UNIVERSIDADE FEDERAL DE SANTA CATARINA
CAMPUS FLORIANÓPOLIS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE AUTOMAÇÃO E SISTEMAS

Suelen Macedo Laurindo

Relay Selection Technique Applied to Wireless Sensor Network: The ORST Approach

Florianópolis
2021

Suelen Macedo Laurindo

Relay Selection Technique Applied to Wireless Sensor Network: The ORST Approach

Tese submetida ao Programa de Pós-Graduação em Engenharia de Automação e Sistemas da Universidade Federal de Santa Catarina para a obtenção do título de doutora em Engenharia de Automação e Sistemas.
Supervisor: Prof. Ricardo Moraes, Dr.
Co-supervisor: Prof. Carlos Barros Montez, Dr.

Florianópolis
2021

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Laurindo, Suelen
Relay Selection Technique Applied to Wireless Sensor Network: : The ORST Approach / Suelen Laurindo ; orientador, Ricardo Moraes, coorientador, Carlos Montez, 2021.
152 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia de Automação e Sistemas, Florianópolis, 2021.

Inclui referências.

1. Engenharia de Automação e Sistemas. 2. Rede de Sensores Sem Fio. 3. Seleção de cooperantes. 4. Comunicação Confiável. I. Moraes, Ricardo . II. Montez, Carlos. III. Universidade Federal de Santa Catarina. Programa de Pós Graduação em Engenharia de Automação e Sistemas. IV. Título.

Suelen Macedo Laurindo

Relay Selection Technique Applied to Wireless Sensor Network: The ORST Approach

O presente trabalho em nível de doutorado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Daniel Gouveia Costa, Dr.
Universidade Estadual de Feira de Santana

Prof. Érico Meneses Leão, Dr.
Universidade Federal do Piauí

Prof. Richard Demo de Souza, Dr.
Universidade Federal de Santa Catarina

Prof. Marcelo Ricardo Stemmer, Dr.
Universidade Federal de Santa Catarina

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de doutora em Engenharia de Automação e Sistemas.

ICP-Edu
Documento assinado digitalmente
Werner Kraus Junior
Data: 08/09/2021 23:08:06-0300
CPF: 531.085.239-53
Verifique as assinaturas em https://v.ufsc.br

Coordenação do Programa de
Pós-Graduação

ICP-Edu
Documento assinado digitalmente
Ricardo Alexandre Reinaldo de Moraes
Data: 09/09/2021 20:38:27-0300
CPF: 844.812.939-34
Verifique as assinaturas em https://v.ufsc.br

Ricardo Alexandre Reinaldo de Moraes
Orientador

Florianópolis, 2021.

Este trabalho é dedicado aos meus queridos pais.

AGRADECIMENTOS

A conclusão deste trabalho é um momento muito especial para mim. Tenho a consciência de que muitas pessoas foram fundamentais para que ele se concretizasse. Assim, eu gostaria de agradecer especialmente:

A Deus, por ter colocado pessoas maravilhosas ao meu lado nesta jornada e me capacitado para percorrê-la!

A Ricardo Moraes, professor na Universidade Federal de Santa Catarina, que muito além de me orientar academicamente ajudou-me a crescer como ser humano. Foram aprendizados incontáveis. Sempre disposto a ajudar com os problemas que surgiam, sempre com dicas e conversas valiosas. Muito obrigado por ser esse profissional incrível e inspirador!

A Carlos Montez, professor na Universidade Federal de Santa Catarina, pelo acompanhamento e pela excelente contribuição científica nos trabalhos realizados no âmbito desta tese.

A Francisco Vasques, professor na Universidade do Porto, ao qual sou grata por suas ideias e definições durante este trabalho. Recebeu-me com muito entusiasmo em Portugal e com certeza fez a diferença naquele período fora da minha terra natal.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), que se responsabilizaram pelo financiamento da minha bolsa e também por proporcionar um período de sete meses na Universidade do Porto em Portugal.

A minha família, que sempre me apoiou. Minha mãe Fátima, meu pai José e meus irmãos Lucas e Josimar, que sempre se esforçaram para entender minha pesquisa e por motivar-me em momentos difíceis.

A Luan Nagel Constante, por toda sua paciência e companhia. Pelas inúmeras conversas sobre os artigos científicos e simulações; por assistir incansavelmente meus ensaios de apresentações, por dedicar inúmeras vezes seus finais de semana ao meu lado nos estudos. E principalmente pelos momentos de alegria, que fizeram que com esta caminhada fosse leve.

Aos amigos que fiz nesta jornada, Renata Martins, Benedito Bitencourt, Renata Oliveira, Thiago Jesus e Rogério Casagrande, que permitiram compartilhar momentos de descontração e crescimento pessoal. Foram bons amigos desde o dia que cheguei em Portugal, me ensinaram muito e criaram um ambiente familiar, que foi fundamental para eu ter vivido meses incríveis.

Aos colegas do Laboratório de Automação e Robótica Móvel, em especial ao acadêmico Guilherme Pereira, pelas discussões técnicas, apoio nos experimentos e por sempre ser tão prestativo.

Aos amigos que mesmo fora da universidade fizeram suas colaborações neste trabalho, Patrícia Bordignon, por ser essa amiga mais que especial, que sempre trouxe momentos de descontração para minha jornada! Érica Dal Ponte, por toda sua ajuda nas gravações dos vídeos para as apresentações dos artigos e a Gerson budke, pelas diversas conversas técnicas que contribuíram para este trabalho.

Por fim, agradeço a Enio Snoeijer, colaborador da secretaria do Programa de Pós Graduação em Engenharia de Automação e Sistemas, por ser esse profissional prestativo e dedicado, que sempre me auxiliou com documentações e dúvidas que surgiram.

"Know all the theories,
master all the techniques,
but as you touch a human soul
be just another human soul."
(Jung, 1928)

RESUMO

Redes de Sensores Sem Fio (RSSF) estão cada vez mais presentes no dia a dia das pessoas (saúde, cidades inteligentes, agricultura inteligente, Indústria 4.0 e o uso sustentável de ecossistemas terrestres), facilitando o desenvolvimento das atividades diárias e melhorando a qualidade de vida. Porém, em todas essas aplicações, a confiabilidade da comunicação ainda é um desafio devido à natureza dos canais de comunicação sem fio, que estão sujeitos a ruídos eletromagnéticos e obstáculos entre os nós que podem atenuar ou refletir o sinal. Para minimizar este desafio, técnicas de diversidade cooperativa e mecanismos de retransmissão podem ser usados como estratégias eficientes. Nas técnicas de diversidade cooperativa, alguns nós são selecionados como cooperantes e são responsáveis por transmitir seus próprios dados e os dados previamente armazenados de outros nós. A seleção dos nós cooperantes é uma etapa crítica, que pode afetar a qualidade da comunicação e é importante encontrar os critérios de seleção mais adequados para garantir o funcionamento da rede. Além disso, mecanismos de retransmissão que são amplamente utilizados e que podem ser citados são técnicas de codificação de rede, onde nós intermediários atuam sobre as mensagens retransmitidas, modificando-as de acordo com técnicas específicas com codificação matemática. As técnicas de codificação de rede são soluções promissoras para minimizar o atraso de transmissão, equilibrar a carga da rede e também melhorar a taxa de sucesso da rede. Nesse contexto, esta tese pretende atuar sobre esta limitação de confiabilidade nas comunicações em RSSF, tratando a comunicação de forma holística. Como contribuição desta tese, destacamos (i) uma revisão sistemática sobre técnicas de seleção de nós cooperantes; (ii) a proposta e validação de uma nova técnica de seleção de nós cooperantes, que chamamos de ORST (Técnica de Seleção de Cooperantes Otimizada); (iii) uma análise dos parâmetros utilizados na seleção dos cooperantes da técnica ORST; (iv) uma análise das soluções utilizadas no problema de seleção de cooperantes da técnica ORST; (v) uma revisão sistemática sobre técnicas de codificação de rede; (vi) a proposta de uma técnica de codificação de rede em conjunto com a proposta de quatro mecanismos de retransmissão, sendo que três deles consideram a técnica de codificação de rede e um não utiliza codificação de rede; e por fim, (vii) uma análise do funcionamento de uma RSSF quando utilizada a técnica ORST e os mecanismos de retransmissão, com e sem codificação de rede. Para avaliar a técnica de seleção de cooperantes e os mecanismos de retransmissão foi utilizada a ferramenta de simulação de rede OMNeT++ e o framework Castalia. O pressuposto inicial desta tese era que a técnica proposta de seleção de cooperantes operando em conjunto com um mecanismo de retransmissão que utilizasse codificação de rede melhoraria a confiabilidade da comunicação das RSSF. No entanto, a avaliação da simulação mostrou que, ao usar a técnica de seleção de cooperantes, a retransmissão sem codificação de rede é a melhor solução.

Palavras-chave: Seleção de cooperantes. Codificação de rede. Rede de Sensores Sem Fio. Comunicação Confiável.

RESUMO EXPANDIDO

Introdução

Uma Rede de Sensores Sem Fio (RSSF) é uma rede composta de dezenas a milhares de nós sensores, que são implantados em um ambiente com recursos de detecção, coleta de dados e comunicações sem fio. Desde a origem das RSSF, em meados dos anos 2000, uma grande evolução tecnológica ocorreu e a tecnologia de semicondutores continua a seguir a lei de Moore, fornecendo dispositivos sem fio menores, com maior poder computacional e com um menor custo, o que contribui para a popularização das RSSF nas mais diversas aplicações (FAHMY, 2021b).

As RSSF podem ser encontradas em diversas áreas, por exemplo, na agricultura, na biomedicina, na indústria, no meio ambiente, nas residências, na saúde e no controle de tráfego (SRIVASTAVA; MISHRA, 2021). Cada tipo de aplicação pode apresentar alguns requisitos específicos. Um exemplo ocorre na indústria, onde o bom funcionamento do sistema depende de medições periódicas, e os dados de monitoramento de alguns sensores são enviados ao controlador em intervalos de tempo restritos. Portanto, as especificações de tempo desses sistemas podem ser mais rígidas. Já em uma aplicação na agricultura, sensores que realizam o monitoramento do solo, por exemplo, não precisam fazer o monitoramento em intervalos de tempo tão restritos. Nesse tipo de aplicação a especificação de tempo não é um requisito crítico (POLAVARAPU; PANDA, 2020).

Desta forma, considera-se que as RSSF são orientadas a aplicações, onde cada aplicação necessita de características diferentes em sua RSSF. O ambiente em que uma RSSF será implantada permitirá determinar o tamanho da rede, a forma como os nós serão implantados, e a topologia da rede. As RSSF podem ser implantadas em diferentes ambientes, como: ambiente terrestre, subterrâneo, subaquático e no próprio corpo humano. Por exemplo, em um ambiente terrestre interno, um número menor de nós pode ser necessário, já que o ambiente é limitado, enquanto ambientes externos podem exigir uma grande quantidade de nós para cobrir uma área maior. Cada aplicação fornece diretrizes que podem levar à construção de uma RSSF ideal, que satisfaça os requisitos da aplicação e atenda as limitações da rede sem fio (FAHMY, 2021a).

No entanto, para todas as aplicações em RSSF, a confiabilidade da comunicação ainda é um desafio devido à natureza dos canais de comunicação sem fio, que estão sujeitos a ruídos eletromagnéticos e obstáculos entre os nós que podem atenuar ou refletir o sinal. Para minimizar esse problema, dois tipos de técnicas de comunicação podem ser usados: técnicas de diversidade cooperativa e mecanismos de retransmissão. Técnicas de diversidade cooperativa envolvem cooperação entre nós para melhorar a taxa de sucesso de mensagens enviadas (LAURINDO, S. et al., 2017); e mecanismos retransmissão permitem determinar como ocorrerá a etapa de retransmissão, como mecanismos de retransmissão podem ser citadas as técnicas de codificação de rede, que permitem que os nós cooperantes combinem várias mensagens e retransmitam esta combinação em uma única mensagem.

Técnicas de diversidade cooperativa utilizam um ou vários nós para atuar como cooperantes na RSSF. Esses nós exploram a natureza das transmissões sem fio, ouvindo e armazenando mensagens enviadas por seus vizinhos, de modo que podem retransmitir com êxito as mensagens ouvidas para o nó de destino. Assim, as mensagens que não foram recebidas por transmissão direta podem ser recebidas por retransmissão (HIMANSHU et al., 2015; SONKAR et al., 2016). Este tipo de comunicação, além de promover melhorias na diversidade espacial e temporal, facilita a melhoria da taxa de sucesso das mensagens enviadas sem aumentar a complexidade do hardware. Porém, sempre que utilizar diversidade cooperativa, um problema importante a ser abordado é a seleção do conjunto de nós cooperantes. O desempenho de toda a rede pode ser melhorado se os nós de cooperantes forem selecionados de forma otimizada. Portanto, um tópico de pesquisa relevante em redes de sensores sem fio é o uso de técnicas otimizadas de seleção de cooperantes (GUO, Q.; LI, Xin, 2017).

Técnicas de codificação de rede têm se mostrado uma solução promissora para minimizar o atraso de transmissão, para equilibrar a carga da rede e também para melhorar o rendimento da rede. Além disso, as características de transmissão das redes sem fio permitem que a codificação da rede forneça melhores resultados quando usadas em aplicações de retransmissão cooperativa, pois melhora a eficiência do espectro e a capacidade do sistema (HO, Tracey; LUN, 2008; FRAGOULI; SOLJANIN, et al., 2007). No entanto, ao usar a codificação de rede em RSSF, é necessário levar em consideração a metodologia utilizada para codificar e decodificar as mensagens. Na codificação de rede linear, as mensagens são combinadas linearmente usando coeficientes de codificação. É necessário selecionar coeficientes de codificação para cada mensagem que será codificada e o nó de destino precisa saber quais coeficientes foram usados. O estado da arte cita que esses coeficientes são enviados junto com a mensagem codificada ou em uma mensagem extra, que gera sobrecarga. Portanto, o tópico de pesquisa relevante em codificação de rede é o aprimoramento dos métodos de envio de coeficientes de codificação (GUO, B. et al., 2014; VALLE, Odilson T et al., 2016).

Objetivo

Esta tese tem como objetivo contribuir para o avanço do estado da arte em comunicação RSSF, propondo, desenvolvendo e avaliando técnicas baseadas na diversidade cooperativa e em mecanismos de retransmissão que utilizem codificação de rede. Seu principal objetivo é demonstrar que técnicas de seleção de cooperantes, usando critérios relevantes para a operação da rede, juntamente com mecanismos de retransmissão que considerem codificação de rede, podem aumentar a confiabilidade da comunicação em RSSFs implementadas no padrão IEEE 802.15.4e.

Metodologia

Esta tese caracteriza-se como uma pesquisa de abordagem mista (qualitativa - quantitativa), pois combina o estudo bibliográfico sobre as técnicas de seleção de cooperantes e codificação de redes com a análise numérica dos resultados obtidos pela realização de simulações. No que se refere ao domínio das ciências, este trabalho se enquadra como uma pesquisa prática, uma vez que busca solucionar problemas práticos de comunicação em RSSF. Em relação aos objetivos da pesquisa, este trabalho caracteriza-se como uma pesquisa exploratória e experimental. Pesquisa exploratória porque realiza uma revisão sistemática da literatura com foco no objeto de estudo. E experimental, porque avaliará os resultados quantitativos produzidos pela simulação (FREIRE, 2013). No que diz respeito aos procedimentos, como em qualquer pesquisa acadêmica, este trabalho se caracteriza como uma pesquisa bibliográfica. É também uma pesquisa experimental, pois experimentos virtuais serão realizados para analisar e avaliar as técnicas propostas em cenários semelhantes a um ambiente real.

Resultados e Discussões

Os resultados apresentados nesta tese de doutorado podem ser divididos em (i) uma revisão sistemática sobre técnicas de seleção de nós cooperantes; (ii) proposta e validação de uma técnica de seleção de nós cooperantes, que chamamos de ORST (Técnica de Seleção de Cooperantes Otimizada); (iii) uma revisão sistemática sobre técnicas de codificação de rede; (iv) proposta de uma técnica de codificação de rede em conjunto com a proposta de mecanismos de retransmissão; e por fim, (v) uma análise do funcionamento de uma RSSF quando utilizado a técnica ORST e os mecanismos de retransmissão, com e sem codificação de rede.

O item (i) foi inicializado no Capítulo 2, onde apresenta trabalhos do estado da arte sobre técnicas de seleção de nós cooperantes entre os anos de 2011 e 2017, e complementado nos Capítulos 3 e 5 com trabalhos entre os anos 2017 e 2021. Como resultado, esta tese de doutorado apresenta uma síntese dos trabalhos do estado da arte dos últimos dez anos. Sobre o item (ii), o Capítulo 2 desta tese apresenta o desenvolvimento da técnica ORST, a qual é formulada como um problema de otimização e que considera dois esquemas de atualização de nós cooperantes, PRS (Seleção de Cooperantes Periódica) e ARS (Seleção de Cooperantes Adaptativa). Os Capítulos 3 e 4 apresentam estudos complementares que melhoram a técnica ORST. No Capítulo 3, os parâmetros utilizados na função objetivo do problema de otimização foram analisados, o que permitiu otimizar a função objetivo, reduzindo o número de critérios considerados e simplificando o problema de seleção de cooperantes. No Capítulo 4, a solução utilizada para realizar a seleção de cooperantes foi analisada. ORST é uma técnica de seleção de cooperantes que pode ser reduzida à aplicação no problema clássico de cobertura de conjuntos (SCP) para RSSF. Desta forma, mais de uma solução pode ser utilizada para resolver o problema de seleção de cooperantes, neste estudo foi possível determinar a solução que garante que os melhores cooperantes sejam selecionados e pode ser implementada em nós sensores de baixo custo.

O item (iii) é apresentado no Capítulo 5, onde trabalhos do estado da arte sobre codificação de rede, entre os anos 2005 e 2021, são apresentados e comentados. Além disso também é proposta uma estrutura que classifica as diferentes formas de realizar a codificação de rede. Sobre o item (iv), o Capítulo 5 desta tese de doutorado apresenta a proposta de uma técnica de codificação de rede, a qual utiliza uma equação, previamente definida entre o nó destino que os nós intermediários, como regra para formar os coeficientes. Os coeficientes são enviados para o destino com base em uma representação de mapa de bits, o que induz a uma redução na sobrecarga gerada pelo envio dos coeficientes. Além da técnica, quatro mecanismos de retransmissão foram propostos, sendo que três deles consideram a técnica de codificação de rede e um não utiliza codificação de rede. O item (v) é apresentado no Capítulo 5, onde a técnica ORST e a técnica de codificação de rede juntamente com os diferentes mecanismos de retransmissão são avaliados. Diferentemente do que foi assumido ao iniciar esta avaliação, a técnica ORST trabalhando em conjunto com abordagens de codificação de rede não apresentou resultados interessantes. No entanto, a eficiência da técnica ORST é demonstrada quando os nós cooperantes retransmitem um conjunto de mensagens ouvidas em *slots* individuais sem usar codificação de rede; a taxa de sucesso é maior do que 90% e o consumo de energia está apenas ligeiramente acima de um cenário que não utiliza diversidade cooperativa.

Considerações Finais

As RSSFs são muito importantes no mundo moderno, pois permitem a interconexão de dispositivos computacionais sem a necessidade de um cabo para intermediar a comunicação. Apesar disso, essas redes apresentam uma grande limitação quanto a confiabilidade das comunicações. O que torna a utilização destas redes em comunicações críticas desafiador. Esta tese de doutorado tratou sobre esse problema, gerando cinco publicações científicas (até a data de escrita deste documento). Os resultados obtidos nesta pesquisa, permitem aumentar a confiabilidade das comunicações em RSSF e ainda permitem o desenvolvimento de trabalhos futuros que incluam outras topologias de rede.

ABSTRACT

Wireless sensor networks (WSNs) are increasingly prevalent in everyday life (for example, in healthcare, smart cities, smart agriculture, Industry 4.0, and the sustainable use of terrestrial ecosystems), as they can facilitate daily activities and improve quality of life. However, in all these applications, the reliability of communication remains a challenge due to the nature of wireless communication channels, which are subject to electromagnetic noise and obstacles between nodes that can attenuate or reflect the signal. To minimize these problems, cooperative diversity techniques and retransmission mechanisms are efficient strategies that can be applied. In cooperative diversity techniques, several nodes are selected as relays and are responsible for transmitting both their own data and data from other nodes that have been previously stored. Relay selection is a critical step that may affect the quality of transmission, and it is therefore important to find the most appropriate selection criteria for the operation of the network. In addition, network coding techniques are widely used retransmission mechanisms in which intermediary nodes act upon the retransmitted messages and modify them using specific mathematical coding techniques. Network coding techniques are a promising solution for minimizing transmission delay, balancing the network load, and improving the network throughput. In this context, this thesis examines the limitations on reliable communication in WSNs by treating the communication in a holistic way. As a contribution in this thesis, we highlight: (i) a systematic review on relay selection techniques; (ii) the proposal and validation of a new technique for selecting relay nodes, which we call ORST (Optimized Relay Selection Technique); (iii) an analysis of the parameters used in the selection of relay nodes in the ORST technique; (iv) an analysis of the solutions used in the relay selection problem of the ORST technique; (v) a systematic review of network coding techniques; (vi) the proposal of a network coding technique together with the proposal of four retransmission mechanisms, three of which consider the network coding technique and one do not use network coding; and finally, (vii) an analysis of the operation of a WSN when using the ORST technique and retransmission mechanisms, with and without network coding. To evaluate the cooperative selection technique and the retransmission mechanisms, the network simulation tool OMNeT++ and the WSN framework Castalia were used. The initial assumption underlying this research work was that the use of the proposed relay selection technique in conjunction with a retransmission mechanism that used network coding would improve the communication reliability of a WNS; however, the simulation results showed that when the proposed relay selection technique was used, retransmission without network coding was a better solution.

Keywords: Relay Selection. Network Coding. Wireless Sensor Network. Communication Reliability.

