep. CS-2003-11, November 2003.

[80] "http ://wsnet.gforge.inria.fr/," July 2011.

[81] L. Lovasz, "Random walks on graphs : A survey," vol. 2, 1993.

[82] R. Motwani and P. Raghavan, "Markov chains and random walks," *Cambridge University Press*, vol. 6, pp. 127–160, 1995.

[83] L. Rodero-Merino, A. F. Anta, L. Lopez, and V. C. c, "Performance of random walks in one-hop replication networks," *Computer Networks*, pp. 781–796, 2010.

[84] P. G. Doyle and J. L. Snell, "Random walks and electric networks," *MAA*, 1984.

[85] C. Thomassen, "Resistences and currents in infinite electrical networks," *J. Comb. Theory*, pp. 87–102, 1990.

[86] R. S. Finch, "Pólya's random walk constant," *Mathematical Constants, Cambridge University Press*, pp. 322–331, 2003.

[87] D. Aldous and J. Fill, *Reversible Markov Chains and Random Walks on Graphs*. University of California, Berkeley, 2002.

[88] P. Diaconis, "Group representations in probability and statistics," *Inst. of Math. Statistics, Hayward, Californis*, 1988.

[89] J. Friedman, "On the second eigenvalue and random walks in randomd-regular graphs," *Combinatorica*, vol. 11, pp. 331–362, 1991.

[90] G. Slade, "The self-avoiding walk : A brief survey." in *Surveys in Stochastic Processes, In Proceedings of the Thirty-third SPA Conference*, Berlin, Germany, 2009.

[91] K. Altisen, S. Devismes, A. Gerbaud, and P. Lafourcade, "Analysis of random walks using tabu lists," in *In 19th International Colloquium of Structural Information and Communication Complexity (SIROCCO)*, Reykjavik, Iceland, July 2012.

[92] Y. Azar, A. Z. Broder, A. R. Karlin, N. Linial, and S. Phillips, "Biased random walks," vol. 16, no. 1, pp. 1–18, 1996.

[93] R. Beraldi, "Service siscovery in manet via biased random walks," *Autonomics*, 2007.

[94] R. Berald, R. Baldoni, and R. Prakash, "A biased random walk routing protocol for wireless sensor networks : The lukewarm potato protocol," *IEEE Transactions on Mobile Computing*, vol. 9, November 2010.

[95] P. Trimintzios, "Measurement frameworks and metrics for resilient networks and services : Technical report," European Network and Information Security Agency (ENISA), Tech. Rep., February 2011.

[96] F. Xing and W. Wang, "Analyzing resilience to node misbehaviors in wireless multihop networks," in *Wireless Communications and Networking Conference (WCNC)*, Hong-Kong, March 2007.

[97] D. J. Rosenkrantz, S. Goel, S. S. Ravi, and J. Gangolly, "Structure-based resilience metrics for service-oriented networks," *IEEE Transactions on Services Computing*, vol. 2, no. 3, pp. 183–196, 2009.

[98] A. Jabba, H. Narra, and J. P. Sterbenz, "An approach to quantifying resilience in mobile ad hoc networks," in *Design of Reliable Communication Networks (DRCN)*, Krakow, Poland, October 2011.

[99] E. Koller, *Encyclopedic Dictionary of the Materials Science*, Dunod, Ed., 2008.

[100] A. M. Madni and S. Jackson, "Toward a conceptual framework for resilience engineering," *IEEE Systems Journals*, vol. 3, no. 2, pp. 181–191, June 2009.

[101] T. R. Farley and C. J. Colbourn, "Multiterminal resilience for series-parallel networks," *Networks*, vol. 50, no. 2, pp. 164–172, September 2007.

[102] W. Najjar and J. Gaudiot, "Network resilience : A measure of network fault tolerance," *IEEE Transactions on Computers*, vol. 39, no. 2, pp. 174–181, February 1990.

[103] X. Li and D. Yang, "A quantitative survivability evaluation model for wireless sensor networks," in *IEEE International Conference on Networking, Sensing and Control*, Okayama, Japan, March 2006.

[104] H. Yang, Y. Yuan, S. Lu, and W. Arbaugh, "Toward resilient security in wireless sensor networks," in *MobiHoc*, Urbana-Champaign, Illinois, USA, May 2005.

[105] J. P. Sterbenz, D. Hutchison, E. K. Cetinkaya, A. Jabbar, J. P. Rohrer, M. Scholler, and P.Smith, "Resilience and survivability in communication networks : Strategies, principles, and survey of disciplines," *Computer Networks*, no. 54, pp. 1245–1265, March 2010.

[106] P. Cholda, A. Mykkeltveit, B. E. Helvik, O. J. Wittner, and A. Jajszczyk, "A survey of resilience differentiation frameworks in communication networks," *IEEE Communications Surveys & Tutorials*, vol. 9, no. 4, pp. 32–55, 2007.

[107] M. Lima, A. dos Santos, and G. Pujolle, "A survey of survivability in mobile ad hoc networks," *IEEE Communications Surveys and Tutorials*, vol. 11, no. 1, pp. 1–3, January 2009.

[108] A. Avizienis, J.-C. Laprie, B. Randell, and C. Landwehr, "Basic concepts and taxonomy of dependable and secure computing," *IEEE Transactions on Dependable and Secure Computing*, vol. 1, pp. 11–33, 2004.

[109] J. Laprie, "Dependable computing and fault tolerance : Concepts and terminology," *Proc. 15th IEEE Int. Symp. on Fault-Tolerant Computing*, 1985.

[110] R. Bourgeois, H. Chavel, and J. Kessler, *Memotech, the Materials Engineering*, Casteilla, Ed., 2001.

[111] Y. Deswarte and D. Powell, "Internet security : an intusion tolerance approach," in *Proceedings of the IEEE*, vol. 2, no. 94, New York, USA, February 2006.

[112] J. Lu and F. Valois, "Modélisation stochastique de réseaux radio," *Rapport de recherche INRIA RR-5518*, 2005.

[113] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion : a scalable and robust communication paradigm for sensor networks," in *In Proceedings of the 6th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, Boston, MA, August 2000.

[114] A. Brandt, J. Hui, R. Kelsey, P. Levis, K. Pister, R. Struik, J. Vasseur, and R. Alexander, "Rpl : Ipv6 routing protocol for low-power and lossy networks," March 2012. [Online]. Available : http://www.rfc-editor.org/rfc/rfc6550.txt

[115] B. D. Hugues, "Random walks and random environments," *Oxford University Press*, vol. 1, pp. 119–121, 1995.

[116] L. Lamport, R. Shostak, and M. Pease, "The byzantine generals problem," *ACM Trans. Programming Languages and Systems*, vol. 4, no. 3, pp. 382–401, July 1982.

[117] A. Kountouris, O. Erdene-Ochir, M. Minier, and F. Valois, "Methode resiliente a la presence des noeuds compromis pour la determination des routes par un protocole de routage dans un reseau," France Patent 1 158 828, 2011.

[118] D. Wagner, "Resilient aggregation in sensor networks," in *ACM Workshop on Security of Ad Hoc and Sensor Networks*, S. Setia and V. Swarup, Eds. Washington, USA : ACM, October 2004.

[119] J. Dong, R. Curtmola, and C. Nita-rotaru, "Secure network coding for wireless mesh networks : Threats, challenges, and directions," *Journal Computer Communications*, vol. 32, pp. 1790–1801, 2009.

[120] Y. Zhang, M. Minier, and W. Znaidi, *Security FOR Network Coding*, K. A. Agha, Ed. ISTE Ltd and Wiley, 2012.