LIST OF FIGURES

Figure 1 – WSNs Application. 24
Figure 2 – Operation of cooperative retransmission with network coding. 26
Figure 3 – Types of scientific research. 31
Figure 4 – Wireless sensor network that needs a relay. 36
Figure 5 – Relay selection categories based on the criteria used to select the cooperating nodes. 38
Figure 6 – Beacon frame format. 47
Figure 7 – Data frame format. 47
Figure 8 – Transmission steps. 48
Figure 9 – Retransmission steps. 49
Figure 10 – Network with the neighborhood of each node. 52
Figure 11 – Operation of the periodic relay selection scheme. 53
Figure 12 – Operation of the adaptive relay selection scheme. 54
Figure 13 – Success Rate—static topology. 59
Figure 14 – Success Rate— dynamic topology. 60
Figure 15 – Energy consumption—static topology. 60
Figure 16 – Energy consumption—dynamic topology. 61
Figure 17 – Average Number of Relay Selection—static topology. 61
Figure 18 – Average Number of Relay Selection—dynamic topology. 62
Figure 19 – Average number of cooperation exchanges for a dynamic scenario. . 62
Figure 20 – Percentage of useless retransmission messages for a dynamic scenario. 63
Figure 21 – Correlation between energy consumption and useless retransmission messages. 63
Figure 22 – Cooperative communication. 66
Figure 23 – Success Rate. 74
Figure 24 – Energy Consumption. 75
Figure 25 – Average number of cooperation exchanges. 75
Figure 26 – Percentage of useless retransmission messages. 76
Figure 27 – Success Rate compared to RAC technique. 77
Figure 28 – Energy Consumption compared to RAC technique. 77
Figure 29 – Average number of cooperation exchanges compared to RAC technique. 78
Figure 30 – Percentage of useless retransmission messages compared to RAC technique. 78
Figure 31 – WSN using a cooperative communication. 82
Figure 32 – Correlation between success rate and time used by relay selection. . 93

Figure 33 – Useless retransmission messages. 94
Figure 34 – Correlation between Average number of cooperation exchanges and energy consumption. 94
Figure 35 – Communication over a wireless network: (**a**)—without NC and (**b**)—with NC. 98
Figure 36 – Communication over a wireless sensor network considering cooperative diversity and NC. 99
Figure 37 – Network coding categories. 101
Figure 38 – Network coding on the physical layer. 104
Figure 39 – Node N_2 performing the network coding and creating the bit-map of the coding coefficients. 116
Figure 40 – Energy consumption. 120
Figure 41 – Success rate. 120
Figure 42 – Correlation between the average number of retransmitted messages per node and the average number of recovered messages in the decoding process. 121
Figure 43 – Average number of relay nodes. 122
Figure 44 – Success rate. 125
Figure 45 – Energy consumption. 126

LIST OF TABLES

Table 1 – Relay selection techniques. 45
Table 2 – Simulation setting. 56
Table 3 – Objective Functions. 72
Table 4 – Simulation Setting. 73
Table 5 – Genetic Algorithm Setting. 90
Table 6 – Simulation Setting. 91
Table 7 – Network coding techniques. 111
Table 8 – Simulation Setting. 117

LIST OF ABBREVIATIONS AND ACRONYMS

ACK	Acknowledgments
AF	Amplify-Forward
B&*B*	Branch and Bound
ANCC	Adaptive Network Coded Cooperation
ARNS	All Relay Nodes Selection
ARQ	Automatic Repeat Request
ARR	Available Relay Selection
ARS	Adaptive Relay Selection
BCSO	Binary Cat Swarm Optimization
BER	Bit Error Rate
BI	Beacon Interval
BIP	Binary Integer Problem
BRNS	Best Relay Node Selection
BTS	Batch Transmission Size
CAP	Contention Access Period
CAR	Candidate Receivers
CFP	Collision Free Period
COTS	Commercial Off-The-Shelf
CPCM	Coded Packet Coefficients Matrix
CR	Completely Random
CSI	Channel State Information
CTS	Clear-To-Send
CV	Coding Vector
DF	Decode-Forward
DFNC	Decode-and-Forward Network Coding
DH	Descent Heuristic
DLNC	Deterministic Linear Network Coding
DSN	Data Sequence Numbers
EPB	Energy-Per-Bit
FEC	Forward Error Correction
FER	Frame Error Rate
GA	Genetic Algorithm
GACK	Group Acknowledgments
GBR	Gradient-Based Routing
GF	Galois Field
GTS	Guaranteed Time Slot
IEEE	Institute of Electrical and Electronics Engineers
IIoT	Industrial Internet of Things

ITH	Interested-to-Help
LAR	List and Remove
LDPC	Low-Density-Parity-Check
LLDN	Low Latency Deterministic Networks
LQI	Link Quality Indicator
MAC	Medium Access Control
MDP	Markov Decision Process
MIMO	Multiple-Input-Multiple-Output
NACK	Negative-Acknowledgment
NC	Network Coding
ORST	Optimized Relay Selection Technique
PAN	Personal Area Network
PRR	Packet Reception Ratio
PRS	Periodic Relay Selection
QR	Relay Receivers
RAC	Random Around the Coordinator
REI	Residual Energy Information
RLNC	Random Linear Network Coding
RPCM	Redundant Packet Coefficients Matrix
RRNS	Random Relay Node Selection
RSSI	Received Signal Strength Indicator
RTH	Request-To-Help
RTS	Request-To-Send
SCP	Set-Covering Problem
SD	Superframe Duration
SEQ	Sequence Number
SLNC	Sparse Linear Network Coding
SNR	Signal-to-Noise Ratio
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
WSNs	Wireless Sensor Networks
WVEFC	Weighted Vandermonde Echelon Fast Coding

CONTENTS

1 INTRODUCTION . . . 23
1.1 RESEARCH CONTEXT . . . 27
1.2 RESEARCH OBJECTIVES . . . 29
1.2.1 Main Objective . . . 29
1.2.2 Specific Objectives . . . 29
1.3 RESEARCH CONTRIBUTIONS . . . 29
1.4 METHODOLOGY . . . 30
1.5 THESIS OUTLINE . . . 31
2 A NEW RELAY SELECTION TECHNIQUE AND ITS APPLICATION TO A WSN . . . 33
2.1 INTRODUCTION . . . 34
2.2 RELATED WORK . . . 37
2.2.1 Link Quality Based . . . 38
2.2.2 Link Quality and Neighborhood Based . . . 41
2.2.3 Link Quality and Energy Based . . . 42
2.2.4 Link Quality and Data Rate Based . . . 43
2.2.5 Random Relay Selection . . . 44
2.2.6 Wrap-Up . . . 44
2.3 ORST-OPTIMIZED RELAY SELECTION TECHNIQUE . . . 46
2.3.1 System Model . . . 46
2.3.2 Brief Explanation of the Relay Selection Operation . . . 47
2.3.3 Problem Formulation . . . 49
2.3.4 Relay Updating Schemes . . . 53
2.4 SIMULATION ASSESSMENT . . . 55
2.4.1 Characteristics of the Model Implementation . . . 55
2.4.2 Simulation Settings . . . 55
2.4.3 Compared Techniques . . . 56
2.4.4 Simulation Assessment . . . 58
2.5 CONCLUSIONS . . . 63
3 IMPACT OF PARAMETERS ON THE OBJECTIVE FUNCTION . . . 65
3.1 INTRODUCTION . . . 66
3.2 RELATED WORK . . . 67
3.3 RELAY SELECTION TECHNIQUE . . . 69
3.3.1 Analysis of Criteria for the Relay Selection Technique . . . 71
3.4 SIMULATION ASSESSMENT . . . 72
3.4.1 Simulation Settings . . . 72
3.4.2 Simulation Assessment . . . 73

3.5 CONCLUSION . . . 78
4 AN ANALYSIS OF DIFFERENT SOLUTIONS TO SOLVE THE RELAY SELECTION PROBLEM . . . 80
4.1 INTRODUCTION . . . 81
4.2 RELATED WORKS . . . 83
4.3 THE ORST TECHNIQUE AND PROBLEM FORMULATION . . . 84
4.4 ALGORITHMS USED TO SOLVE THE RELAY SELECTION PROBLEM IN THE ORST TECHNIQUE . . . 87
4.4.1 Branch-and-Bound Algorithm - *B&B* **. . . 87**
4.4.2 Greedy Algorithm . . . 88
4.4.3 Genetic Algorithm . . . 89
4.5 SIMULATION ASSESSMENT AND DISCUSSIONS OF RESULTS . . . 91
4.5.1 Simulation Results . . . 92
4.6 CONCLUSION . . . 95
5 RETRANSMISSION MECHANISMS COMBINED WITH THE ORST TECHNIQUE . . . 96
5.1 INTRODUCTION . . . 97
5.2 RELATED WORK . . . 100
5.2.1 Relay Selection Related Work . . . 101
5.2.2 Brief Explanation of the Main Types of Network Coding . . . 103
5.2.3 Network Coding Related Work . . . 106
5.2.4 Wrap-Up . . . 110
5.3 WSN COMMUNICATION . . . 111
5.3.1 System Model . . . 111
5.3.2 The Optimized Relay Selection Technique . . . 112
5.3.3 Network Coding Technique . . . 115
5.4 SIMULATION ASSESSMENTS . . . 117
5.4.1 Simulation Settings . . . 117
5.4.2 Network Coding Technique Application Scenarios . . . 118
5.4.3 Simulation Assessment . . . 119
5.5 CONCLUSIONS . . . 126
6 CONCLUSIONS . . . 128
6.1 REVIEW OF MOTIVATIONS AND OBJECTIVES . . . 128
6.2 THESIS OVERVIEW . . . 128
6.3 LIMITATIONS OF THE RESEARCH . . . 130
6.4 FUTURE WORK . . . 130
REFERENCES . . . 133
APPENDIX A – LIST OF PUBLICATIONS . . . 151
A.1 JOURNAL PUBLICATIONS . . . 151

A.2 CONFERENCE PUBLICATIONS 151

1 INTRODUCTION

Wireless Sensor Networks (WSNs) are composed of nodes that contain tiny, low-power sensors, processors, memory, power supplies (usually batteries) and radios (YICK et al., 2008). These WSN nodes link the physical and digital worlds by sensing, capturing and sending the collected data via wireless links to a predetermined destination (BURATTI et al., 2011), thus integrating machines, humans and/or environments into the cyber-physical world (ZHANG, C.; CHEN, Yong, 2020).

Cyber-physical networks emerged in the late 1990s, and were consolidated after 2003, when the first version of the Institute of Electrical and Electronics Engineers (IEEE) 802.15.4 communication network standard was released. This standard put forward specifications for the physical layer and the medium access control in wireless networks (IEEE COMPUTER, 2015). The IEEE 802.15.4 standard was designed to provide wireless connectivity at low data rates, with low complexity, low cost and low energy consumption, as these are important features of WSNs, and was adopted as a *de facto* standard for WSNs (IEEE COMPUTER, 2015).

The use of WSNs was boosted by the technological development of electronics based on smaller, cheaper components that could exchange messages both among themselves and with a coordinator via radio frequency communication (AKYILDIZ et al., 2002; ALEMDAR; IBNKAHLA, 2018; FAHMY, 2021b). In recent years, WSNs have become popular in everyday life due to the widespread development of the Internet of Things (IoT). A WSN can be adopted as a connectivity solution for several types of applications, ranging from control of the simplest types of system to those operating in hostile environments (DARGIE; POELLABAUER, 2010; YETGIN et al., 2017; TOMIĆ; MCCANN, 2017).

A WSN may contain several different types of sensors, such as humidity, temperature, seismic, radar, acoustic and magnetic sensors. Considering that a WSN is application-oriented and performs a specific task (KUORILEHTO et al., 2005), it is necessary to define the requirements for the WSN based on the network application. Figure 1 illustrates examples of solutions based on a WSN.

The applications of WSNs mainly include intelligent transportation systems, smart homes, military and industrial applications, precision agriculture, healthcare monitoring and environmental monitoring.

Intelligent transportation systems are used to improve efficiency and safety on the roads. A WSN can monitor road conditions and report information to drivers, thus making it possible to detect situations in which the driver needs to pay extra attention, such as icy or flooded roads (ZHANG, W. et al., 2019).

In smart homes, WSNs are applied in monitoring and control systems such as lighting, air conditioning and appliances. This type of control system allows for

Figure 1 – WSNs Application.

Source – Author.

improvements in energy efficiency, as it can closely monitor the environment (light, temperature, etc.). When the residence is empty, the lights and air conditioning can be switched off, and can be switched on again at a time close to the arrival of the residents. This can offer greater control and comfort in daily life (GAIKWAD et al., 2015).

In military applications, WSNs can be used in intrusion detection systems, where sensor nodes scan the environment and send alerts to troops. In addition, borders can be easily monitored using sensors to obtain information about enemy activity within the monitored area, thus allowing for a rapid response (AHMAD et al., 2016; ALEMDAR; IBNKAHLA, 2018).

WSNs are also employed to monitor and control production processes in industry, such as pressure, humidity, temperature, flows, levels, densities and vibration intensities. In addition, specific equipment can be used to monitor and detect possible problems at an early stage, to help prevent catastrophic failure (ZHAO, 2011).

In precision agriculture, WSNs are widely utilized in meteorological monitoring to measure parameters such as temperature, humidity, and radiation, which are necessary for management of a harvest. WSNs can also be used as actuators in irrigation applications, and can make the irrigation process more efficient, thus increasing the yield of food production (BARCELO-ORDINAS et al., 2013).

Healthcare is another field of application of WSNs, where information is periodically captured about the vital signals of patients and sent to a medical professional to carry out online monitoring of several patients. There are also healthcare services that provide special care to the elderly, thus allowing their health to be monitored without causing them to lose their independence; this type of monitoring system can permit the patient greater mobility, due to the fact that it is based on a wireless technology (AMMARI et al., 2015; YETGIN et al., 2017; SADIKU et al., 2018).

Finally, WSNs sensors can be employed to monitor environmental areas, for example to track animal behavior and the movements of a particular animal species within a natural environment, with minimal human interference. An analysis of the collected data may allow researchers to find new ways of managing the populations under study, for example, to determine the impacts of human development on these populations, or to understand whether there are enough individuals of a species in a given area to ensure survival. Another application is forest monitoring, which permits the efficient detection of forest fires and the avoidance of major catastrophes (BOUABDELLAH et al., 2013).

Although many of these applications have different operating requirements, they all need energy resources and reliable communication for proper operation. These are the two of the most important restrictions on WSNs (DARGIE; POELLABAUER, 2010).

The energy consumption of the nodes is an acute problem; it is not always feasible to replace batteries, since WSN nodes may be located in areas that are difficult to access. In view of this, it is necessary to develop use cases in which not all of the nodes are active at the same time (by imposing 'sleep' times during which nodes turn off radio communication), hence reducing the power consumption and increasing the lifetime of the overall network (GAO et al., 2011; LAURINDO, S. M. et al., 2016).

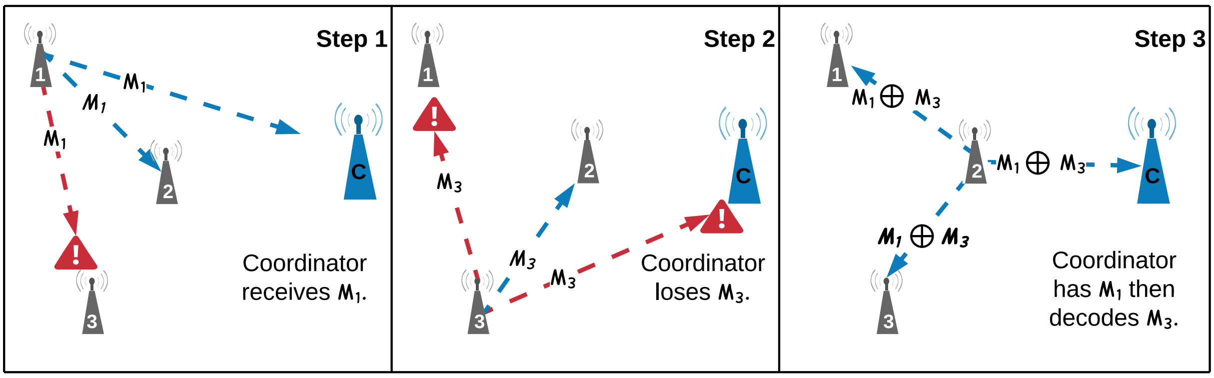
The problem of low reliability associated with message transfer arises from the characteristics of the wireless channel. Electromagnetic noise and obstacles between nodes can attenuate or reflect the signal, thus impairing the arrival of messages at their destinations (VALLE, Odilson T et al., 2016). To minimize these problems, two types of communication techniques can be used: cooperative diversity and retransmission mechanisms. Cooperative diversity techniques exploit cooperation between nodes to improve the success rates of sent messages (LAURINDO, S. et al., 2017), whereas retransmission mechanisms exploit how will occur the retransmission step. Between the retransmission mechanism, the Network Coding (NC) techniques have been shown to be a promising solution for minimizing transmission delay, balancing the network load and improving the network throughput. The broadcast characteristics of wireless networks enable network coding to provide better results when used in cooperative retransmission applications, as this can improve both the spectrum efficiency and system capacity (HO, Tracey; LUN, 2008; FRAGOULI; SOLJANIN, et al., 2007).

Cooperative retransmission is a technique in which one or more nodes act as relays in the WSN. These nodes exploit the broadcast nature of wireless transmissions by listening to and storing the messages sent by their neighbors, so they can retransmit successfully heard messages to the destination node. In this way, messages that were not received via direct transmission can be received via retransmission (HIMANSHU et al., 2015; SONKAR et al., 2016). This type of communication, in addition to creating improvements in spatial and temporal diversity, can improve the success rate of sent messages without increasing the complexity of the hardware. In this way, the relay nodes can form a Multiple-Input-Multiple-Output (MIMO) system, as multiple antenna systems are not a viable option in low-cost nodes due to their high energy consumption (WANG, C. L.; SYUE, 2009; IMAM et al., 2017).

Network coding is a technique in which nodes that have heard messages from their neighbors can generate a new data packet using an algebraic combination. Each node constructs a new data packet by applying an encoding operation to successfully heard packets (FRAGOULI; SOLJANIN, 2006; SUDHA et al., 2016). The use of network coding allows for the optimization of transmission rates, and can also improve throughput, since more information (a larger number of packets) is transmitted using a smaller number of transmissions (HO, Tracey; LUN, 2008).

Cooperative retransmission and network coding techniques can be used together. Figure 2 illustrates the principle of operation of a WSN that combines both techniques in a three-step scenario.

Figure 2 – Operation of cooperative retransmission with network coding.



Source – Author.

In this scenario, $Node_1$ and $Node_3$ want to send data to a coordinator node. In Step 1, $Node_1$ broadcasts a message m_1 that is correctly received by the coordinator and is also heard by $Node_2$. In Step 2, $Node_3$ broadcasts the message m_3, which is not received by the coordinator but is heard by $Node_2$. In Step 3, $Node_2$, which has stored both m_1 and m_3, acts as a relay and encodes both messages, transmitting them as a single encoded message. The coordinator recovers the original message m_3, by performing the related decoding operation (VALLE, Odilson T et al., 2016).

This example illustrates how the use of cooperative retransmission in conjunction with network coding can improve the network success rate by enabling the destination node to retrieve messages that were not received via direct transmission. In addition, network coding allows multiple messages to be packaged into a single data packet and sent to the destination. However, even with so many benefits, when using NC, it is necessary to consider that, in some scenarios, this type of technique can present a high computational cost to encode and decode messages. Furthermore, some scenarios may not achieve the benefits mentioned. For example, in the scenario presented in Figure 2, if the Coordinator had not received the m_1 message, it would be disadvantageous to send a coded message, because it would not be able to decode any message.

In addition, when cooperative diversity techniques are used, an important problem arises in regard to the selection of the set of cooperating nodes. The performance of the overall network can be improved if the relay nodes are optimally selected. An important research topic in the area of WSNs is therefore the optimization of relay selection techniques (GUO, Q.; LI, Xin, 2017).

When network coding is used, the methodology used to encode and decode the messages must be taken into account. In linear methods of network coding, messages are linearly combined using encoding coefficients. In this case, it is necessary to select encoding coefficients for each message to be encoded, and the destination node then needs to know which coefficients were used. In most state-of-the-art schemes (MIGABO et al., 2017; HEIDE et al., 2011; AKHTARI et al., 2020; WU, H. et al., 2019; LI, Ye et al., 2018; DONG et al., 2019; XU, C. et al., 2017; HAN, C. et al., 2017), these coefficients are sent alongside the encoded message or in an extra message, which generates overhead. An important research topic in the area of network coding is therefore the search for better methods for sending coding coefficients (GUO, B. et al., 2014; VALLE, Odilson T et al., 2016).

1.1 RESEARCH CONTEXT

In a cooperative diversity technique, relay nodes transmit both their own data and data previously stored from other nodes (SONKAR et al., 2016). The selection of an appropriate relay set is also a highly relevant issue in terms of improving the performance of the overall network communication (GUO, Q.; LI, Xin, 2017). An optimal solution will not treat all neighboring nodes as relays, and the random selection of a set of relay nodes is also a poor solution. These types of solution may be inefficient due to problems related to high bandwidth usage, increased energy consumption, and synchronization problems among network nodes. There is therefore a need to optimize the selection of relays based on specific and appropriate criteria.

In view of these problems, and since relay selection is a critical step that may affect the quality of transmission, it is important to consider the most suitable relay selection criteria for the operation of the network (GUO, Q.; LI, Xin, 2017). The majority of state-of-the-art techniques consider only quality estimators, such as the Received Signal Strength Indicator (RSSI), Channel State Information (CSI), Signal-to-Noise Ratio (SNR) or Link Quality Indicator (LQI) when selecting relay nodes (LI, Yubo et al., 2011; IKKI; AHMED, M. H., 2009; SUN et al., 2009; MARCHENKO et al., 2014; WANG, H. et al., 2010).

According to Baccour et al. (2010), quality estimators that use hardware measurements such as LQI, RSSI or SNR are based on a sample of just the first eight symbols of a received packet, meaning that the use of quality estimators alone may lead to an inaccurate selection of the set of relays. Although hardware metrics may provide a quick way of classifying a link as good or bad, they are unable to provide an accurate estimate of its quality, and must therefore be considered in combination with other metrics (BACCOUR et al., 2010).

In addition to the relay selection criteria, the updating techniques used can also influence the performance of the network. Updating of the relay set may be periodic, adaptive or reactive (MARCHENKO et al., 2014). In a periodic approach, a selection is made at regular intervals, regardless of the requirements of the network. An adaptive approach considers the performance of the current relay set, and will only update the selection if this falls below a predefined threshold. Finally, in a reactive approach, relay selection is only carried out if a direct transmission between the source and destination nodes fails.

In addition to defining the optimal set of relay nodes and the period after which the selection of a set of relays is triggered, it is also necessary to determine the way in which the retransmission of messages overheard by the relays is carried out. Studies of the state-of-the-art in this area (YUE et al., 2016; VALLE, Odilson T et al., 2016; LIU, X. et al., 2014) report that the use of a network coding technique can allow intermediary nodes to act upon the retransmitted messages and to modify them using specific mathematical coding techniques. Thus, the use of these techniques can increase the effective transmission rate of the network and the overall communication reliability.

The aim of this doctoral thesis is to address these limitations by treating the selection of relay nodes in a holistic way that combines both cooperative diversity techniques and the use of retransmission mechanisms with network coding algorithms, in order to enhance the reliability of communication in WSN networks.

In view of this, one of the fundamental issues considered in this thesis can be summarized in the form of the following research question:

– "Is it possible to increase the reliability of WSN communication by proposing a relay selection technique, considering a set of relevant criteria combined with retransmission techniques?"

In this context, this doctoral thesis investigates the use of new solutions and technologies based on cooperative diversity and retransmission mechanisms to improve WSN communication reliability.

1.2 RESEARCH OBJECTIVES

Considering that this doctoral thesis focuses on the improvement of the communication reliability in WSNs, the general and specific objectives of this thesis are presented in the following subsections.

1.2.1 Main Objective

The overarching objective of this research is to contribute to advancing state-of-the-art in WSN communication by proposing, developing and evaluating techniques based on cooperative diversity and retransmission mechanism.

The primary objective is to demonstrate that the use of relay selection techniques with relevant criteria for the operation of the network and retransmission mechanism can increase the reliability of communication in WSNs implemented based on the IEEE 802.15.4e standard.

1.2.2 Specific Objectives

The specific objectives of this doctoral thesis are as follows:

- To propose and to implement an innovative relay selection technique;
- To propose and to implement a retransmission mechanisms considering network coding technique;
- To combine the operation of both of these proposed techniques in a WSN to improve the reliability of communication;
- To propose a set of simulation models for the OMNeT++ simulator that will enable a simulation assessment of the proposed techniques;
- To define a set of performance metrics for the assessment of the proposed techniques, that can also be used to carry out a comparison with some of the most closely related techniques available in the literature.

1.3 RESEARCH CONTRIBUTIONS

The primary contribution of this work is the proposal of innovative relay selection techniques combined with retransmission techniques. The major overall contributions of this study can be summarized as follows:

- An overview of the state-of-the-art in relay selection techniques for WSNs, in which the main challenges and impairments in relay selection are identified and discussed;
- The proposal of a new relay selection technique;
- An analysis of which parameters should be considered in the relay selection process;
- A study of the best solution to the relay selection problem, and a comparative evaluation with other state-of-the-art relay selection techniques;
- An overview of the state-of-the-art in the field of network coding for WSNs.
- The proposal of an innovative retransmission mechanism, using a network coding technique allied to a relay selection technique, and an assessment of the WSN operation using both techniques.

1.4 METHODOLOGY

Scientific methodological procedures consist of a set of techniques and processes that are used to achieve an objective, understand the research process and allow the study to be replicated (MATIAS-PEREIRA, 2012). This section presents the methodological procedures that guided this research, based on the taxonomy of methodological procedures proposed by Freire (2013), as illustrated in 3.

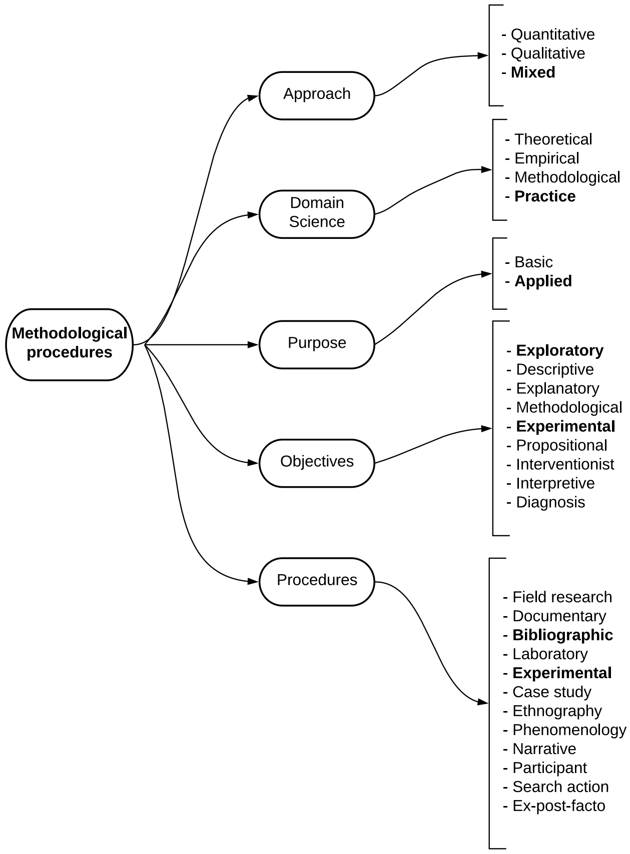
According to this classification, this work can be characterized as a mixed (qualitative-quantitative) research study, since it combines a bibliographic study of relay selection and network coding techniques with a numerical analysis of the results obtained from simulations (TOZONI-REIS, 2009).

In terms of the science domain, this work fits into the category of practical research, since it seeks to solve practical problems related to the communication in WSNs. As for the purpose of the research, this work can be characterized as an applied research study, as it aims to generate knowledge for practical applications and to solve specific problems in WSNs (FREIRE, 2013).

Regarding the objectives of the research, this work can be characterized as an exploratory and experimental study. It is exploratory because a systematic literature review is performed of the subject under study, and experimental because an evaluation is carried out of the quantitative results produced by a simulation (FREIRE, 2013).

With regard to the procedures used, as with any academic research, this work can be categorized as bibliographic research. It also falls into the class of experimental research, since virtual experiments are carried out to analyze and evaluate the proposed techniques using scenarios that are similar to real environments.

Figure 3 – Types of scientific research.



Source – Author.

1.5 THESIS OUTLINE

The remainder of this thesis consists of four previously published papers, which are presented in Chapters 2 to 5. Each paper makes a set of contributions towards improving communication in WSNs. Finally, Chapter 6 discusses the main conclusions of the research project and presents suggestions for future research.

More specifically, Chapter 2 consists of a paper published in *Sensors MDPI* (LAURINDO, S. et al., 2018), in which a new relay selection technique is proposed with the aim of increasing the reliability of communication in WSNs. The proposed scheme uses optimization techniques to ensure selection of the smallest number of relay nodes, and to ensure that each node is linked to at least one corresponding relay node, thus enabling messages to reach their destinations in noisy environments. In addition, two schemes are investigated for relay updating, which define when a new relay selection is triggered, based on periodic and adaptive approaches. The effectiveness of the proposed technique is assessed through an extensive set of simulations, and it is compared with other state-of-the-art relay selection techniques. The results are discussed at the end of this chapter.

Chapter 3 consists of a paper presented at the *Conference on Ad Hoc Networks* *and Wireless* (LAURINDO, S. et al., 2019), which contains an improvement to the proposed relay selection technique. Initially, the proposed technique was modeled as an optimization problem, and four parameters were considered in the objective function. To improve this approach, an analysis was performed to determine the impact of each parameter of the objective function, in which new objective functions were modeled based on each possible combination of parameters. As a result of this analysis, it was possible to optimize the objective function by reducing the number of criteria, thus simplifying the problem of relay selection. A discussion of the results is presented at the end of the chapter.

Chapter 4 presents a reproduction of a paper presented at the *Conference on Emerging Technologies and Factory Automation* (LAURINDO, S. et al., 2020), in which three different algorithms for solving the optimization problem of the proposed relay selection technique were assessed. The proposed relay selection technique can be characterized as a resource allocation algorithm that may be reduced to the classic Set-Covering Problem (SCP) applied to WSNs. In this study, the computational time required to solve the set-covering problem generated in the WSN relay selection was analyzed, using each of the three algorithms. The general operation of the WSN using each of the three resolution algorithms was also analyzed. A discussion of the results is presented at the end of this chapter.

Chapter 5 contains a paper published in the MDPI journal *Information* (LAURINDO, S. et al., 2021) that proposes four different retransmission mechanisms, three of which are based on network coding algorithms. These mechanisms were assessed operating together with the proposed relay selection technique, with the aim of analyzing the impact of each technique on the reliability of communication in WSN applications. A discussion of the results is presented at the end of the chapter. In addition, the chapter presents an extensive study of the state-of-the-art in the use of network coding techniques in the context of WSNs.

Finally, Chapter 6 presents the main conclusions of this thesis and suggests some new directions for future research.

2 A NEW RELAY SELECTION TECHNIQUE AND ITS APPLICATION TO A WSN

In this chapter, a new relay selection technique is proposed with the aim of increasing the reliability of communication in WSNs. This technique selects the set of relay nodes based on an optimization technique, in order to allow messages to be delivered from nodes that do not have direct communication with the coordinator node, thus increasing the throughput of the network. This chapter contains a transcript of a paper called "An optimized relay selection technique to improve the communication reliability in wireless sensor networks", which was published in *Sensors* (LAURINDO, S. et al., 2018). The next paragraph presents the abstract of the cited paper.

WSNs are enabling technologies for the implementation of the concept of the Internet of Things (IoT). Although WSNs can provide an adequate infrastructure for the last-link communication with smart objects, the wireless communication medium is inherently unreliable, and there is a need to increase the reliability of this communication. Techniques based on the use of cooperative communication concepts offer one way to achieve this aim. Cooperative communication techniques can be used in which certain nodes selected as relays transmit not only their own data, but also cooperate by retransmitting data from other nodes. A fundamental step towards improving the reliability of communication in WSNs is related to the use of efficient relay selection techniques. This chapter proposes a relay selection technique based on multiple criteria for selecting the smallest number of relay nodes while simultaneously ensuring adequate operation of the network. In addition, two relay updating schemes are also investigated, based on periodic and adaptive updating policies. Simulation results show that both proposed schemes, referred to here as periodic relay selection and adaptive relay selection, can significantly improve the communication reliability of the network when compared to other state-of-the-art relay selection schemes.

As mentioned above, this chapter contains a previously published paper, and is organized as follows. Section 2.1 presents a contextualization for the challenges faced by existing WSNs in terms of the exhaustion of energy resources and the unreliability of message communication, and possible ways of minimizing these challenges, such as the use of cooperative diversity techniques. Section 2.2 presents a study of state-of-the-art relay selection techniques, and introduces a classification scheme for relay selection techniques based on the parameters used during the relay selection process. Section 2.3 describes a novel relay selection technique that aims to improve the reliability of communication in WSNs. Section 2.4 reports the results of a simulation assessment of the proposed approach. Finally, conclusions are presented in Section 2.5.

2.1 INTRODUCTION

One of the most pursued goals in today's computing environments is the ability to run and access computing data, anywhere, anytime. This concept—commonly known as ubiquitous computing—can be achieved using the emergent Internet of Things (IoT) paradigm, where sensors and actuators are seamlessly integrated into the environment (GUBBI et al., 2013). One of the major objectives of IoT is to enable smart objects or "things" to communicate with each other and cooperate to achieve a common objective (LAZARESCU, 2013; HAN, G. et al., 2018). Due to their characteristics, such as low cost, ease of deployment, ability to closely monitor physical phenomena of interest, Wireless Sensor Networks (WSNs) provide a suitable support for low-rate monitoring applications, and are considered as one of the enabling technologies for the dissemination of the IoT paradigm (LAZARESCU, 2013; LI, Xiong et al., 2018).

Nevertheless, WSNs are subject to restrictions in what concerns the exhaustion of energy resources and the unreliability of message communication (GUNGOR, V.; HANCKE, G., 2009; VASEGHI et al., 2017). Energy consumption is a problem because nodes have limited energy resources. Normally, they are powered by chemical batteries, that when discharged must be either replaced or recharged. However, it is not always easy to replace batteries, especially in large-scale networks. To reduce the energy consumption of WSN nodes, these type of networks are usually configured to follow a duty-cycle operation. That is, nodes will periodically switch off parts of their circuits, becoming idle during significant periods of time. Moreover, low communication reliability may lead to a severe reduction of the achievable throughput in WSNs. This problem is associated with messages being lost due to electromagnetic noise and/or other devices that operate in the same frequency range or obstacles between nodes (VALLE, Odilson T et al., 2016).

A possible solution to improve the reliability of WSNs is by providing multiple paths to transmit data from the source to the destination node. As a consequence, a message transfer would still be possible if there are other available links whenever a link fails. This type of communication is referred to as cooperative diversity, and it allows the enhancement of communication over unreliable channels (NHON; KIM, D. s., 2014).

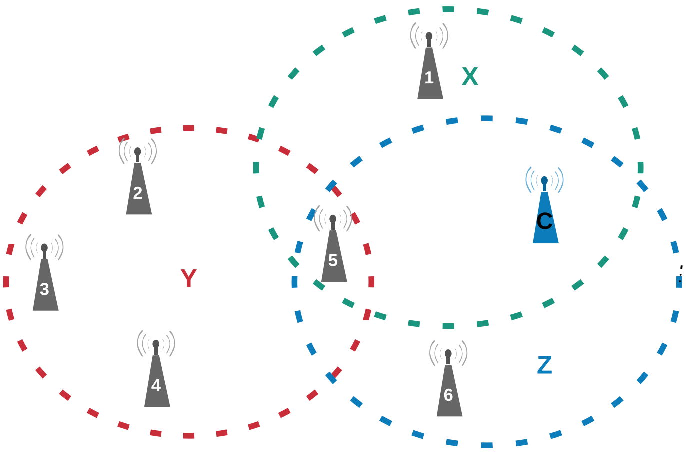
In cooperative diversity techniques, networks may use single-antenna equipments to achieve the same benefits of multiple-input-multiple-output (MIMO) systems. It allows a set of nodes to exploit their antennas to form a virtual MIMO system, avoiding the deployment of multiple antennas that would increase the overall energy consumption (WANG, C. L.; SYUE, 2009). Unlike traditional WSN communication, where packets have a destination address and the transmission involves only one transmitter and one receiver, cooperative diversity considers the existence of nodes that will cooperate with the transmitter–receiver pair, listening and storing messages received from their neighbors, and then retransmitting them to the destination node. Thus, the destination node has a higher probability of receiving the sent packets because messages that were not directly received by the destination node may be received during the retransmission phase (KHAN; KARL, 2014). This communication behavior allows a better usage of the broadcast nature of wireless transmissions (LIANG et al., 2009), since the message diffusion can be heard by the neighboring nodes (as long as they have their radios on), improving the network success rate without increasing its hardware complexity.

Whenever using cooperative diversity, the selection of the set of relays is of paramount importance to achieve a good performance level. It is important to point out that selecting all nodes as relays would allow a greater cooperative diversity. However, at the same time, it would increase the number of message collisions, the overall energy consumption, and would affect the synchronization among all those nodes (WANG, C. L.; SYUE, 2009). Consequently, one of the major challenges of cooperative retransmission techniques is the selection of the best set of relay nodes, in order to improve the overall communication reliability (JAMAL; MENDES, 2010).

The selection criterion is one of the most relevant aspects when selecting the set of relays nodes. When selecting the cooperator nodes, many studies consider just quality estimators for this purpose, such as Received Signal Strength Indicator (RSSI), Channel State Information (CSI), Signal-to-noise ratio (SNR) or Link Quality Indicator (LQI) (LI, Yubo et al., 2011; FERDOUSE; ANPALAGAN, 2015; ALKHAYYAT et al., 2015; MARCHENKO et al., 2014; PHAM; KIM, D. S., 2016; VALLE, Odilson T et al., 2016). However, the consideration of just this type of quality estimators may generate inaccurate decisions (BACCOUR et al., 2010). In fact, hardware metrics such as LQI, RSSI and SNR are based on just the first eight symbols of a received packet, and are therefore unable to provide an accurate estimate of the link quality. More importantly, they are only measured for successfully received packets. As a consequence, whenever there are frequent packet losses, the above metrics will overestimate the quality of the link. Therefore, despite providing a quick way to classify communication links as good or bad, this type of link quality estimators are unable to provide accurate estimates. Thus, they should not be independently considered, but rather by combining them with other metrics (BACCOUR et al., 2010).

Figure 4 presents a WSN composed of seven nodes, one being the coordinator node. In this WSN, there are three regions identified as X, Y and Z. All nodes in the same region are considered to be neighbors and can hear each other. Nodes 2, 3 and 4 that are in the Y region are not able to directly communicate with the coordinator. In this network, it is possible to visualize the need for using relays, so that messages from all nodes of the network can reach the coordinator node. In addition, it is also possible to observe that to select adequate relays it would be necessary to consider other parameters besides the channel quality. For instance, nodes 1 and 6 have a good communication link with the coordinator, but, they are far from the nodes that need to use cooperation techniques to reach the coordinator. In this scenario, it would be more appropriate to select node 5 as the relay node.

Figure 4 – Wireless sensor network that needs a relay.



Source – Author.

On the other hand, whatever the parameters used to perform the relay selection, the updating activity can follow a periodic or an adaptive strategy (MARCHENKO et al., 2014). In the periodic strategy, the relay updating always occurs, independently of the network requirements. An adaptive strategy considers a specific updating policy, which considers relevant modifications of the environmental conditions to trigger a new relay selection.

This chapter proposes a new relay selection technique to be used in WSNs, named Optimized Relay Selection Technique (ORST). It also investigates the use of two different relay updating schemes. The term optimized is used in this chapter in the sense of an improved way to select relay nodes, when compared to other state-of-the-art approaches. The periodic scheme, named Periodic Relay Selection (PRS), is an extension of the scheme proposed in Suelen Laurindo et al. (2017), where the relay selection is periodically triggered, without analyzing the need for a new selection. In the adaptive approach, named Adaptive Relay Selection (ARS), the time interval between consecutive relay selections is dynamically determined, according to the network's success rate. The ARS scheme is able to handle dynamic networks, where nodes may randomly join or leave the coverage area of the coordinator node.

For both relay selection schemes, the aim is to maximize the communication success rate and also to minimize the energy consumption of the network, by selecting the smallest number of relay nodes and, at same time, ensuring that all nodes are linked to at least one corresponding relay node. When compared to other techniques, the main scientific contribution of the proposed technique is the selection of a set of relay nodes being based on multiple criteria, namely: the number of neighbors of the candidate node, their remaining battery energy, the quality of the communication link between the candidate node and its neighbor nodes (by using the RSSI) and the success rate's history in recent node transmissions, which provides the adequate selection of relay nodes and improve the communication reliability, since these criteria are critical for the operation of the network. However, this technique is more complex than when using just a single criterion to rank the candidate nodes to become relays. As multiple constraints are included in the relay selection model, the proposed scheme is formulated as an optimization problem, using a specifically selected benefit function.

A simulation assessment of both schemes was performed using the OMNeT++ tool (COMMUNITY, 2011). The proposed approaches were compared with three state-of-the-art techniques: Opportunistic (VALLE, Odilson T et al., 2016), which selects the cooperating nodes according to the network packet error rate, Random Around the Coordinator (ETEZADI et al., 2012), which performs a random selection of the nodes that have an adequate communication link with the destination node and Completely Random relay selection (WILLIG; UHLEMANN, 2012), which performs a random selection from all the nodes of the network.

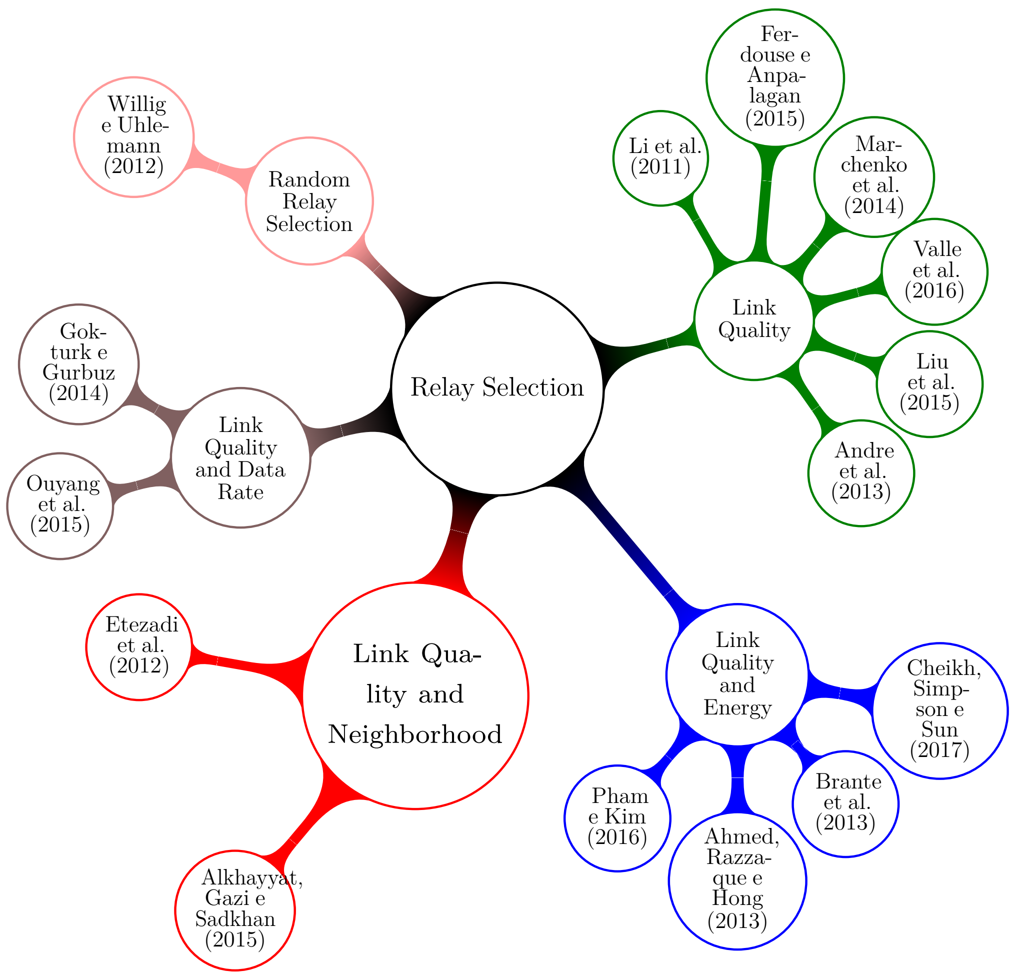
2.2 RELATED WORK

The adequate selection of relay nodes is fundamental to improve the performance of cooperative communication techniques (ETEZADI et al., 2012). There are multiple schemes to select relay nodes in the literature. However, many of these works usually focus on the analysis of the data communication protocols, evaluating and proposing different techniques for sending messages (or signals) by the relay nodes, based on Amplify-Forward (AF), Decode-Forward (DF) or Hybrid approaches (SELVARAJ; MALLIK, 2010; RAJANNA et al., 2017; TUAN et al., 2018). As a limitation of these works, it can be mentioned that they do not deal with the selection of retransmission nodes, or, when they do, they assume only a single criterion for such a selection. This work aims to improve the relay selection in WSNs. In this way, the objective of this section is to describe state-of-the-art approaches that mainly deal with the relay selection procedures. Then, state-of-the-art selection schemes were classified into five categories, according to the criteria used for the relay selection: Link Quality Based (MARCHENKO et al., 2014; LIU, L. et al., 2015; LI, Yubo et al., 2011; FERDOUSE; ANPALAGAN, 2015; ANDRE et al., 2013; VALLE, Odilson T et al., 2016), Link Quality and Energy Based (BRANTE et al., 2013; AHMED, M. H. U. et al., 2013; PHAM; KIM, D. S., 2016; CHEIKH et al., 2017), Link Quality and Neighborhood Based (ETEZADI et al., 2012; ALKHAYYAT et al., 2015), Link Quality and Data Rate Based (GOKTURK; GURBUZ, 2014; OUYANG et al., 2015) and Random Relay Selection (WILLIG; UHLEMANN, 2012) (Figure 5).

2.2.1 Link Quality Based

In this category, the relay selection only considers the use of a link quality estimator. (LIU, L. et al., 2015) proposed a relay selection scheme for clustered wireless networks. The authors considered that there should be a relay set in the source node's cluster (transmitter T) and a relay set in the destination node's cluster (receiver R), with a maximum number of relays K, such that, $|T| + |R| \leq K$.

Figure 5 – Relay selection categories based on the criteria used to select the cooperating nodes.



Source – Author.

The set T is incrementally formed, selecting relays with the maximum channel gain up to the relays of set R. The selection occurs until the number of relays reaches K or until the SNR requested by relay in R is reached. To select the receiver set R, the authors defined that only nodes with a non-negative SNR margin could be qualified as Relay Receivers (QR). Two heuristics were proposed to further reduce the searching space of R. The basic idea is to define a set of Candidate Receivers (CAR) based on *QR*. In the first heuristic, the selected relays are the nodes with more capability to decode the received signal from the transmitting relays. In the second heuristic, the selected relays are the nodes geographically closer to the destination node, which present the maximum achievable SNR margin up to the destination node.

After defining the *CR* set, the receiver relay set *R* is selected considering just the nodes belonging to *CR*. The nodes in *CR* are arranged in decreasing order according to the SNR to the destination node and the set *CR* will be the input to determine the transmitting relays set *T*. If a feasible solution can be found, then both sets *CR* and *T* are the optimal solutions.

Marchenko et al. (2014) presented an adaptive relay selection based on LQI, where the relay selection is triggered depending on the recent delivery ratio performance of the cooperative link. A source node *S* monitors the Acknowledgments (ACK) of the sent packets. If the ACK for a packet is missing, it assumes that neither *S* nor the currently assigned relay R_i could deliver this packet. If the ratio of missing ACKs during a contention window W_a is equal or higher than a threshold value ε_a, a new relay selection is triggered. To perform the selection of relay nodes, *S* broadcasts a relay request message, which also includes the identification of the destination node *D* for the following data packets. All nodes receiving this request will trigger a random timer $T_c = \text{rand}(0, W)$ for a transmission during a contention window of duration *W*. When the timer of a node expires, it sends a candidature message to *D*. This message includes the measured LQI value and the value of W-T_c. Thus, *D* can identify the end of the contention window even if it has not received the relay request message. Nodes whose candidature message is received at *D* will form a relay candidate set *R*. When the contention interval ends, *D* evaluates the end-to-end link quality for each candidate node R_i by taking the minimum of two LQI values (*S* to R_i and R_i to *D*).

Yubo Li et al. (2011) proposed a relay selection technique using CSI as criterion, which can be performed either as a centralized or a decentralized approach. In the centralized approach, the relay selection is done at the destination node. In the decentralized approach, nodes estimate their channel gains through the exchange of handshake messages Request-To-Send (RTS) and Clear-To-Send (CTS) and trigger a timer with a value inversely proportional to the gain. The node that has the first expired timer is selected as relay.

Ferdouse and Anpalagan (2015) proposed a relay selection scheme based on the use of the Bayes's theorem. The authors consider CSI and SNR values between source-destination pairs and the source-relay pairs as criteria for the selection. The authors use two factors for the relay selection. The first performs the relay selection, the destination node or the source node; and the second is responsible to define the retransmission technique, AF (Amplify-and-Forward) and DF (Decode-and-Forward). The authors also assume that the CSI value, the achievable data rate under AF (C_{AF}), and the data rate in direct transmission (C_D) are known. Based on the CSI information, it is calculated the SNR_D, the SNR_{AF} or the SNR_{DF}. Also based on the SNR value, the capacity of the channel from the source node to the destination and from the source node to the possible relays is calculated. Afterwards, it is calculated the prior probability for each source node $P(S_1)\dots P(S_n)$ and the conditional probability for each possible relay $P(R_1|S_1)\dots P(R_n|S_n)$, applying the Bayesian theorem to determine the posterior probability for each source node and possible relays in the network $P(S_1|R_1)\dots P(S_n|R_n)$. The relay selection is based on the maximum posterior probability, if $P(S_n|R_k) > P(S_n|R_j)$ to $k \neq j$, then the R_k will be selected.

The main difference between the source and destination methods is that when a source node performs the relay selection, it collects information about the channel of the possible relay during the beacon period, measures its SNR and estimates the posterior probability, while the destination node collects all channel information between source-relay (CSI and SNR) to calculate the posterior probability and assign a relay to all active transmission (source and destination) pairs.

Andre et al. (2013) presented an adaptive and decentralized relay selection technique, where the relay selection occurs each five failed transmissions. The criteria used for the relay selection is the LQI, where a timer inversely proportional to the link quality is triggered at each node. Thus, the node with the best quality will have its timer expired first and will be the relay node. The operation takes place as follows: the source node broadcasts a relay frame; potential relay nodes that listen to this message will store the channel quality with the source node and wait for the response of the destination node. When the destination node responds, the nodes that listen to the message store the channel quality to the destination. Only the nodes that listen to both messages may dispute the relay selection. The candidate relays will activate a timer inversely proportional to the link quality from the source node to the destination node. When the timer expires, the candidate relay sends a message to the source node to inform that it will be the relay and, after receiving it, the source node sends a message signalling that the relay has already been selected. Thus, if after a source node sending a message, the destination node does not send its ACK, which means that the relay node will retransmit the message.

The authors conclude that the adaptive relay selection is beneficial compared to the periodic selection, improving the delivery rates. However, the improved reliability comes at the cost of an increased overhead, as this technique requires additional message exchanges between the source and the destination nodes. Additionally, as the adaptive relay selection is more frequently triggered than the periodic technique, it also generates an increased overhead.

Another relay selection technique, called Opportunist, was proposed by Odilson T Valle et al. (2016). This technique considers the history of successful transmission rates and the LQI between each node and the coordinator. To select the best relay nodes among the candidates, the authors propose the following equation:

$$CN_i = \frac{SR_i + LQI_i}{2}, \tag{1}$$

where, for each node i, SR_i is the history of recent successful transmission rates and LQI_i is the quality indicator of the link between the node and the coordinator.

Nodes presenting the highest CN_i will be selected as relays. The number of relay nodes is dynamically defined according to the percentage of message losses in the network. The coordinator notifies the relay nodes through a special message called *Blocop*. The selected relays store all messages that are sent by all nodes during each of the T slots (where T is the total number of slots within the transmission period). After the end of the T slots, the relay nodes encode the set of messages stored in their buffers and wait until specific slots to perform the retransmissions.

2.2.2 Link Quality and Neighborhood Based

In this category, Etezadi et al. (2012) proposed three techniques for the selection of relay nodes. The first one uses the source-to-destination SNR as criterion; the second is based in a geometrical analysis, selecting nodes closer to the source node; and the third one randomly selects relay nodes located in an R radius around the coordinator.

The first two techniques have a higher energy consumption. In the first, the authors assume that SNR can vary frequently. Thus, the set of nodes with the highest SNR can be modified between two consecutive transmissions. This fact suggests that, in this scheme, all nodes are possible candidates to be relays, and, therefore, should always switch between the listening and the transmission mode without the privilege of going to sleep and preserving their energy. In the second, the set of nodes closest to the source node rarely undergo updates. Therefore, these nodes continue to switch between the listening and the transmission modes, leading to an early exhaustion of their batteries, while all the other nodes are able to remain in sleep mode and saving energy. The third technique randomly selects a new group of relay nodes among the nodes that are within a radius R around the source node $O(S, R)$. According to the authors, this technique is the one with the lowest energy consumption.

Alkhayyat et al. (2015) presented a relay selection scheme whose aim is to select the best relay node considering the quality of the possible relay with the destination node in consideration. The authors defined a selection area and only the nodes that are within this area can participate in the relay selection. This area is defined as follows: when a source node has a message to transmit, it begins by transmitting an RTS message to the destination. If the destination receives the message correctly, it will transmit a CTS message to the source node. The nodes that listen to both messages

(RTS/CTS) are considered to be within the selection area and can dispute the relay selection.

After the RTS/CTS message exchange, the source node will transmit its message, and the nodes that are both within the selection area and are able to decode the message will dispute the relay selection. Each of these nodes will initiate a contention period, during which a backoff time function decreases based on the distance and the received signal quality from the relay to the destination node. The evaluation of this function is done at each node, based on the information obtained through the handshake messages (RTS/CTS). Thus, the node that gets the lowest value from the backoff time function will access the channel first and will be the relay.

2.2.3 Link Quality and Energy Based

In this category, Brante et al. (2013) presented a decentralized fuzzy-based relay selection technique, that uses the relay-to-destination CSI and the node's battery level (E_i) as parameters. It considers that each node is able to read its battery level(E_i) and to estimate the state of its channel (g_{iD}) using the Negative-Acknowledgment (NACK) message sent by the destination node. Each node evaluates its degree of relevance to act as relay, which is a function of g_{iD} and E_i. The higher $f(g_{iD},E_i)$ is, the higher is the quality of the relay. Thus, when the node gets its degree of relevance, it waits for a time interval T_i to avoid a collision, in case more than one node presents the same degree of relevance. In the sequence, it will act as a relay and retransmit the message to the destination node.

Mohammad Helal Uddin Ahmed et al. (2013) proposed a relay selection protocol that aims to maximize the network lifetime and to minimize the end-to-end delay in packet delivery. The authors consider that each node is able to evaluate a weighted average W from its residual energy and the average delay of sent packets. The protocol operation is as follows: the source node sends an RTS message and the destination node responds with a Request-To-Help (RTH) message containing WH and WL values, which are defined by the destination node and correspond to a highest and lowest threshold W. Nodes that listen to both messages (RTS/RTH) will evaluate their weight W and will contend to send a message Interested-to-Help (ITH) if its W value satisfies the condition $WL \leq W_n \leq WH$.

In this technique, if more than one node has a W value that satisfies the condition $WL \leq W_n \leq WH$ a collision may occur, which will make the relay selection process slower. Whenever a collision occurs, the destination node will increase the value of its lower threshold WL, aiming to reduce the number of nodes that can be within the range of WL and WH. Then, nodes that still satisfy the constraint try to send the ITH message again. The first node that is able to send its own message to their destinations will be the relay.

Pham and Dong Seong Kim (2016) presented a relay selection technique that aims to increase the network lifetime and improve its packet delivery. The node with the highest residual energy level and the lowest Energy-Per-Bit (EPB) value will be the relay node. The basic idea is that, based on the CSI and Residual Energy Information (REI) of each relay, it first decides which protocol is better for the transmission. If the channel SNR between the source node and the cooperating candidate node is greater than a previously defined value, the node will use AF, otherwise, it will use DF. After selecting a suitable forwarding protocol, the node with the higher residual energy e_i^m and the lower EPB (*Energy-Per-Bit*) E_b will be selected as relay. Thus, the authors proposed the calculation of a weight w for each cooperating node r, as: $w_r = \frac{e_i^m}{E_b}$, where, the relay with the highest weight value is selected.

Cheikh et al. (2017) proposed a relay selection technique that aims to decrease the energy consumption of the network. The authors propose an optimization problem that determines which node minimizes the energy consumption per bit in transmissions between source and relay nodes and between relay and destination nodes. The destination node implements an optimization algorithm and selects the node that minimizes the energy consumption.

The operation of this technique is as follows: the source node sends an RTS message, signaling that it has data to transmit; the destination node responds with a CTS message. The nodes that listen to both messages estimate the channel gain and the desired transmission power to reach the target Bit Error Rate (BER). Using an optimization strategy, the destination node selects the relay node, and signals through the beacon message which was the selected node.

2.2.4 Link Quality and Data Rate Based

In this category, Gokturk and Gurbuz (2014) proposed two techniques for the selection of relay nodes, using Frame Error Rate (FER) as selection criterion, one being centralized and the other decentralized. In the centralized technique, the source node will incrementally select the nodes in the relay set. The number of relays in the set is incremented until the set satisfies the target FER, as long as the set presents a minimum power consumption. The nodes that were able to successfully decode messages coming from the source nodes are part of the relay set. In the decentralized technique, each node decides whether it will be relay or not and calculates its own optimal power level. Each node that was able to decode the message from the source node calculates the power allocation required to satisfy the target FER in the transmission. If a node concludes that its inclusion in the relay set can help the achievement of the required FER, it announces its intention for cooperation. Each time that a node advertises its availability to cooperate, the other nodes calculate the optimal power assignment and the energy cost for the new relay set and compares it with the energy cost of the set before its insertion. The algorithm continues until the energy cost per bit can not decrease any more. If no nodes are available for the cooperation or the FER requirement can not be satisfied with the cooperation of the available nodes, the cooperation is aborted.

Ouyang et al. (2015) presented a blind relay selection technique, based in a random medium access mechanism (ALOHA). The idea is that the node able to achieve a required transmission rate at the destination and able to access the channel first will be the relay. Due to high collision probability, the authors determine that only the nodes that are within a given region called Available Relay Selection (ARR) will participate in the selection. To be in this region, the nodes need to satisfy a required transmission rate. The operation of the proposed technique occurs in three phases. In the first phase, the target is to identify the possible relay nodes. Thus, the source nodes transmit a message with the required rate R_r and their locations. The nodes that listen these messages will evaluate which nodes are able to achieve the required rate. The nodes that satisfy the R_r rate respond with a 'ready-to-relay' message. In the second phase, all possible relays will start a competition phase based on a slotted ALOHA medium access so that the relay selection may occur. Each possible relay transmits a 'select-me' message to the source node; if just one node transmits a message within the slot, it means a successful selection and the relay is selected for the retransmission in the current frame. Otherwise, a collision occurs and the nodes will attempt to send the message in the next slot. Phase two ends whenever a selection occurs (that is, when the first 'select-me' message sent by a possible relay successfully reaches the source node). When a selection occurs, phase three begins, starting the retransmission.

2.2.5 Random Relay Selection

Finally, in this category, the relay selection is performed without considering any criteria; therefore, any node in the network can be selected as relay. Willig and Uhlemann (2012) conducted a study about this topic and concluded that relay nodes do not need to be in specific positions, which often requires specific channel quality measurements. According to the authors, randomly selected relay nodes can also provide benefits to a WSN, improving its success rate.

2.2.6 Wrap-Up

Table 1 summarizes the described works, comparing them among themselves with respect to the following set of classifiers: used category to select the relays, if the selection technique uses a periodic (P) or adaptive (A) approach and if it requires the use of additional message exchanges.

From Table 1, it is possible to highlight some common characteristics. For instance, a large number of works consider just the link quality factor, which may generate bad decisions when selecting the set of relays (BACCOUR et al., 2010). The main reason is that hardware metrics are based on the sample of just the first eight symbols of a correctly received packet, not considering the lost packets. Thus, it may provide an inaccurate estimate of the link quality.

Table 1 – Relay selection techniques.

State-of-the-Art	Category	Periodic or Adaptive	Exchange of Additional Messages
Lingya Liu et al. (2015)	Link Quality	P	
Marchenko et al. (2014)	Link Quality	P/A	$\surd$
Yubo Li et al. (2011)	Link Quality	P	$\surd$
Ferdouse and Anpalagan (2015)	Link Quality	P	
Andre et al. (2013)	Link Quality	A	$\surd$
Odilson T Valle et al. (2016)	Link Quality	P	
Etezadi et al. (2012)	Link Quality and Neighborhood	P	$\surd$
Alkhayyat et al. (2015)	Link Quality and Neighborhood	P	$\surd$
Brante et al. (2013)	Link Quality and Energy	P	
Mohammad Helal Uddin Ahmed et al. (2013)	Link Quality and Energy	P	$\surd$
Pham and Dong Seong Kim (2016)	Link Quality and Energy	P	$\surd$
Cheikh et al. (2017)	Link Quality and Energy	P	$\surd$
Gokturk and Gurbuz (2014)	Link Quality and Data Rate	P	
Ouyang et al. (2015)	Link Quality and Data Rate	P	$\surd$
Willig and Uhlemann (2012)	Random Relay Selection	P	
Proposed Technique	Multi Parameters	P/A	

Source – Author.

Another characteristic that can be observed is that most of the proposed techniques require the use of additional message exchanges to make the selection. Many of them use handshake techniques (RTS/CTS) to estimate the channel quality between source-relay-destination nodes. However, when it comes to WSNs, handshake messages may generate a relevant overhead in the network and, in case of collisions, may generate considerable additional delays.

Finally, most of the presented works periodically perform the relay selection. A periodical selection may be inefficient for two reasons. First, the relay selection can be unnecessarily triggered because the current relay remains the best option. The second reason is that, if the selection always occurs at a fixed interval, there is always the possibility of an outdated relay selection to occur. For instance, a node may leave the network, a new node may join the network or even an increase in the level interference may occur. If the selection periodicity is long, there is a risk that the current set of relays may not be the best one. Thus, whenever the network has a dynamic behavior, it is necessary to avoid these problems using, for instance, adaptive selection techniques.

To overcome some of the disadvantages highlighted by this state-of-the-art analysis, the technique proposed in this chapter considers multiple criteria as parameters for the relay selection. In addition, it does not require the exchange of any additional messages to converge.

2.3 ORST-OPTIMIZED RELAY SELECTION TECHNIQUE

In this chapter, it is proposed a centralized relay selection technique, named ORST (Optimized Relay Selection Technique). The selection of relay nodes is formulated as an optimization problem, using a benefit function that considers the following selection parameters: (a) the number of neighbors of the candidate node; (b) its remaining battery energy; (c) the quality of the link between the candidate node and its neighbor nodes; and (d) the history success rate in recent node transmissions. In addition, this chapter also investigates the usage of two different relay update schemes, PRS (Periodic Relay Selection) and ARS (Adaptive Relay Selection).

2.3.1 System Model

The system model considers a network organized in a star topology. It is also considered that nodes without a direct link to the coordinator will use a neighbor relay to establish a communication link with the coordinator. The use of a star topology in industrial applications is commonly justified due to its advantages in terms of latency, synchronization, simplicity and also due to its energy efficient behavior (VALLE, Odilson T et al., 2016; CHEN, Feng et al., 2008).

A slotted communication approach is assumed, where the medium access is based on a Time Division Multiple Access (TDMA) scheme. This type of communication approach is of common use whenever dependable communication is required, as it increases the communication reliability allowing the medium access without contention (GUNGOR, V.; HANCKE, G., 2009; VALLE, Odilson T et al., 2012).

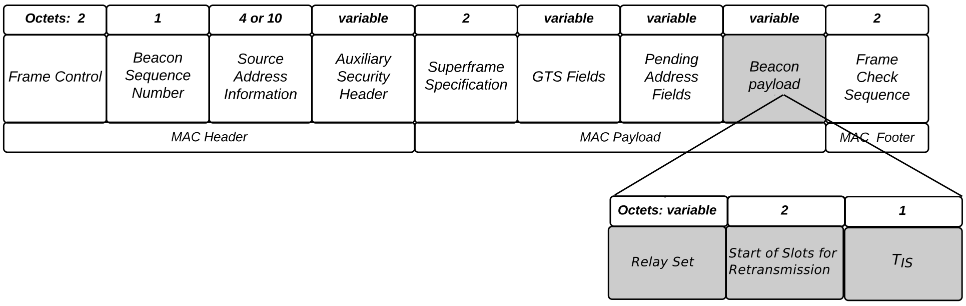
The IEEE 802.15.4 standard operating in beacon-enabled and time-slotted mode is adopted for the PHY (Physical) and Medium Access Control (MAC) layers of the network. It is important to note that this communication scheme is similar to Low Latency Deterministic Networks (LLDN), which was also used in the cooperative retransmission schemes proposed and assessed in Willig et al. (2016).

The proposed relay selection technique assumes that there is specific information exchanged among the coordinator and other nodes during each period of the network, being the information sent by the coordinator piggybacked with the beacon frame and the information sent by the nodes piggybacked with the monitored data.

The payload field of the beacon frame is used to send three parameters, Relay Set, which is the information with the identifiers of the selected relay nodes, Start of Slots for Retransmission, which reveals which is the first slot for retransmission (retransmission slots are contiguous slots that are allocated after the transmission slots) and T_{IS}, which reveals if the nodes should listen to their neighbors in this Beacon Interval (BI) or not. Figure 6 illustrates the beacon frame structure.

The data payload field is also used by nodes to piggyback two types of informa-

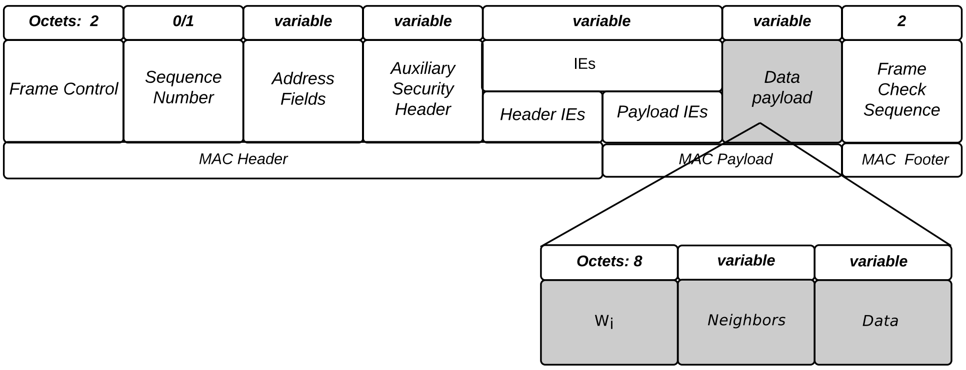
Figure 6 – Beacon frame format.



Source – Adapted from (IEEE COMPUTER, 2015).

tion for the coordinator: Neighbors, which reveals the list of neighbors of each particular node and W_i that reveals its benefit value (described in Section 2.3.3). Figure 7 illustrates the frame structure of a data message.

Figure 7 – Data frame format.



Source – Adapted from (IEEE COMPUTER, 2015).

In the relay message, the payload data field is used to send the list of `heard messages`. This field is formed by a list that contains just the *id* of the transmitting node and the *id* of the transmitted message, enabling the coordinator to assess the message success rate. As the major focus of the proposed ORST is the adequate selection of relay nodes, no data fusion technique (PINTO et al., 2014), nor retransmission protocols as AF, hybrid approaches or network coding techniques have been implemented (SELVARAJ; MALLIK, 2010; SU, W. et al., 2008; AI; CHEFFENA, 2017; VALLE, Odilson T et al., 2016).

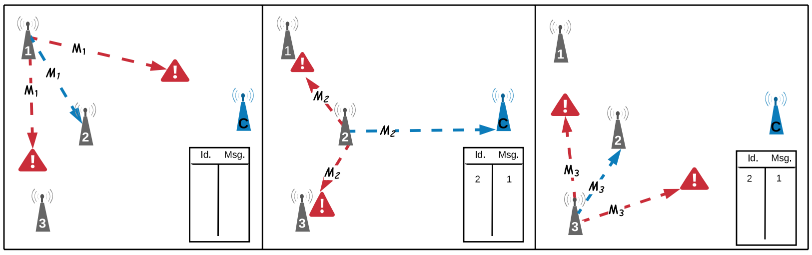
2.3.2 Brief Explanation of the Relay Selection Operation

When a node joins the network, it synchronizes itself with the beacon message sent at the beginning of the beacon interval (BI). Considering a time-slotted medium access, all nodes will make a transmission attempt during their respective time slot. During the first BIs, the network still does not have any relay. After starting the network operation, there will be a configuration period within which nodes will identify their neighbors, calculate the benefit function value (described later), and send this information back to the coordinator node. Only the nodes with RSSI greater than or equal to –87 dBm will be added as neighbors, as this value indicates that there is an adequate communication link between them, as suggested by Srinivasan and Levis (2006).

During the following BIs, each node is already aware of its neighbors, this information being used to evaluate the benefit value for the relay selection. Afterwards, each node sends to the coordinator its benefit value and the related neighborhood information piggybacked with the data message. The coordinator will use this information to perform the relay node selection. The information about which nodes were selected as relays will then be piggybacked by the coordinator in the next beacon.

After notifying all the relay nodes, the communication will occur in two steps: transmission and retransmission. In the first step, as illustrated in Figure 8, each node make a transmission attempt, the set of relay nodes will stand by listening to and storing all messages sent by all nodes, storing both the successfully received messages and the identification (*id*) of the sending node. If the node is not a relay, it will enter into a sleep mode when finishing the transmission step ($Node_1$ and $Node_3$ in Figure 8).

Figure 8 – Transmission steps.

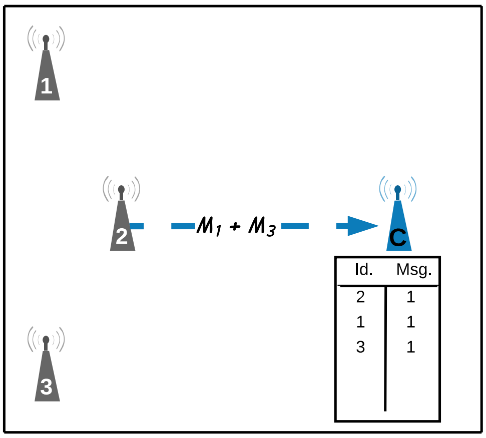


Source – Author.

After the first selection of relay nodes, the interval between selections will be defined according to the used selection scheme (PRS or ARS). Therefore, this interval can be fixed (PRS) or adaptive (ARS). During the two BIs that precede a new relay selection, all nodes of the network will remain active until the end of the transmission step, as they need to listen to the nearby nodes to update their list of neighbors.

In the second step, represented in Figure 9, each relay node will retransmit one message containing all the *Ids* of the successfully received messages to the coordinator. The coordinator stores in a table all incoming messages and compares them with the received relay messages. Whenever the coordinator receives a retransmitted message that it had previously received, it counts this message as a duplicate, which is a useless message. This information is not used by the proposed selection technique; it will be used just as statistical information. The percentage of received duplicates will be used as a metric to evaluate the efficiency of the relay node selection technique. After the retransmission, the relay will enter in sleep mode.

Figure 9 – Retransmission steps.



Source – Author.

2.3.3 Problem Formulation

The proposed ORST scheme aims to find a set of relay nodes $S^\star = \{y_1, y_2, \ldots, y_m\}$ among a set of nodes $X = \{x_1, x_2, \ldots, x_n\}$ in WSNs, ensuring two conditions: (1) each node x_i ($1 \leq i \leq n$) is covered by at least one relay node; (2) the sum of the weights of the relays is minimized. In this scheme, x_i is used as node identifier, n is the total number of nodes in the network, m is the total number of relay nodes and $S^\star \subseteq X$, i.e., relays are selected in the same set of nodes, transmitting not only their own data, but also cooperate by retransmitting data from other nodes. There is one node called a coordinator in the WSN (C).

A node x_i will be a candidate to be a relay if and only if: (a) it is neighbor of the coordinator; and (b) it has at least one more neighbor. The ORST scheme is a kind of resource allocation algorithm that may be reduced to the classic set-covering problem applied to WSNs (XU, K. et al., 2005).

Then, considering a WSN composed of a set of nodes $X = \{x_1, x_2, \ldots, x_n\}$, being that every node has an associate positive weight value and a specific communication range, we construct an undirected and weighted graph $G = (X, E)$ in the following way. Each node x_i corresponds to a vertex $x_i \in X$ and two vertices x_i and x_j have an edge $e_{i,j} \in E$ if x_i is able to hear a message sent by x_j with the value of RSSI ≥ -87 dBm, as defined by Srinivasan and Levis (2006) as the minimum value for an adequate communication in WSNs.

Every graph with X and E has subsets $F = \{S_1, \ldots, S_K\}$, where each subset S_k is known as a set cover of the graph G. The set-covering problem consists of a finite set X and a family F of subsets of X, such that every vertex of X belongs to at least one subset in F. Each subset of F is formed by vertices that accomplish conditions (a) and (b).

It is said that a subset $S \in F$ covers its element. The problem is to find the set with the minimum sum of weight subset $S^{\star} \subseteq F$ whose members cover all vertices of X. The WSN problem treated in this chapter, which consists of finding the set-cover with minimum sum of weights, is a special case of the set-covering problem. The corresponding decision problem generalizes the well-known NP-complete vertex-cover problem and is therefore also NP-hard (CORMEN et al., 2009; KARP, 1972).

Differently from other relay selection techniques, ORST selection methodology is based on multiple criteria, represented by a weight W_i that is assigned for each node x_i. This weight takes into consideration the available energy in the nodes, the number of neighbors that each node can hear (RSSI $\geq$ –87 dBm), the quality of communication between the source node and the candidate relay node, as well as the history of successful transmission rates of node x_i. These parameters were selected because they are highly relevant for the operation of the network. For instance, the residual energy of the nodes is an important parameter, considering that, if a node has a low battery level it will stop being a promising candidate, because soon it will exhaust its own energy resources. The number of neighbors that each node has is also a parameter that must be considered, since if a node does not have neighbors, it does not make sense to select it as a relay. The quality of the channel between the source and the relay nodes is another important parameter because it allows for knowing if there is a good communication link between these nodes, ensuring that the relay node correctly receives the message to be retransmitted. Finally, the history of successful transmission rates is an indication that the selected node has a good communication link with the destination node, ensuring that messages sent by this node will correctly arrive to their destination. Combining these parameters as the selection criterion, the target is to ensure that appropriate nodes will be selected as relay nodes. Each node x_i will evaluate its benefit W_i as follows:

$$W_i \doteq \left(\frac{\beta^{v}}{v_i} + \frac{\beta^{e}}{e_i} + \frac{\beta^{s}}{s_i} + \frac{\beta^{H}}{H_i} \right), \tag{2}$$

where:

- W_i is the benefit value of node x_i;
- v_i is the total number of neighbors of node x_i;
- $e_i = \frac{RE_i}{IE_i}$, where RE_i is the remaining energy and IE_i is the initial energy of node x_i, respectively. The e_i value is the normalized remaining energy of node x_i (an integer value between 0 and 1);

- $s_i = \frac{1}{Limited\_RSSI} \sum_{j=1}^{n_i} RSSI_j$, where $RSSI_j$ is the Received Signal Strength Indicator (RSSI) among node x_i and its neighbors nodes x_j, and the constant *limited_RSSI* is the minimum value of RSSI for an adequate communication (–87 dBm (SRINIVASAN; LEVIS, 2006)).
- $H_i = (1 - \alpha) \times H_i + \alpha \times S_R$ is the history of successful transmission rates adjusted at each beacon interval. The value of variable α is adjusted according to each case, being defined between $0 < \alpha \leq 1$; variable S_R is equal to 1 in case of a successful transmission of node x_i or 0 otherwise;
- β^n, β^e, β^s, β^H are the weights of each parameter for the objective function.

The selection of the set of relay nodes is based on the gain (W_i) provided by a node $x_i(i = 1, 2, \ldots, n)$ up to its set of neighbors $x_j(j = 1, 2, \ldots, v_i)$. In order to select the minimum number of relay nodes, ensuring at the same time that every node has a reachable relay, an optimization problem is formulated as follows:

$$\begin{aligned} & minimize \sum_{i=1}^{n} W_i y_i, && \text{(3a)} \\ & subject \quad to: \quad \mathtt{A} y \geq b, && \text{(3b)} \\ & \mathtt{C} y = d, && \text{(3c)} \\ & y_i \in \{0, 1\}. \end{aligned}$$

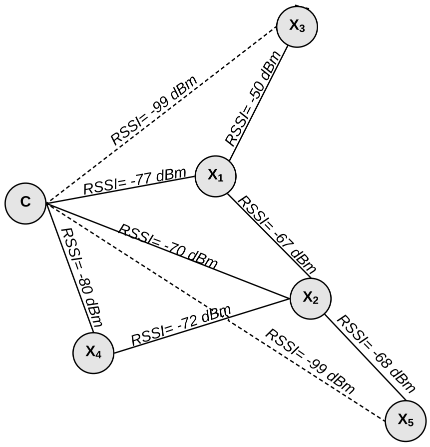
In the constraint presented in Equation (3b), $\mathtt{A}$ is the adjacency matrix of order $n \times n$, where its element $a_{i,j} = 1$ if node x_i is a neighbor of node x_j and $a_{i,j} = 0$, otherwise. Matrix $\mathtt{A}$ is formed on the coordinator node based on the list of neighbors sent by each node of the network. Therefore, whenever the list of neighbors of a node x_j has not been received by the coordinator, all elements of row j of matrix $\mathtt{A}$ will be equal to zero; y is a vector of order nx1, where y_i will be equal to 1 when node x_i is selected as relay and 0 otherwise; and b is a vector whose b_i value has been defined as 1, representing the minimum number of relay nodes of each node x_i. As a consequence, based on the variables of the problem $y_i \in \{0, 1\}$, the ORST scheme can be considered as a Binary Integer Problem (BIP).

The constraint presented in Equation (3c) is determined by the coordinator node, where matrix $\mathtt{C}$ represents the set of nodes that do not have an adequate communication link with the coordinator node. Each row of matrix $\mathtt{C}$ represents a node x_i that does not communicate directly with the coordinator and each column represents a node that is able to hear this node. In this case, d will be equal to 1, in order to guarantee that at least one of these nodes will cooperate with node x_i.

Figure 10 illustrates the operation of these constraints, where nodes are considered to be neighbors if RSSI $\geq$ –87 dBm. In this case, the coordinator C (represented in row 1 of matrix A) received the list of neighbors of nodes x_1, x_2 e x_4. Thus, matrix $\mathtt{A}$ is equal to:

Figure 10 – Network with the neighborhood of each node.

$$\mathtt{A} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$



Source – Author.

As the coordinator node did not receive the list of neighbors from nodes x_3 and x_5, rows 4 and 6 are equal to zero. Constraint (3b) is set up as follows: x_1: $y_1 + y_2 + y_3 \geq 1$, x_2: $y_1 + y_2 + y_4 + y_5 \geq 1$ and x_4: $y_2 + y_4 \geq 1$, where each x_i is composed of the neighborhood of node x_i. This constraint ensures that each node will have at least one relay among its neighbors, that is, one or more binary variables y_i must be equal to one ($y_i = 1$). Constraint (3c) is formed by the nodes from which the coordinator did not receive the list of neighbors, that is, x_3 and x_5.

For each node x_i that does not communicate with the coordinator, a constraint is set up with the binary variables of nodes x_j that listen to node x_i, and one of these nodes must be relay. That is, the binary variable of this node must be equal to one ($y_i = 1$). In this case, the list of nodes that listen to nodes x_3 and x_5 is formed only by, respectively, nodes x_1 and x_2, being: x_3: $y_1 = 1$ and x_5: $y_2 = 1$. The benefit function will prioritize the selection of nodes with smaller profit that respect the constraints. The relative weights of the parameters of the benefit function can be modified, according to the values of β^v, β^e, β^s e β^H.

As previously mentioned, the ORST proposal represents a Binary Integer Problem. This type of problem can be solved by a Branch and Bound approach, which in the worst case runs on exponential time (WOLSEY, 1998). In the simulation assessment, a solver integrated to the simulation tool has been used to solve the formulated problem for all of the analyzed scenarios. The analysis of its execution time and the comparison with other solving methods are out of the scope of this chapter.

Finally, in ORST, an adequate number of slots is reserved for the communication, enabling all nodes to transmit their messages, and also all relay nodes to retransmit their messages, at each period of the network. The proposed ORST relay selection scheme assumes that there is a limited number of relay nodes. Considering a maximum number of 100 nodes in the network, it is assumed that up to 40% of these nodes can act as relays. Therefore, a total of 140 slots should be reserved in the network for the transmission and retransmission of all messages.

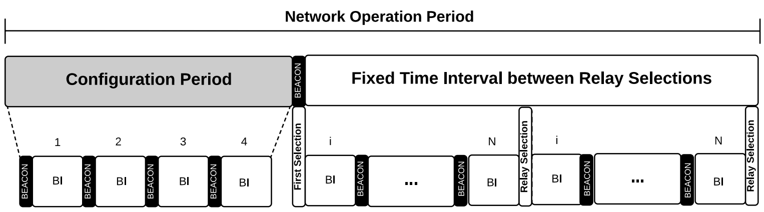
2.3.4 Relay Updating Schemes

In this chapter, two relay updating schemes have been investigated, the Periodic Relay Selection (PRS) and the Adaptive Relay Selection (ARS). The periodic relay selection scheme is an extension of preliminary work reported in Suelen Laurindo et al. (2017), where the problem of high energy consumption has been corrected by adjusting the number of BIs that the nodes need to stay awake (described in Section 2.3.2). In the adaptive relay selection scheme, a new relay selection will be triggered based on the network success rate. That is, if the success rate is smaller, the relay selection will be more frequently triggered, if the success rate is higher, the relay selection will be performed with a lower frequency.

Periodic Relay Selection (PRS) — In the PRS scheme, the time interval between two consecutive relay selections T_{IS} is fixed and independent of the current relay performance.

The operation of the PRS scheme is illustrated in Figure 11, where the first four BIs (Beacon Intervals) characterize the configuration phase. At the end of this phase, the first relay selection will occur, the time interval for the subsequent selections being periodic.

Figure 11 – Operation of the periodic relay selection scheme.



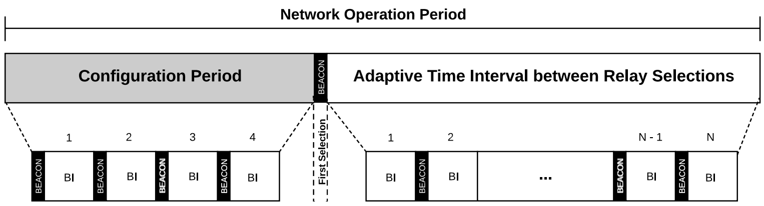
Source – Author.

In a network with a static topology, T_{IS} may have a higher value, considering that the set of nodes remain in the same localization. On the other hand, when considering dynamic topology networks, lower values for T_{IS} allow new relays to be selected sooner if any modification in the network occurs. Nevertheless, the proposed scheme does not take into account the performance of the current set of relays. That is, even if the network success rate is 100%, a new selection will be triggered every T_{IS}.

Adaptive Relay Selection (ARS) — The ARS scheme allows an adaptive selection of the relay nodes for networks with dynamic topology, where several nodes may join or leave the coverage area of the coordinator node. In the proposed scheme, the interval between two consecutive relay selections T_{IS} is dynamically determined, according to the success rate of the network. The reason for considering the network's success rate as a whole and not just the link between the cooperating nodes and the coordinator is that the communication of all these nodes with the coordinator can remain acceptable over time. However, with nodes joining and/or leaving the coverage area of the coordinator node, it may occur that other nodes require cooperation and the current set of relays nodes may no longer be adequate. Thus, if all messages successfully reach their destination, there is no need for a new relay selection. However, if the success rate decreases, it means that the current set of relay nodes is not meeting the communication requirements of the network and a new relay selection procedure must be triggered.

The operation of ARS is illustrated in Figure 12. When the network starts, as in the periodic scheme, a configuration phase of four BIs is also considered. At the end of the configuration phase, the first relay selection will occur, and the interval for the subsequent selections will be dynamically determined according to the success rate of the network.

Figure 12 – Operation of the adaptive relay selection scheme.



Source – Author.

Before sending a new beacon, the coordinator node checks the network success rate. If it has increased more than δ, the interval between selections (T_{IS}) is incremented by one BI, respecting an upper bound of 10 BIs. However, if the success rate has decreased more than δ, the T_{IS} value is reduced by half of its current value, respecting a lower bound of 2 BIs. If the success rate remains δ-stable, the T_{IS} value will be kept at its previous value. The maximum time interval between two consecutive selections has been set to 10 BIs to ensure the responsiveness of the network. These up and down selection rates were obtained by simulation using the OMNeT++/Castalia simulator.

After defining the T_{IS} value, the coordinator node needs to inform the nodes about when the relay selection will be performed. Thus, the next two beacons will be used to implement the relay selection. Firstly, the coordinator informs the network nodes through the first beacon, then notifying them to listen to their neighbors. In this way, the nodes know that they need to listen and update their list of neighbors during this BI. In the next BI, each node will send its updated list of neighbors piggybacked with the data message. After receiving the message from the nodes containing the list of neighbors, the coordinator will select a new set of relay nodes.

2.4 SIMULATION ASSESSMENT

The network simulation tool OMNeT++ (COMMUNITY, 2011) and the WSN framework Castalia (CASTALIA, 2006) were used to evaluate the proposed ORST scheme, with the PRS and ARS relay updating schemes. The open source Solve Library *lp_solve* (SOLVE, 2007) was used to implement the relay selection in ORST, solving the resulting optimization problem.

2.4.1 Characteristics of the Model Implementation

In Castalia, several extensions were added to the available IEEE 802.15.4 LLDN model, including the *aNumSuperframeSlots* parameter, the number of Guaranteed Time Slot (GTS) slots, and the size of Contention Access Period (CAP). This was necessary because Castalia still does not have a fully functional implementation of the LLDN communication mode.

The *aNumSuperframeSlots* parameter determines the size of the active portion of the superframe. The default value in the standard is 16. As previously mentioned, we consider 140 slots for the transmission and retransmission of messages. The superframe also considers the time corresponding to five time-slots during the CAP for node association, according to the IEEE 802.15.4 standard. This value of *aNumSuperframeSlots* parameter (145 slots) constrains the values used for both the BI and the Superframe Duration (SD). These BI and SD values define the duty-cycle of the network, that is, its periodicity and the duration of its inactive period.

2.4.2 Simulation Settings

Five simulation scenarios were defined with 21, 41, 61, 81 and 101 nodes, one of the nodes being the Personal Area Network (PAN) coordinator. Nodes were randomly deployed in an area of 50×50 m^2, with the PAN coordinator positioned in the center. The used channel model was the free space model without time-varying. Other simulation parameters are described in Table 2.

The simulation execution time was set to 450 seconds, during which the coordinator is able to send up to 50 *beacons*. The radio model used was CC2420, which is compliant with the IEEE 802.15.4 PHY Standard. For the PRS updating scheme, the interval between relay selections (T_{IS}) was defined to four BIs. This value was obtained by simulation using the OMNeT++/Castalia simulator. Thus, in the case of a modification of the network, a new selection of cooperators will be quickly initiated. To reduce the statistical bias, each simulation was performed 60 times with a confidence interval of 95%.

Table 2 – Simulation setting.

Parameters	Values	Parameters	Values
Node distribution	Random with coordinator in center	Beacon Order (BO)	6
Radio	CC2420	Superframe Order (SO)	4
MAC layer	IEEE 802.15.4	β^n	0.5
Number of superframe slots	145 (5 are used by the CAP)	β^e	1.5
Data rate	250 kbps	β^s	1.0
Initial energy per nodo	18,720 J	β^H	1.5
TxOutputPower	0 dBm	T_{IS}	4 (for PRS)

Source – Author.

The simulations were performed considering two modes of operation. Firstly, a static topology has been considered, where all nodes remain connected to the network until the end of the simulation. Then, a dynamic topology was also considered, where only 50% of nodes were associated with the network at time zero and the remainder were subsequently associated in groups of 5 by 5 nodes. The first group at time instant 50 s and then all the other groups every 30 s. Considering the scenario with the highest number of nodes (100 nodes), after 320 s, all nodes were associated. Later, from the time instant 320 s of simulation, 20% of the nodes of the network randomly left the coverage of the coordinator node. This leaving operation was performed in groups of four nodes, every 10 s of simulation. Finally, all nodes again joined the network, in the same order they have left (groups of 4 in 4), from the time instant 350 s of simulation, respecting an interval of 10 s for each group, except for the case of the network with 100 nodes, where only 10% of the outgoing nodes returned.

The dynamic topology mode was designed to force the list of neighbors to undergo multiple changes during the simulation time, in order to assess the reliability of the relay selection procedure.

2.4.3 Compared Techniques

To validate the relevance/pertinence of the proposed PRS and ARS relay updating schemes, in addition to comparing the two techniques among themselves, their performance was also compared to three state-of-the-art techniques: Completely Random selection (WILLIG; UHLEMANN, 2012), Random selection Around the Coordinator (ETEZADI et al., 2012) and Opportunistic selection (VALLE, Odilson T et al., 2016). This subsection briefly describes these three state-of-the-art relay selection techniques.

Completely Random (CR) — A totally random technique (WILLIG; UHLEMANN, 2012) was considered for the selection of the relay nodes, without considering the quality of the communication link with the coordinator. To determine the number of relay nodes, it must be verified how many nodes are associated with the network. If the number of associated nodes is smaller than the maximum number of relays (n_r) (defined in Section 2.3), then the maximum number of relays is the number of associated nodes. Otherwise, a random number of relays between 1 and the maximum number of relays is selected (Equation (4)):

$$numCoop = random[1, n\_r]. \tag{4}$$

In this case, the set of *numCoop* relay nodes is randomly selected.

Random Around the Coordinator (RAC) — The random selection technique is a simple technique (ETEZADI et al., 2012) that randomly selects relay nodes that are closely located around the coordinator, by using the RSSI metric as the selection criterium. For the selection of cooperating nodes, the signal strength, between the node and the PAN coordinator was lower bounded to –87 dBm.

At each selection, the maximum number of relays (*n_r*) is determined by Equation (5):

$$n\_r = min(min(n\_cl, n\_ncl), n\_cm), \tag{5}$$

where:

- *n_cl* is the number of nodes that the coordinator is able to listen, considering the RSSI lower bound of –87 dBm;
- *n_ncl* is the number of nodes that the coordinator is not able to hear;
- *n_cm* is the maximum number of relay nodes that can be selected; this upper bound was set to 40, as previously mentioned in Section 2.3.

Equation (5) determines the maximum number of relay nodes, according to the smallest value between *n_cl* and *n_ncl* to prevent a large number of relays from being selected. The second part of the equation ensures that if the two cited variables are greater than 40, then this upper bound value is selected. In this way, the number of selected relays will always be a value between 1 and *n_r* (Equation (4)).

Finally, the set of *numCoop* relay nodes is randomly selected among nodes that have a good communication ratio with the coordinator.

Opportunistic — Finally, the technique proposed by Odilson T Valle et al. (2016) presents an opportunistic selection of relay nodes. In this technique, the number of relay nodes is determined according to the network error rate. As the error rate can quickly fluctuate over time, an exponentially weighted moving average is used, which keeps the "memory" of the last instances. The calculation of the relative weights between the last measured instance is done using two constants: α and β.

The evaluation of the new set of relay nodes is executed at each relay selection. The upper bound for the number of relays is given by the number of potential cooperative nodes (n_p). In this evaluation, the proposed technique considers the number of previous unsuccessfully transmitted messages. It involves the estimation of the number of message losses (E_L) and its standard deviation (D_L):

$$numCoop = min(n_p, (\delta \times E_L) + D_L). \tag{6}$$

The estimated number of message losses (E_L) is an exponential moving average based on a weighted combination of the previous value of E_L and the new value of S_L, which is the number of messages that have not been successfully delivered in the previous beacon interval:

$$E_L = (1 - \alpha) \times E_L + \alpha \times S_L. \tag{7}$$

The standard deviation (D_L) is an estimation of how much S_L typically deviates from E_L:

$$D_L = (1 - \beta) \times D_L + \beta \times |S_L - E_L|. \tag{8}$$

To determine the best set of relays, the authors use a predefined communication quality index (Q_i). The mean value of these two estimators is used for the Q_i calculation: the success rate (H_i) and the normalized link quality indicator (L_i). The L_i value is used mainly because it has a good correlation with the success rate. Therefore, for each node i, its communication quality index (Q_i) is evaluated at the coordinator as:

$$Q_i = \frac{H_i + L_i}{2}. \tag{9}$$

The H_i value is evaluated as:

$$H_i = (1 - \alpha) \times H_i + \alpha \times S_R, \tag{10}$$

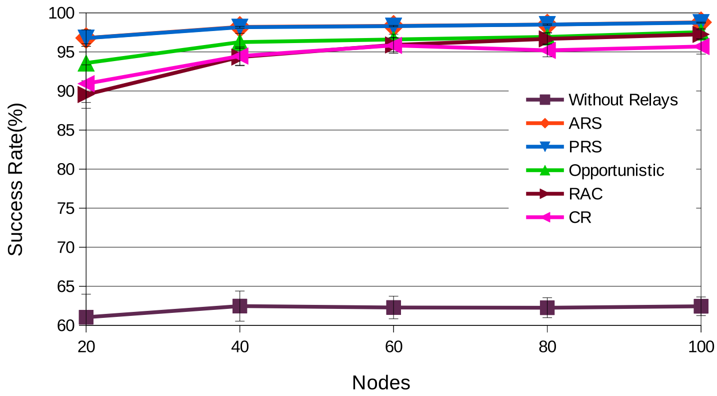
where $S_R = 1 - S_L$.

The values of δ, α and β constants for Equations (6)–(8) and (10) were tuned, and the selected value was 2 for δ and 0.2 for both α and β (VALLE, Odilson T et al., 2016). The coordinator node maintains an ordered list of Q_i values and the set of *numCoop* nodes with higher Q_i value will be elected as relays.

2.4.4 Simulation Assessment

The simulation assessment considered the following metrics: success rate, number of cooperation per node, energy consumption, number of relay selections during the system's runtime and the percentage of duplicate (useless) messages. The success rate represents the ratio between the number of sent messages and the number of messages that actually reach the coordinator. This metric considers messages transmitted in both the transmission attempt and the retransmission attempts performed by relayers. The number of cooperations represents the average number of cooperations performed per node, i.e., it is based on the number of retransmission messages sent

Figure 13 – Success Rate—static topology.



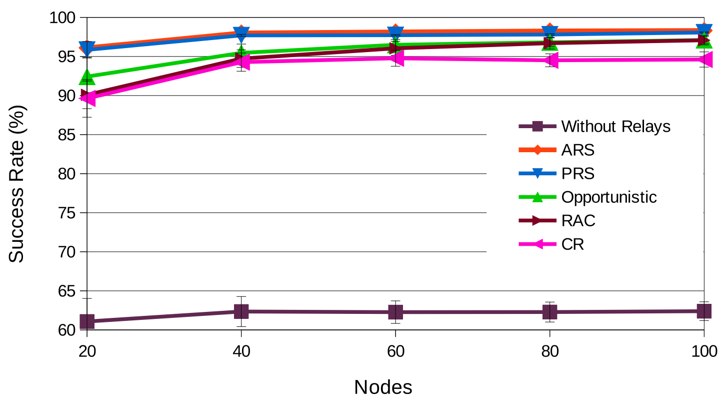
Source – Author.

by each relay node. Energy consumption represents the average amount of energy spent by each node, obtained through the resource management module available in Castalia framework. The average number of relay selections represents the average number of times a new relay selection was triggered during the simulation time. Finally, the percentage of duplicate (useless) messages represents the percentage of cooperation's messages that were not used, i.e., all messages that the relay node listened to and inserted in the cooperation message that had already arrived with success to the coordinator.

Figures 13 and 14 illustrate the success rate of all assessed relay selection techniques for static and dynamic network topologies, respectively. Both figures clearly highlight the importance of the relay retransmission techniques, as when retransmission techniques are not used ("Without Relays"), the obtained success rate is smaller than 65%. It is worth noting that, despite the good performance of state-of-the-art techniques, the proposed selection schemes ARS and PRS present better performance, with success rate above 95%, independently of the number of nodes and topology. For the static topology, ARS and PRS have basically the same probability of successful transmissions. For the dynamic topology network, where nodes may join/leave the network, the ARS updating scheme has a slightly better performance than the PRS. Among the state-of-the-art techniques, the Opportunistic scheme presented the best success rate results, and the Completely Random (CR) scheme presented the worst results.

Figures 15 and 16 present the energy consumption in the network with static and dynamic topology scenarios, respectively. For both scenarios, the ARS technique presents the lower energy consumption, followed by the PRS and RAC techniques in the static topology scenario and the PRS technique in the dynamic topology scenario. The ARS technique presents an energy consumption lower than PRS because nodes do

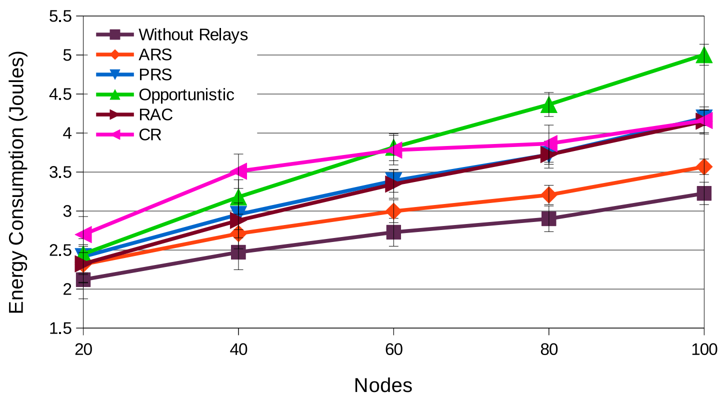
Figure 14 – Success Rate— dynamic topology.



Source – Author.

not need to listen and update the list of neighbors so frequently, it only being necessary to update the list of neighbors whenever a new relay selection is really necessary.

Figure 15 – Energy consumption—static topology.

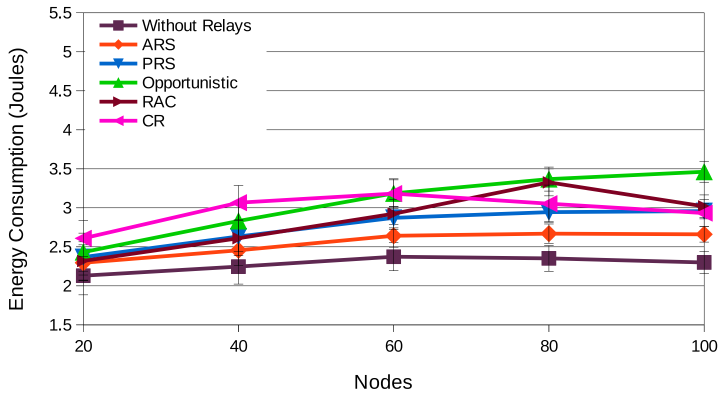


Source – Author.

As the previous results have already shown the advantage of using cooperative diversity techniques, the next results discuss only metrics related to cooperation techniques. Figures 17 and 18 present the average number of relay selection operations in the network with static topology and dynamic topologies, respectively. It is possible to verify that the average number of relay selections made by the ARS technique in the static scenario is much smaller than in the dynamic scenario. This is due to the fact that the success rate of the network is just slightly changing during the simulation time for the static topology. On the other hand, for the dynamic topology scenario, the network suffers a significant number of modifications, imposing a higher rate for the relay selections.

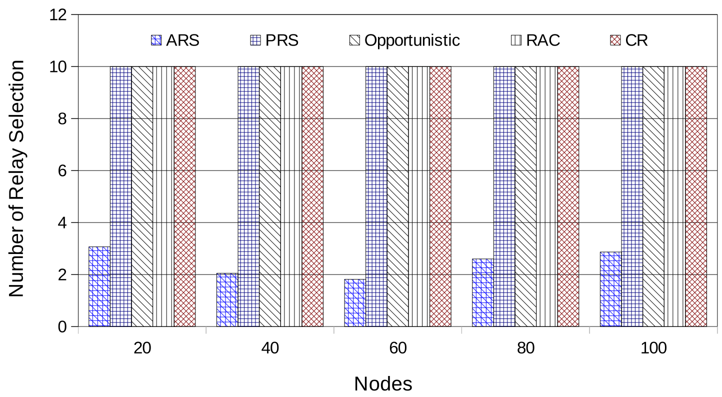
As dynamic topology scenarios, where nodes can dynamically join and leave the network, are more challenging for the retransmission techniques, for the sake of

Figure 16 – Energy consumption—dynamic topology.



Source – Author.

Figure 17 – Average Number of Relay Selection—static topology.



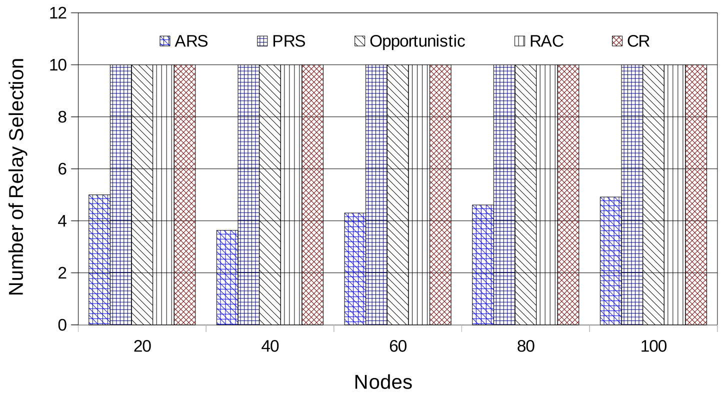
Source – Author.

simplicity, hereafter in this chapter, only the results for the dynamic topology scenarios will be represented. Figure 19 illustrates the average number of cooperation per node. As expected, the ARS and PRS have the smaller number of cooperations. This behavior is a direct consequence of the smaller number of selected relays, due to the employed optimization technique. This metric associated with the success rate allows for verifying how many cooperations per node were required to reach the adequate level of success rate. For instance, in the ARS and the PRS techniques when the number of nodes is 20, the average number of cooperations performed per node is about 6 reaching a success rate about 96%, while, in the CR technique, the average number of cooperations performed per node is about 20 and the success rate is about 89%.

The CR technique is the one that presents the highest average number of cooperations per node in scenarios with fewer nodes (less than 60) because the probability that the same nodes will be selected as a relay is greater.

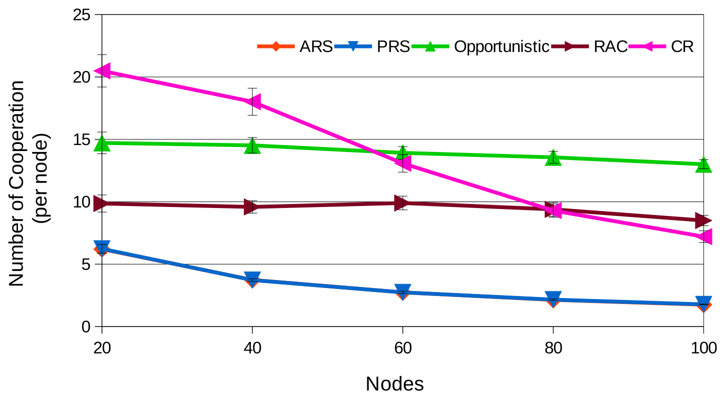
Figure 20 presents the percentage of useless retransmission messages, where

Figure 18 – Average Number of Relay Selection—dynamic topology.



Source – Author.

Figure 19 – Average number of cooperation exchanges for a dynamic scenario.

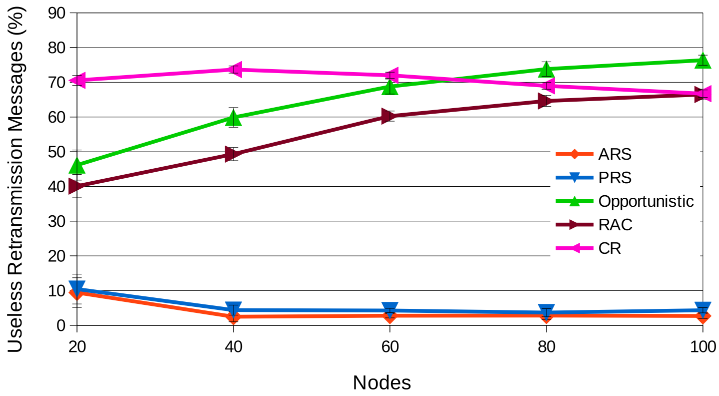


Source – Author.

ARS is the relay selection technique that presents the smallest value, followed by PRS. That is, in both schemes, the set of relay nodes has been optimized by the proposed optimization procedure. Other techniques have more than 50% of useless retransmission messages, i.e., these messages were already received in transmissions or retransmissions performed by other nodes. For the case of the Opportunistic technique, the justification is that the number of relays is determined by the success rate of the network. Therefore, whenever the success rate decreases, a higher number of nodes will be selected. For the case of the Completely Random (CR) and Random Around the Coordinator (RAC) techniques, it occurs because they do not consider any criteria to determine how many nodes will be selected as relays nodes, selecting an unnecessary number of relay nodes. Again, among the state-of-the-art techniques, the RAC scheme presented the best results.

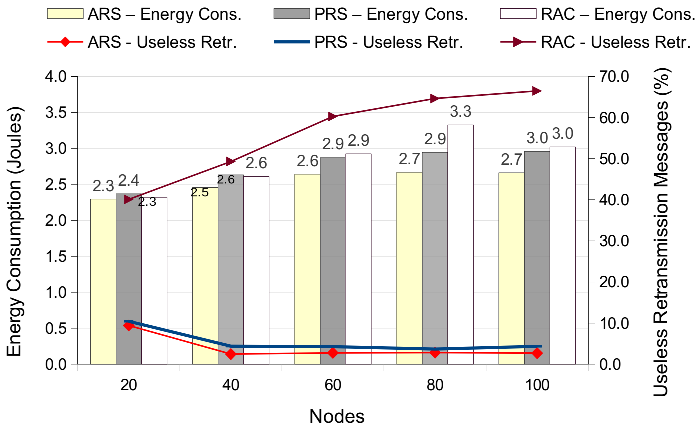
Finally, Figure 21 illustrates a correlation between energy consumption and useless retransmission messages of the two schemes proposed in this work (ARS and

Figure 20 – Percentage of useless retransmission messages for a dynamic scenario.



Source – Author.

Figure 21 – Correlation between energy consumption and useless retransmission messages.



Source – Author.

PRS) and RAC, which was the-state-of-the-art technique that obtained the best performance. It is possible to observe that ARS is the technique that presents the lowest energy consumption, being one of the main reasons the transmission of the smallest number of useless messages.

2.5 CONCLUSIONS

Smart objects with communicating actuating capabilities are becoming more common, bringing the IoT paradigm closer, where sensors and actuators are completely integrated into the environment and communicate transparently. The use of WSNs have been pointed out as a promising technology for the IoT paradigm. However, and due to the inherently unreliable wireless communication medium, new communication schemes are needed to increase WSNs' communication reliability.

Cooperative diversity techniques can be used to improve the reliability of WSN communications. Nonetheless, an important step for the use of these techniques is to perform an adequate relay selection. This chapter focuses on the adequate selection of relay nodes. It proposes the use of the ORST scheme, whose target is to adequately select relay nodes without generating overheads or excessive energy consumption.

The proposed ORST relay selection scheme was applied using two relay updating schemes: PRS and ARS. Aiming to investigate which is more adequate to improve the reliability of WSN, the following metrics were considered: success rate, average number of cooperation exchanges per node, percentage of duplicate (useless) messages, energy consumption and average number of relay selection.

Both PRS and ARS schemes were assessed by simulation against other three state-of-the-art relay selection techniques. The performed simulation assessment highlighted that ORST demonstrates a significantly improved reliability behavior for both updating schemes: PRS and ARS. The proposed ARS relay selection technique outperformed the other state-of-the-art techniques, increasing the message transfer success rate, decreasing the average number of cooperation exchanges per node, and presenting the smallest percentage of useless retransmission message and a lower energy consumption. That is, the use of an optimized relay selection technique is able to more efficiently select the set of relay nodes.

As future work, we intend to assess the implementation feasibility of the proposed scheme using available Commercial Off-The-Shelf (COTS) WSN nodes. In this case, it will be necessary to evaluate and to optimize the resolution method implemented to solve the generated optimization problem, as typical WSN nodes have scarce available resources—mainly related to reduced footprint memory and processing capabilities.

3 IMPACT OF PARAMETERS ON THE OBJECTIVE FUNCTION

In Chapter 2, we presented a new relay selection technique in which the set of relays was chosen by modeling the selection as an optimization problem, using an objective function involving four parameters (number of neighbors; remaining energy; Received Signal Strength Indicator (RSSI) between node x_i and its neighboring nodes x_j; and the history of successful transmission rates) to select the cooperating nodes. This chapter proposes an improvement to the technique set out in the previous chapter. An analysis of the impact of each of the parameters used in the objective function is presented. As a result of this analysis, an update to the objective function is proposed. This chapter contains a transcription of a paper entitled "Multicriteria analysis to select relay nodes in the ORST technique" presented at the *Conference on Ad Hoc Networks and Wireless* (LAURINDO, S. et al., 2019). The following paragraph presents a transcript of the abstract of the cited paper.

Cooperative diversity techniques can be used to improve the reliability of communication in Wireless Sensor Networks (WSN). Typically, these techniques use relay nodes to retransmit messages that otherwise would not be heard by their destination nodes. Thus, the relay selection techniques are fundamental to improve WSN's communication behavior. However, to perform the adequate relay selection, it is necessary to identify which are the most relevant parameters for the operation of the network and analyze their impact when used in the relay selection, that is, it is necessary to define which are the best parameters to use as selection criteria. In this context, this chapter performs an analysis of the impact of each of the parameters used to perform the relay selection in the Optimized Relay Selection Technique (ORST). This analysis was assessed by simulation using the OMNeT++ tool and the WSN framework Castalia. It was considered a set of parameters, aiming to identify their relevance and possibly optimize the objective function used in this technique. Simulation results show that the objective function can be optimized considering a small number of parameters to perform the relay selection.

As described in the abstract above, this study aimed to determine the most important parameter for use in the objective function. Although it was known that four parameters had particular relevance to the operation of the network, the specific relevance of each in terms of the objective function was unknown. The idea for this assessment arose from a discussion of which had the greatest impact on the objective function, or in other words, which of these parameters needed to be improved to guarantee that a given node was a suitable cooperator.

As a result of this assessment, the objective function was updated to include only a single parameter, the amount of energy remaining in each node; this made it simpler without affecting the performance of ORST. Since the optimization model of the problem already takes into consideration the number of neighbors for each node in the network and the issue of which nodes do not communicate with the coordinator node, as part of the restrictions on the problem, there is no need for the objective function to take this information into account.

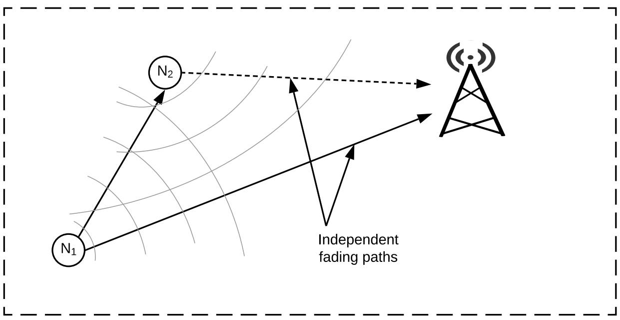
This chapter is structured as follows. Section 3.1 presents an introduction to cooperative diversity and the importance of using a suitable relay selection technique. Section 3.2 reviews the state-of-the-art in regard to relay selection techniques for WSNs. Section 3.3 briefly describes ORST and analyzes the criteria used to perform the relay selection, with a particular focus on improvements in the reliability of communication in WSNs. Section 3.4 presents the simulation assessment and the results. Finally, the conclusions are presented in Section 3.5.

3.1 INTRODUCTION

In wireless sensor networks (WSN), the wireless communication medium is inherently unreliable, where messages may be lost due to electromagnetic noise, other devices that operate in the same frequency range or obstacles between nodes. This problem may lead to a severe reduction of achievable throughput (IQBAL et al., 2017). A possible solution to improve the reliability of WSNs is by providing multiple paths to transmit data from the source to the destination node; this type of communication is called cooperative diversity (PHAN et al., 2009).

Cooperative diversity considers the existence of nodes that will cooperate with the transmitter-receiver pair, in order to increase the chances of the sent message being received in the destination (KHAN; KARL, 2014). Thus, nodes with a single antenna can share their antennas and produce a virtual multiple-antenna transmitter, as illustrated in Figure 22, where there are two source nodes (N_1 and N_2) and one destination node. As the fading paths from two source nodes are statistically independent, the use of the cooperative diversity generates spatial diversity, improving the communication in the network. In the cooperative diversity technique, nodes that share their antennas retransmitting messages assume the role of relay nodes (NOSRATINIA et al., 2004).

Figure 22 – Cooperative communication.



Source – Author.

This behavior provided by cooperative diversity allows better use of the broadcast nature of wireless transmissions, in which nodes are usually able to hear messages transmitted by their neighborhood (LIANG et al., 2009). However, in a conventional communication in WSN, nodes discard listened messages that are not intentionally sent to them (KHAN; KARL, 2014; NOSRATINIA et al., 2004).

In the cooperative communication research area, the adequate relay selection is a key problem that must be addressed. The performance of cooperative retransmission techniques depends heavily on the efficiency of the process used to select one or more relay nodes (AZIZ; GHANI, 2019; VALLE, Odilson T et al., 2012, 2018).

An adequate relay selection will allow messages from a source node to reach the destination node even though the quality of the communication channel is degraded, as the source node will have a relay that will aid retransmitting the message to the destination.

In order to perform the relay selection, it is necessary to consider some important criteria for the operation of the network. However, there are few state-of-the-art works evaluating a different set of criteria when selecting the relay nodes. In a recent work (LAURINDO, S. et al., 2018), we proposed a relay selection technique to be used in WSN, named Optimized Relay Selection Technique (ORST). In this technique the selection of a set of relay nodes is based on multiple criteria, namely: the number of neighbors of the candidate node, their remaining battery energy, the quality of the communication link between the candidate node and its neighbor nodes (by using the RSSI) and the success rate's history in recent node transmissions. The ORST technique was formulated as an optimization problem, using a specifically selected objective function. When the technique was proposed, the aim was to consider all the criteria that are highly relevant for the operation of the network, but without analyzing the real impact of each criterion upon the network performance.

This chapter performs an analysis considering each of the parameters individually and combining them, aiming to identify the importance of each one of them and possibly to optimize the objective function previously used (LAURINDO, S. et al., 2018).

3.2 RELATED WORK

In spite of the existence of multiple works related to relay selection techniques in the literature, many of these works do not give due relevance to the criteria used for the relay node selection. Nevertheless, the performance of the relay selection technique can be highly improved if the selected criteria are adequate.

In Zhu et al. (2018), the authors proposed a technique to select the best relay to cooperate in the transmission from the source node to the destination node. The relay that has the highest instantaneous SNR (Signal-to-Noise Ratio) of the combiner output at the destination is selected as the cooperative relay. The authors proposed in

Yu Zhang et al. (2017) a relay selection scheme to maximize the network lifetime. The authors considered the energy consumption rate in transmission and residual energy of each sensor node as criteria for the relay selection. The nodes that maximize the network lifetime are selected as relay nodes.

In Jianrong Bao et al. (2017), a relay selection technique using both AF (Amplify-Forward) and DF (Decode-Forward) protocols is proposed, which selects the AF or DF schemes to forward signals adaptively according to the CSI (Channel State Information) information. If the channel status of link $Source - Relay_i$ is good enough for the relay to decode the source information, the DF protocol is selected to forward signals in the relay. Otherwise, the AF protocol is selected. The relay selection occurs in the destination node based on the SNR value. The node that maximizes the SNR in the destination node is selected as relay and notified to forward the source information.

In Senanayake et al. (2018), a relay selection based on the SNR is designed. The aim is to maximize the minimum received SNR for all users. The authors considered a multi-user multihop relay network where each hop is equipped with multiple relays that assist users to communicate with their designated destinations.

In Pham and Dong Seong Kim (2016), the authors presented a relay selection technique that aims to increase the network lifetime and to improve its packet delivery rate. The node with the highest residual energy level e_i^m and the lowest Energy-Per-Bit (EPB) value E_b will be the relay node. The authors proposed the calculation of a weight w for each cooperating node r, as: $w_r = \frac{e_i^m}{E_b}$, where, the relay with the highest weight value is selected.

In Cheikh et al. (2017), a relay technique that selects the node that minimizes the energy consumption per bit in transmissions between source and relay nodes and between relay and destination nodes is presented. The nodes that listen to the RTS (Request to Send) and CTS (Clear to Send) messages (between source node and destination node) estimate the channel gain and the desired transmission power to reach the target BER (Bit Error Rate). Using an optimization strategy, the destination node selects the relay node, and signals through the beacon message which was the selected node.

In Jun Wang and Qilin Wu (2017), it is proposed a technique that selects as relay the node that maximizes the number of packets successfully transmitted. It was considered that a transmission fails when the SNR signal arriving at the destination node is less than a predetermined threshold. This proposal considered that a data packet of size L is transmitted with an R rate and that each node can determine the number of packets successfully transmitted, called K. Thus, using an optimization technique, each node maximizes the K value, the node that presents the highest value should be selected as a relay.

In Odilson T Valle et al. (2016), the authors proposed a technique that considers the history of successful transmission rates and the LQI between each node and the coordinator. Nodes presenting the highest average between the history of successful transmissions and LQI will be selected as relays. The number of relay nodes is dynamically defined according to the percentage of message losses in the network.

Most of the works found in the state-of-the-art use as criteria of selection quality estimators based on hardware (LQI, RSSI,CSI, etc) (ZHU et al., 2018; BAO, J. et al., 2017; SENANAYAKE et al., 2018; CHEIKH et al., 2017; VALLE, Odilson T et al., 2016). However, these metrics consider only the received frames. Thus, if a radio link presents excessive losses, the quality estimators may overestimate the quality of the link.

It is worth to mention that the different combination of relay selection criteria, how they were modeled and what parameters they use, directly impacts on the relay selection performance. However, analyzing the state-of-the-art approaches, it is possible to attest they generally do not justify their choices.

3.3 RELAY SELECTION TECHNIQUE

The Optimized Relay Selection Technique (ORST) is a centralized technique recently proposed in Suelen Laurindo et al. (2018). This technique considers an IEEE 802.15.4 network operating in time-slotted and beacon-enabled modes. It allows the adaptive selection of relay nodes in dynamic networks, where nodes may randomly leave/join the coverage area of the coordinator node. The time interval between two consecutive relay selections is dynamically determined, according to the message transfer success rate. If all messages successfully reach their destination, there is no need for a new relay selection. If the success rate decreases, it means that the current set of relay nodes is not meeting the communication requirements and a new relay selection procedure must be performed. This behavior may be a consequence of: relay nodes that left the coordinator coverage; new nodes that joined the network and there are no enough relay nodes; or an abrupt increase of the interference level in the network.

The ORST scheme aims to find a set of relay nodes $S^{\star} = \{y_1, y_2, \dots, y_m\}$ among a set of nodes $X = \{x_1, x_2, ..., x_n\}$ in WSNs, ensuring two conditions: 1) each node x_i ($1 \leq i \leq n$) is covered by at least one relay node, 2) the sum of the weights (W_i) of the relays is minimized. In this scheme, x_i is used as node identifier, n is the total number of nodes in the network, m is the total number of relay nodes and $S^{\star} \subseteq X$, i.e. relays are selected in the same set of nodes, transmitting not only their own data, but also cooperating by retransmitting data from other nodes. There is one node called coordinator in the WSN (C).

This technique was designed as an optimization problem using an objective function. The objective function (Equation 11) takes into consideration the number of neighbors that each node can hear (v) (RSSI ≥ -87dBm (SRINIVASAN; LEVIS, 2006)), the available energy in the nodes (e), the quality of communication between the source node and the candidate relay node (s), as well as the history of successful transmission rates (H) of node x_i. These parameters were selected because they are highly relevant to the operation of the network. For instance, the residual energy of the nodes is an important parameter, considering that if a node has a low battery level it will stop being a promising candidate because soon it will exhaust its own energy resources. The number of neighbors that each node has is also a parameter that must be considered, since if a node does not have neighbors, it does not make sense to select it as a relay. The quality of the channel between the source and the relay nodes is another important parameter, because it allows knowing if there is a good communication link between these nodes, ensuring that the relay node correctly receives messages to be retransmitted. And finally the history of successful transmission rates is an indication that the selected node has a good communication link with the destination node, ensuring that messages sent by this node will correctly arrive at their destination. Combining these parameters as the selection criterion, the aim is to ensure that the appropriate nodes are selected as cooperating. Each node x_i will calculate its objective function value W_i and this information will be sent to the coordinator.

$$W_i \doteq \left(\frac{\beta^v}{v_i} + \frac{\beta^e}{e_i} + \frac{\beta^s}{s_i} + \frac{\beta^H}{H_i} \right) \tag{11}$$

where:

- W_i is the objective function value of the node x_i;
- v_i is the total number of neighbors of node x_i;
- $e_i = \frac{RE_i}{IE_i}$, where RE_i is the remaining energy and IE_i is the initial energy of node x_i, respectively. The e_i value is the normalized remaining energy of node x_i (a real value between 0 and 1);
- $s_i = \frac{1}{Limited\_RSSI} \sum_{j=1}^{n_i} RSSI_j$, where $RSSI_j$ is the Received Signal Strength Indicator (RSSI) among node x_i and its neighbors nodes x_j, and the constant *Limited_RSSI* is the minimum value of RSSI for an adequate communication (–87 dBm (SRINIVASAN; LEVIS, 2006));
- $H_i = (1 - \alpha) \times H_i + \alpha \times S_R$ is the history of successful transmission rates adjusted at each beacon interval. The value of variable α is adjusted according to each case, being defined between $0 < \alpha \leq 1$; variable S_R is equal to 1 in case of a successful transmission of node x_i or 0 otherwise;
- β^v, β^e, β^s, β^H are the weights of each parameter for the objective function.

In order to select the minimum number of relay nodes, ensuring at the same time that every node has a reachable relay, an optimization problem is formulated as follows:

$$minimize \sum_{i=1}^{n} W_i y_i \quad (12a)$$

$$subject \quad to: \quad \mathtt{A}y \geq b \quad (12b)$$

$$\mathtt{C}y = d \quad (12c)$$

$$y_i \in \{0, 1\}$$

In the constraint presented in Equation 12b, A is the adjacency matrix of order nxn, where its element $a_{i,j}$ = 1 if node x_i is a neighbor of node x_j and $a_{i,j}$ = 0 otherwise. Matrix A is formed in the coordinator node based on the list of neighbors sent by each node of the network. Therefore, whenever the list of neighbors of a node x_j has not been received by the coordinator, all elements of row j of matrix A will be equal to zero; y is a vector of order nx1, where y_i will be equal to 1 when node x_i is selected as relay and 0 otherwise and; b is a vector whose b_i value has been defined as 1, representing the minimum number of relay nodes of each node x_i. As a consequence, based on the variables of the problem $y_i \in \{0, 1\}$, the ORST scheme can be considered as a Binary Integer Problem (BIP).

The constraint presented in Equation 12c is determined by the coordinator node, where matrix C represents the set of nodes that do not have an adequate communication link with the coordinator node. Each row of matrix C represents a node x_i that does not communicate directly with the coordinator and each column represents a node that is able to hear this node. In this case, d will be equal to 1, in order to guarantee that at least one of these nodes will cooperate with node x_i.

3.3.1 Analysis of Criteria for the Relay Selection Technique

In this chapter, the impact of each parameter to perform the relay selection, in the ORST scheme, will be analyzed. This technique considers a set of parameters that may have a significative impact upon an adequate relay selection (which are: e, v, s and H). This chapter targets to identify the relative importance of each of them and possibly to optimize the used objective function, without reducing the quality of communication.

For this analysis, new objective functions were modeled considering each of the possible combinations of parameters, as presented in Table 3.

Table 3 shows fourteen possible objective functions, where functions 1, 2, 3 and 4 consider the parameters individually, functions 5, 6, 7, 8, 9 and 10 consider combinations two-by-two and functions 11, 12, 13 and 14 consider combinations three-by-three. When performing the relay selection considering each of these functions, it will be possible to identify which parameters have the greatest influence upon the selection and thus it may be possible to simplify the problem of relay selection.

Table 3 – Objective Functions.

1	$W_i \doteq \left(\frac{\beta^H}{H_i}\right)$
2	$W_i \doteq \left(\frac{\beta^s}{s_i}\right)$
3	$W_i \doteq \left(\frac{\beta^e}{e_i}\right)$
4	$W_i \doteq \left(\frac{\beta^v}{v_i}\right)$
5	$W_i \doteq \left(\frac{\beta^v}{v_i} + \frac{\beta^e}{e_i}\right)$
6	$W_i \doteq \left(\frac{\beta^v}{v_i} + \frac{\beta^s}{s_i}\right)$
7	$W_i \doteq \left(\frac{\beta^v}{v_i} + \frac{\beta^H}{H_i}\right)$
8	$W_i \doteq \left(\frac{\beta^e}{e_i} + \frac{\beta^s}{s_i}\right)$
9	$W_i \doteq \left(\frac{\beta^e}{e_i} + \frac{\beta^H}{H_i}\right)$
10	$W_i \doteq \left(\frac{\beta^s}{s_i} + \frac{\beta^H}{H_i}\right)$
11	$W_i \doteq \left(\frac{\beta^v}{v_i} + \frac{\beta^e}{e_i} + \frac{\beta^s}{s_i}\right)$
12	$W_i \doteq \left(\frac{\beta^v}{v_i} + \frac{\beta^s}{s_i} + \frac{\beta^H}{H_i}\right)$
13	$W_i \doteq \left(\frac{\beta^v}{v_i} + \frac{\beta^e}{e_i} + \frac{\beta^H}{H_i}\right)$
14	$W_i \doteq \left(\frac{\beta^e}{e_i} + \frac{\beta^s}{s_i} + \frac{\beta^H}{H_i}\right)$

3.4 SIMULATION ASSESSMENT

A simulation assessment was performed using the network simulation tool OMNET++ (COMMUNITY, 2011) and the WSN framework Castalia (CASTALIA, 2006). The open source Solve Library *lp_solve* (SOLVE, 2007) was used to implement the relay selection in ORST, solving the resulting optimization problem.

3.4.1 Simulation Settings

Each function was performed in five scenarios defined with 21, 41, 61, 81 and 101 nodes, with one of the nodes being configured as the Personal Area Network (PAN) coordinator. Nodes were randomly deployed in an area of 50x50m^2, with the PAN coordinator positioned in the center. The used channel model was the free space model without time-varying. Others simulation parameters are described in Table 4.

The simulation execution time was set to 450 seconds, during which the coordinator is able to send up to 50 *beacons*. The radio model used was CC2420, which is compliant with the IEEE 802.15.4 PHY standard. The β^n, β^e, β^s and β^H values were obtained through experiments performed in the simulator, where values were tested in a range of 0.5 to 5 for each of the parameters. To reduce the statistical bias, each simulation was performed 60 times with a confidence interval of 95%.

Table 4 – Simulation Setting.

Parameters	Values	Parameters	Values
Node distribution	Random with coordinator in center	BO	6
Radio	CC2420	SO	4
MAC layer	IEEE 802.15.4	β^n	0.5
Number of superframe slots	145 (5 are used by the CAP)	β^e	1.5
Data rate	250 kbps	β^s	1.0
Initial energy per node	18720 J	β^H	1.5
TxOutputPower	0 dBm	T_{IS}	4 (for PRS)

Source – Author.

Simulations were performed considering a dynamic topology, where only 50% of nodes were associated to the network at time zero and the remainder were subsequently associated in groups of 5 by 5 nodes. The first group at time instant 50 seconds and then, all the other groups every 30 seconds. Considering the scenario with the highest number of nodes (100 nodes), after 320 seconds, all nodes were associated. Later, from the time instant 320 seconds of simulation, 20% of the nodes of the network randomly left the coverage of the coordinator node. This leaving operation was performed in groups of 4 nodes, every 10 seconds of simulation. Finally, all nodes again joined the network, in the same order (groups of 4 in 4), from the time instant 350 seconds of simulation, respecting an interval of 10 seconds for each group, except for the case of the network with 100 nodes, where only 10% of the outgoing nodes returned.

This topology was designed to force the list of neighbors to undergo multiple changes during the simulation time, in order to assess the reliability of the relay selection procedure.

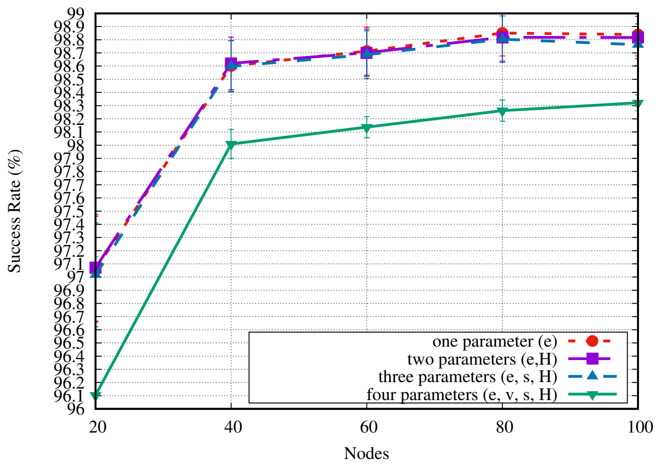
3.4.2 Simulation Assessment

It was considered the following metrics in the evaluation scenarios: success rate, number of cooperations per node, energy consumption and the percentage of duplicate (useless) messages. The success rate represents the ratio between the number of sent messages and the number of messages that actually reach the coordinator. This metric considers messages transmitted in both the transmission attempt and the retransmission attempts performed by relayers. The number of cooperations represents the average number of cooperations performed per node, i.e., it is based on the number of retransmission messages sent by each relay node. Energy consumption represents the average amount of energy spent by each node, obtained through the resource management module available in Castalia framework. Finally, the percentage of duplicate (useless) messages represents the percentage of cooperation's messages that were not used, i.e., all messages that the relay node listened to and inserted in the cooperation message which had already arrived with success in the coordinator.

Simulations were performed evaluating the fourteen objective functions. However, only the most relevant results considering one, two, three or four parameters will be presented. When considering one parameter, the best results were obtained with Energy (e) one, as presented in Equation 3 (Table 3). For two parameters, the best results were obtained with Energy (e) and History of Successful Transmission Rate (H), as presented in Equation 9 (Table 3). And the combination of three parameters that presented the greatest impact was Energy (e), the RSSI among node x_i and its neighbors node x_j (s) and History of Successful Transmission Rate (H), as defined in Equation 14 (Table 3).

Finally, these results will be compared with the results obtained with the ORST technique, where four parameters were considered together. Figure 23 presents the success rate results. It can be observed that the success rate can even be improved when the selection is performed, considering only the parameters of greatest impact. In this case, three new objective functions were advantageous when compared to the function that considers the four parameters. Among the three new modeled objective functions, we can highlight the function that considers only the energy resource as a parameter, because besides presenting a good performance in relation to the success rate is the function that best simplifies the optimization problem, due to the restricted number of variables involved in its formulation.

Figure 23 – Success Rate.

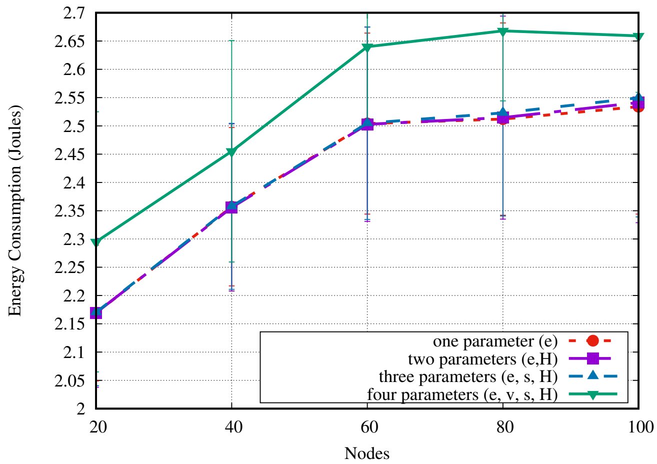


Source – Author.

Figure 24 illustrates the energy consumption of the overall network when considered different parameters for the relay selection. Again it is possible to observe that considering a smaller number of parameters can bring good results, considering that there was no increase in energy consumption.

Figure 25 presents the average number of cooperations per node. It is possible to verify that the average number of cooperations per node made by all the objective

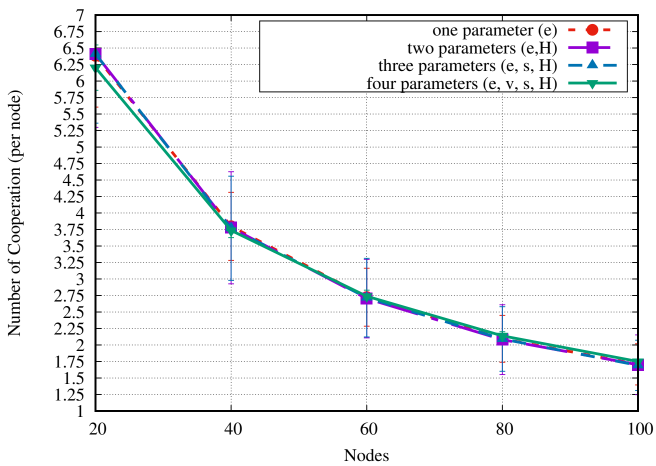
Figure 24 – Energy Consumption.



Source – Author.

functions was very similar. This behavior is a direct consequence of the smaller number of selected relays, due to the optimization technique used for all the objective functions.

Figure 25 – Average number of cooperation exchanges.

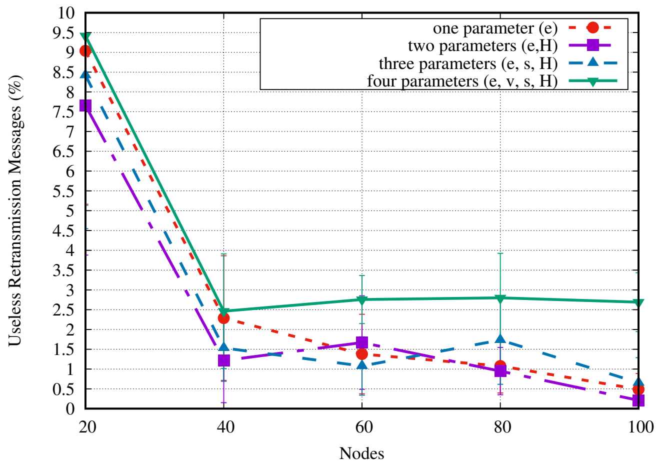


Source – Authors.

Figure 26 presents the percentage of useless retransmission messages. The objective function that combines two parameters (*e* and *H*) presented the smaller value in the scenarios with 80 and 100 nodes when compared with the objective function that considered the four possible parameters. However, when compared to other objective functions, the percentage of useless retransmissions messages is very similar.

Considering that the purpose of this analysis is to identify the relevance of each of the parameters to simplify the objective function, it can be stated that the objective function that presented best results was the objective function that considers only the energy as a selection parameter.

Figure 26 – Percentage of useless retransmission messages.



Source – Author.

The energy consumption parameter has great relevance to maximize the lifetime of the network, considering that if a node has a low battery level, it should not be selected as a relay. According to relays selection technique (Equation 12), the nodes that have a higher energy load and comply the constraints will be selected as relays. In this way, the overall purpose of not contributing for the exhausting of the energy of a relay node is fulfilled.

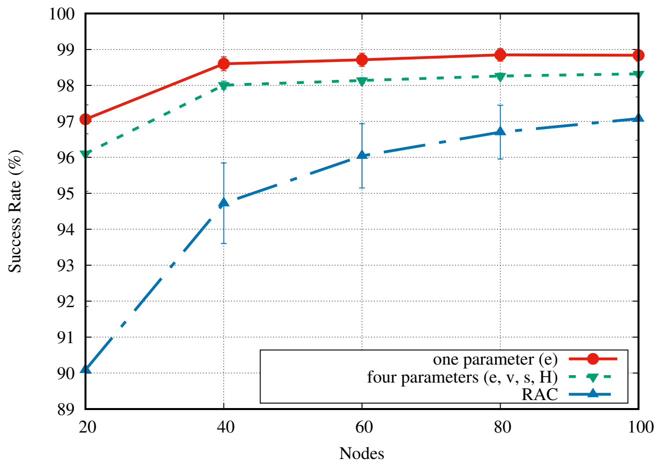
According to the results obtained from the assessment, it was possible to optimize the ORST technique's objective function, reducing the number of considered criteria. The new objective function that will be considered for the ORST technique is the one that considers only the energy resource, as presented in Equation 13:

$$W_i \doteq \left(\frac{\beta^e}{e_i}\right) \tag{13}$$

In Suelen Laurindo et al. (2018), the ORST technique was compared with three state-of-the-art techniques: *Opportunistic* (VALLE, Odilson T et al., 2016), which selects the cooperating nodes according to the network packet error rate, Random Around the Coordinator (*RAC*) (ETEZADI et al., 2012), which performs a random selection of the nodes that have an adequate communication link with the destination node and Completely Random relay selection (*CR*) (WILLIG; UHLEMANN, 2012), which performs a random selection from all the nodes of the network. In this assessment, the ORST technique outperformed the other state-of-the-art techniques. We selected the state-of-the-art technique with the best performance in this evaluation (the *RAC* technique) to compare with the ORST technique using the new objective function and the ORST technique using the old objective function.

Figure 27 presents the success rate compared to RAC technique. It is possible to observe that the ORST technique with the new objective function achieve a high level of success rate, independently of the number of nodes.

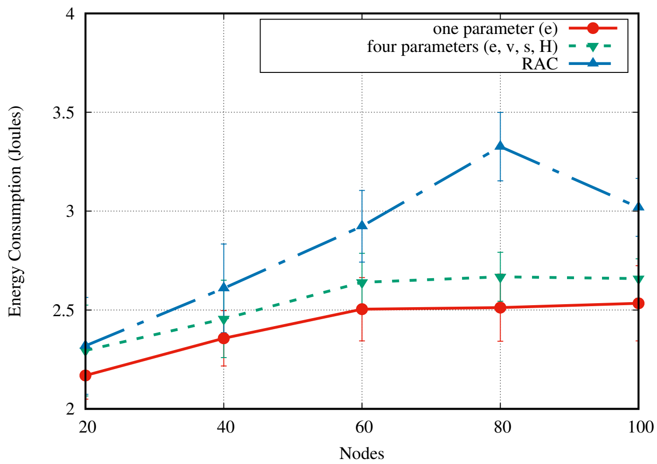
Figure 27 – Success Rate compared to RAC technique.



Source – Author.

Figure 28 illustrates the energy consumption compared to the RAC technique, again the ORST technique implementing the new objective function presented very promising results, having a lower energy consumption than the RAC technique.

Figure 28 – Energy Consumption compared to RAC technique.

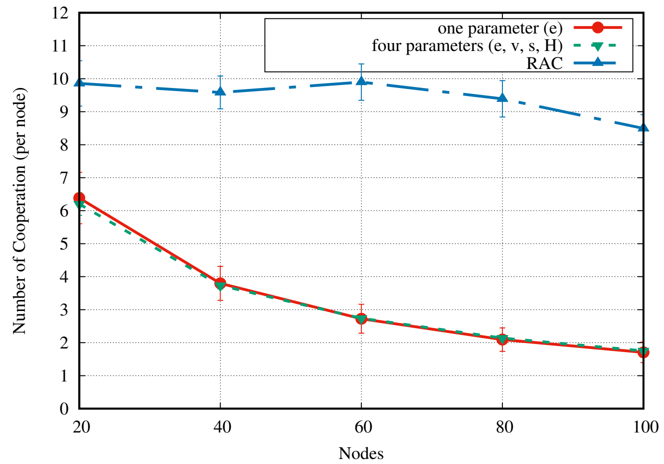


Source – Author.

Figure 29 presents the average number of cooperations per node compared to the RAC technique, this metric presents a great difference in results between the ORST (old and new) techniques and the RAC technique, being justified by the goal of the ORST technique, which is to select the lowest number of cooperating nodes, different from RAC, which randomly selects relay nodes.

Figure 30 presents the percentage of useless retransmission messages compared to RAC technique, it is possible to observe that the ORST technique with the

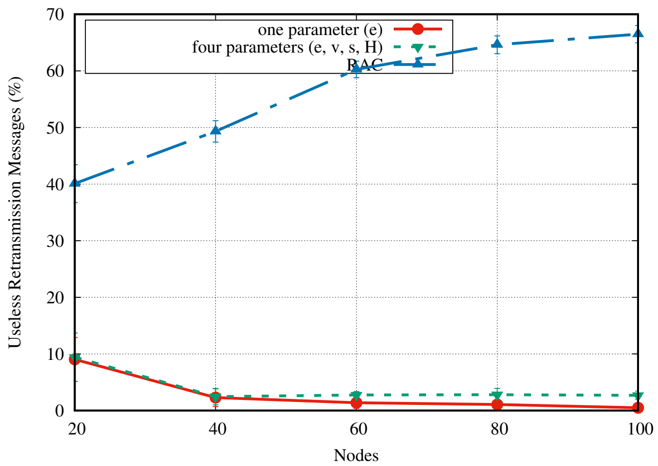
Figure 29 – Average number of cooperation exchanges compared to RAC technique.



Source – Author.

new objective function presents the lowest percentage of useless retransmission messages, approaching zero in the scenario with 100 nodes. However, the RAC technique presents a high percentage of useless retransmission messages, being greater than 60 % in the scenario with 100 nodes.

Figure 30 – Percentage of useless retransmission messages compared to RAC technique.



Source – Author.

3.5 CONCLUSION

The relays selection is a decisive step to guarantee adequate cooperative communication. With the analysis performed in this chapter, it was possible to simplify the objective function used by the ORST selection technique, reducing the number of parameters considered in this function and still properly selecting the set of relay nodes, promoting improvements in the success rate of the network.

Previously, we believed that it was necessary to consider all the parameters (v, e, s and H) in the objective function to obtain an adequate relay selection. However, what guarantees the adequate relay selection is the modeling of the problem of the ORST technique as an optimization problem (Equations 12a, 12b and 12c), that provides two guarantees. The first is that only a node that has at least one neighboring node will be selected as relay. The second is that when a node has no neighbors, it must necessarily have a relay node that listens to it. Both constraints are determined in the coordinator node, thus, the coordinator node knows the neighborhood of each node and also can define which nodes have good communication with it. In this way, considering only the energy resource in the objective function allows to complement the information that the coordinator already has, significantly improving the used relay selection technique.

4 AN ANALYSIS OF DIFFERENT SOLUTIONS TO SOLVE THE RELAY SELECTION PROBLEM

In the previous chapters, we have presented ORST and described the simplification of the objective function used in this technique. In this chapter, we will analyze alternative ways of solving the relay selection problem. To analyze different solutions, in terms of the criteria that affect the operation of the technique in real nodes, can pave the way for an analysis of the possibility of implementing ORST using low-cost commercial nodes, in view of the time constraints imposed by the communication process. We will analyze three solutions: a greedy algorithm, a genetic algorithm and the branch and bound approach. This chapter contains a transcription of a paper called "Assessment of different algorithms to solve the set-covering problem in a relay selection technique" presented at the *Conference on Emerging Technologies and Factory Automation* (LAURINDO, S. et al., 2020). The following paragraph presents a transcript of the abstract to this chapter.

The use of adequate relay selection techniques is crucial to improve the behavior of cooperation based approaches in Wireless Sensor Networks (WSN). The Optimized Relay Selection Technique (ORST) is a relay selection technique that may be reduced to the application on classic set-covering problem (SCP) to WSN. The SCP seeks to find a minimum number of sets that contain all elements of all data sets. The SCP can be solved with different types of algorithms. This chapter assesses the performance and quality of three different algorithms to solve the SCP generated by the previously proposed ORST technique, considering performance metrics relevant within WSNs context. The analysis was performed by simulation using the OMNeT++ tool and the WSN framework Castalia. The simulation results show that the branch and bound algorithm excels when compared to other state-of-the-art approaches.

As described in the abstract above, we aimed to find a better solution for solving the SCP generated by ORST. The idea for this research arose from an analysis of the computational complexity of the solution that was initially used (the branch and bound algorithm), which in the worst case is exponential. We therefore decided to analyze other solutions that are commonly used to solve the SCP. In addition to the computational complexity, we explored the average times required by these solutions to find the relay nodes in ORST during the simulations, and investigated whether the relays selected by each algorithm would ensure the proper operation of the network. The results would then enable us to determine the best solution for use with the ORST technique when this was implemented within real prototypes.

The results of this analysis allowed us to observe that although the computational complexity of the branch and bound algorithm is higher than for the other algorithms, it does not necessarily require more time to solve the SCP in the ORST technique. It is shown that in the worst case, all possible solutions will be explored by the algorithm, but that in practice, the algorithm eliminates potential solutions that will not lead to the optimal solution. Thus, in ORST, the time constraints on the operation of the network are respected and the best solution is found when the branch and bound algorithm is used. Since this is an exact method, the branch and bound algorithm searches for the optimal solution, unlike heuristic methods, which seek to find a satisfactory solution to the problem that may or may not be the optimal solution.

This chapter is structured as follows. Section 4.1 presents an introduction and shows that the performance of cooperative retransmission techniques depends heavily on the efficiency of the relay selection technique. In addition, we discuss the importance of solutions that are suitable for real-time applications such as IoT and Industry 4.0. Section 4.2 gives an overview of state-of-the-art solutions to the SCP. Section 4.3 briefly describes ORST and formulates the problem. Section 4.4 reviews the state-of-the-art algorithms selected to solve the SCP in ORST. Section 4.5 reports the results of a simulation assessment. Finally, conclusions are presented in Section 4.6.

4.1 INTRODUCTION

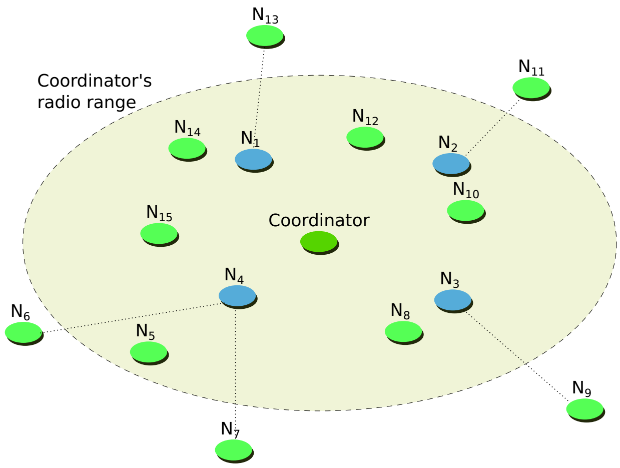
Due to their importance in wireless sensor networks (WSN), relay selection techniques have been extensively studied in the last few years (AZIZ; GHANI, 2019; VALLE, Odilson T et al., 2018; BAO, J. et al., 2017; PHAM; KIM, D. S., 2016). These techniques consider that some neighboring nodes will cooperate with the transmitter-receiver pair, in order to increase the probability of sent messages being correctly received at the destination (VALLE, Odilson T et al., 2018).

An adequate relay selection technique may improve the communication reliability without generating excessive power consumption. Figure 31 shows a WSN organized in a star topology, with a coordinator node in the center and fifteen nodes around it. Nodes N_6, N_7, N_9, N_{11} and N_{13} are outside of the area that has good communication links with the coordinator. This network uses relay nodes (N_1, N_2, N_3 and N_4), which are able to directly communicate with the coordinator node and also to listen nodes that are outside of the coordinator coverage area. Thus, relay nodes may cooperate in communication by retransmitting the listened messages to the coordinator node.

In this example, the smallest number of relays was selected and at the same time it was ensured a good communication link between all nodes and the coordinator. A problem may arise when more relays than required are selected, because this will increase the number of repeated messages and the overall energy consumption (VALLE, Odilson T et al., 2016; VALLE, O. T. et al., 2015). In this way, the performance of cooperative retransmission techniques depends heavily on the efficiency of the relay selection technique (AZIZ; GHANI, 2019; VALLE, Odilson T et al., 2018).

In a recent work Suelen Laurindo et al. (2018), we proposed a relay selection technique to be used in WSNs organized in star topology, named Optimized Relay

Figure 31 – WSN using a cooperative communication.



Source – Author.

Selection Technique (ORST). In this technique, the selection of the relay nodes is based on multiple criteria, namely: the number of neighbors of the candidate node, its remaining battery energy, the quality of the communication link between the candidate node and its neighbor nodes (based on RSSI – Received Signal Strength Indication) and the success rate's history in recent node transmissions. The ORST technique was formulated as an optimization problem, using an objective function and respecting two restrictions. A node x_i will be a candidate to be a relay, if and only if: (a) it is neighbor of the coordinator; and (b) it has at least one more neighbor. The ORST technique seeks to select the smallest number of relay nodes through which all nodes can communicate with the coordinator node. In this way, the ORST technique can be characterized as a resource allocation algorithm that may be reduced to the classic set-covering problem (SCP) applied to WSNs. The SCP seeks to find a minimum number of sets that contain all elements of all data sets (XU, K. et al., 2005).

In a complementary work Suelen Laurindo et al. (2019), we verified that there is no need to make the selection based on multiple criteria, when considered the above mentioned restrictions. In this way, the update version of objective function will be considered in this chapter.

With the rapid development of IoT applications and Industry 4.0 paradigms, where a large number of mobile and static sensors coexist, it is extremely challenging to achieve real-time and reliable communication between the sensors. In this way, considering these complex scenarios, it is crucial to use adequate relay selection algorithms that allow getting better solutions in a short time, without compromising network performance (PILLONI, 2018).

This chapter assesses the performance and quality of different SCP algorithms, which are suitable for solving the optimization problem of the ORST technique. In this context, this chapter presents the following contributions: (i) an analysis of the computational time required to solve the set-covering problem generated in the WSN relay selection, through a simulation comparison of three state-of-the-art resolution algorithms, which are: greedy algorithm, Genetic Algorithm (GA) and Branch and Bound ($B\&B$) algorithm, and (ii) a simulation assessment using the discrete event simulation tool OMNET++ to analyze the operation of the WSN when using each of the above resolution algorithms.

4.2 RELATED WORKS

This section aims to present some of the most relevant state-of-the-art solutions to deal with the set-covering problem.

In Qi Yang et al. (2015), the authors proposed an algorithm based on the greedy algorithm, named List and Remove (LAR). The authors optimized the process to find the best set covering result. They insert a count of the number of uncovered elements by the result set for each set. Then the best set to be added can be found more easily.

In Crawford et al. (2016), an algorithm that emulates cat behavior to solve optimization problems is proposed, which is called Binary Cat Swarm Optimization (BCSO). The BCSO use two behaviors of cats, the seeking mode, where the cats spend most of the time resting, but in fact, they are constantly alert and moving slowly. And the tracing mode, when cats detect a prey and they do fast movements. In BCSO these two behaviors are mathematically modeled to solve complex optimization problems. The authors showed that the proposed metaheuristic performs well in most observed cases. However, it may also occur a premature convergence problem, converging the solution to a local optimum.

The authors of Bilal et al. (2013) proposed a new formulation for the SCP with a maximization objective. The gain attributed to each element must be higher than the cost of at least one of the sets that covers the element. Otherwise, there is no benefit of covering that element. In addition, the authors propose a Descent Heuristic (DH) for this formulation of the SCP. The proposed DH is an adaptation of the classical greedy heuristic. In the DH algorithm, the authors seek to find the set s_j with the maximum ratio $R_j = \delta_j/c_j$, where c_j is the cost of the set s_j and δ_j the variation in the objective function associated with adding (removing) the set s_j. The algorithm stops when the current configuration is better than all of its neighbors ($R_j \leq 0$, for all j). The authors compare the DH version to the classical greedy heuristic version using 88 set covering problems. According to the authors, the version based on the new formulation found better solutions than the original version for 69 test problems, equal solutions for 10 problems, and worse solutions for 09 problems.

In Bara'a and Hameed (2014), the authors described a set covering problem applied to sensor wake-up scheduling, to alternate between active and idle state in a WSN. A genetic algorithm (GA) is proposed to construct a reliable set cover, where the total number of sensors used in the set is to be minimized. The GA simulates the biological processes of natural selection, reproduction, and mutation to iteratively evolve species of individual solutions to become more and more adapted to the problem environment. The authors aim to iteratively evolve a population of solutions, using genetic operators, toward the best set cover solution in terms of minimum number of sensors (i.e., sensor cost) that reliably cover all targets. According to the authors, most of the time, the proposed GA finds a minimum number of active sensors from the whole sensors set.

In Constantino et al. (2010), a new GA is proposed. In a common genetic algorithm, the representation of a chromosome is a binary array with the same size of the number of columns of the matrix, where the position j has value 1 if the column j is in the solution, and 0 otherwise. In the proposed approach, the authors implemented a different representation through a matrix. In addition to the column list (matrix rows), the authors keep an auxiliary structure. Storing for each row, the number of columns that covers this row and the indexes of these columns. Each individual (solution) is represented by a set of the index (columns) implemented as a linked list structure, instead of a binary representation. The authors compared the proposed solution with three other state-of-the-art solutions and the proposed solution proved to has better quality than just one of the compared algorithms.

In Eremeev (2018), the authors proposed an improvement in a GA to solve the set cover problem. A new restart rule is proposed. The rule uses the Schnabel Census method for estimation of the number of solutions which may be visited, whenever the distribution of offspring in the GA remains unchanged. According to the authors, the computational experiments show a significant advantage of the GA with the new restart rule over the GA without restarting.

According to the works presented in the literature, it is possible to observe that there are different approaches for solving a set covering problem. Most of the mentioned papers use genetic or greedy algorithms to solve the set covering problem and show good results. Therefore, the genetic and greedy algorithms will be used and evaluated together with the $B\&B$ to solve the optimization problem of the ORST technique. The branch and bound algorithm was the approach originally used when the ORST technique was proposed in Suelen Laurindo et al. (2018).

4.3 THE ORST TECHNIQUE AND PROBLEM FORMULATION

The ORST technique was designed as an optimization problem using an objective function. The objective function (Equation 14) takes into consideration the available energy in the nodes (e). This parameter was selected among the set of available parameters because it was demonstrated that this is the parameter which has the highest impact upon the quality of the network operation (LAURINDO, S. et al., 2019).

The objective function aims to ensure that appropriate nodes are selected as cooperating nodes. Each node x_i will calculate its objective function value W_i and this information will be sent to coordinator.

$$W_i \doteq \left(\frac{\beta^e}{e_i}\right) \tag{14}$$

where:

- $e_i = \frac{RE_i}{IE_i}$, being RE_i the remaining energy and IE_i is the initial energy of node x_i, respectively. The e_i value is the normalized remaining energy of node x_i (a real number between 0 and 1);
- β^e is a constant used so that the node with the largest amount of energy has the lowest cost in the objective function.

In order to select the minimum number of relay nodes, ensuring at the same time every node has a reachable relay, an optimization problem is formulated as follows:

$$minimize \sum_{i=1}^{n} W_i y_i \tag{15a}$$

$$subject \quad to: \quad \mathtt{A}y \geq b \tag{15b}$$

$$\mathtt{C}y = d \tag{15c}$$

$$y_i \in \{0, 1\}$$

In the constraint presented in Equation 15b, $\mathtt{A}$ is the adjacency matrix of order nxn, where its element $a_{i,j} = 1$ if node x_i is a neighbor of node x_j and $a_{i,j} = 0$ otherwise. Matrix $\mathtt{A}$ is formed in the coordinator node based on the list of neighbors sent by each node of the network. Therefore, whenever the list of neighbors of a node x_j has not been received by the coordinator, all elements of row j of matrix $\mathtt{A}$ will be equal to zero; y is a vector of order nx1, where y_i will be equal to 1 when node x_i is selected as relay and 0 otherwise and; b is a vector whose b_i value has been defined as 1, representing the minimum number of relay nodes of each node x_i. Considering the WSN presented in Figure 31, the coordinator will build matrix $\mathtt{A}$ from the list of neighbors of nodes $N_1, N_2, N_3, N_4, N_5, N_8, N_{10}, N_{12}, N_{14}$ and N_{15}, which are the nodes that the coordinator can hear and consequently receives the message with the list of neighbors. All elements of N_6, N_7, N_9, N_{11} and N_{13} rows of matrix $\mathtt{A}$ will be equal to zero.

The constraint presented in Equation 15c is determined by the coordinator node, where matrix $\mathtt{C}$ represents the set of nodes that do not have an adequate communication link with the coordinator node. Each row of matrix $\mathtt{C}$ represents a node x_i that is not able to directly communicate with the coordinator and each column represents a node that is able to hear this node. In this case, d will be equal to 1, in order to guarantee that at least one of these nodes will cooperate with node x_i.

The proposed ORST scheme aims to find a set of relay nodes among the WSN nodes, ensuring two conditions: (1) each node x_i ($1 \leq i \leq n$) is covered by at least one relay node; (2) the sum of the weights of the relays is minimized. In this scheme, x_i is used as node identifier and n is the total number of nodes in the network. There is one node called a coordinator in the WSN (C). The ORST scheme is a resource allocation algorithm that may be reduced to the classic set-covering problem applied to WSNs (XU, K. et al., 2005). Considering the WSN presented in Figure 31, the rows of matrix $\mathtt{C}$ will be filled with nodes N_6, N_7, N_9, N_{11} and N_{13}, which are the nodes that do not communicate with the coordinator node. The columns will be filled with the nodes that listen to each of these nodes, that is, N_4 for the nodes N_6 and N_7; N_3 for the node N_9; N_2 for the node N_{11}; and N_1 for the node N_{13}. In this way, the coordinator will find the relay nodes solving the optimization problem with the mentioned constraints.

The set-covering problem seeks to find a minimum number of sets that contain all elements of all data sets. According to (CORMEN et al., 2009), the set covering problem can be formally defined as follows. An instance (X, $\mathcal{F}$) of a set covering problem consists of a finite set X and a family $\mathcal{F} = s_1, s_2, ...s_z$ of subsets of X (z is the total number of subsets in $\mathcal{F}$), such that every vertex of X belongs to at least one subset in $\mathcal{F}$:

$$X = \bigcup_{s \in \mathcal{F}} s \tag{16}$$

A subset $s \in \mathcal{F}$ covers its elements. Thus, the problem is to find a minimum-size subset $C \subset \mathcal{F}$ whose members cover all of X:

$$X = \bigcup_{s \in \mathcal{C}} s \tag{17}$$

When a subset C satisfies the Equation 17, it covers X.

The ORST problem considers a WSN composed of a set of nodes $X = \{x_1, x_2, \ldots, x_n\}$, being that every node has an associate positive weight value (W_i) and a specific communication range. We construct a directed and weighted graph $G = (X, E)$ in the following way. Each node x_i corresponds to a vertex $x_i \in X$ and two vertices x_i and x_j have an edge $e_{i,j} \in E$ if x_i is able to hear a message sent by x_j with the value of RSSI ≥ -87 dBm, as defined by Srinivasan and Levis (2006) as the minimum value for adequate communication in WSNs.

Every graph with X and E has subsets $\mathcal{F} = \{s_1, \ldots, s_k\}$, where each subset s_k is known as a set cover of the graph G. Each subset of $\mathcal{F}$ is formed by vertices that accomplish conditions (1) and (2).

The WSN problem treated in this chapter consists of finding the set-cover with minimum sum of weights. The corresponding decision problem generalizes the well-known NP-complete vertex-cover problem and is therefore also NP-hard (CORMEN et al., 2009; KARP, 1972).

4.4 ALGORITHMS USED TO SOLVE THE RELAY SELECTION PROBLEM IN THE ORST TECHNIQUE

As mentioned in the Section 4.3, the ORST scheme can be reduced to the classic set-covering problem and this kind of problem can be solved by different algorithms. In a previously work (LAURINDO, S. et al., 2018), we modeled the ORST problem as a Binary Integer Problem (BIP), which was solved with $B\&B$ algorithm. This solution presented good results. However, other solutions could also be explored, such as the use of greedy algorithms or genetic algorithms. In this section, we describe the algorithms that will be evaluated to solve the relay selection problem. The $B\&B$ algorithm is described, in sequence the Greedy algorithm and finally, the Genetic algorithm.

4.4.1 Branch-and-Bound Algorithm - $B\&B$

Based on the $y_i \in \{0, 1\}$ variables of the ORST problem, cited in the section 4.3, the minimum set cover problem can be formulated as a Binary Integer Problem (BIP) and solved with the $B\&B$ algorithm.

The $B\&B$ algorithm allows to produce exact solutions to NP-hard optimization problems. The goal of this algorithm is to partition the feasible region into subdivisions and then, if necessary, to further partition the subdivisions (WOLSEY, 1998). This algorithm uses a tree search strategy to implicitly enumerate each of the possible solutions of a given problem, applying pruning rules to eliminate branching that cannot lead to a better solution (WOLSEY, 1998).

Considering an optimization problem defined as $P = (X, f)$, where X is the space for feasible solutions and $f : X \to \mathbb{R}$ is the objective function. To solve P, the $B\&B$ algorithm builds a search tree T of subproblems. At each iteration, a new subproblem R $\subseteq X$ is selected from a list L of unexplored subproblems; the first feasible solution is named the incumbent ($\hat{x} \in X$) and is globally stored. If in a subproblem R a new solution is found with a better benefit value, the incumbent solution is updated. However, if it can be proven that the subproblem R cannot provide a better solution than the incumbent, the subproblem is pruned. Otherwise, child subproblems are generated by branching R in other set of subproblems $(R_1, R_2, ..., R_n)$, which are then inserted into T. Once no unexplored subproblems remain, the best incumbent solution is returned (MORRISON et al., 2016). Algorithm 1 presents the pseudocode for a $B\&B$ algorithm.

The complexity of $B\&B$ algorithms is dependent on two factors: the branching factor b of the tree, which is the maximum number of elements (subproblems) generated at any node in the tree, and the search depth d of the tree, which is the length of

Algorithm 1 Branch-and-Bound Algorithm.

```
procedure BRANCH-AND-BOUND(X, f)
    L ← {X}
    x̂ ← 0
    while L is not empty do
        Select a subproblem R from L to explore
        if a solution x̂′ ∈ {x ∈ R | f(x) < f(x̂)} can be found then
            x̂ ← x̂′
            if R cannot be pruned then
                Branching R into R_1, R_2, ..., R_n
                Insert R_1, R_2, ..., R_n into L
            end if
            Remove R from L
        end if
    end while
    return x̂
end procedure
```

the longest path from the root of T to a child element. This way, the $B\&B$ algorithm have a worst-case running time of $O(Mb^d)$, where M is the maximum time to solve a subproblem (MORRISON et al., 2016).

4.4.2 Greedy Algorithm

The greedy algorithm is a simple solution used in the set covering problem. Although greedy algorithms do not result in optimal solutions, they are widely used for obtaining satisfactory results due to their simplicity of implementation (CHANDU, 2015). A greedy algorithm makes a sequence of decisions. Each decision seeks to achieve the optimal for this decision, even though greedy choices does not always lead to the global optimal decision (CHANDU, 2015).

The Algorithm 2 presents a pseudocode for the classical greedy algorithm for set covering problem of the ORST technique. In each step, the greedy algorithm selects the set with the largest number of uncovered elements of X.

Considering a WSN with 10 nodes, i.e., let $X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$ be the universal set and $F = \{\ S_{x1} = \{x_2, x_3, x_4, x_5\}$, $S_{x2} = \{x_1, x_6, x_7\}$, $S_{x3} = \{x_1, x_2, x_6, x_7, x_8\}$, $S_{x4} = \{x_3, x_9, x_{10}\}$, $S_{x5} = \{x_1, x_3, x_7\}$, $S_{x6} = \{x_2, x_3, x_8\}$, $S_{x7} = \{\ \}$, $S_{x8} = \{\ \}$, $S_{x9} = \{x_4\}$, $S_{x10} = \{x_4\}\ \}$ the family of subsets of X. In this case, the subsets of F represent the list of neighbors that each of the nodes listens to, for example the subset $S_{x1} = \{x_2, x_3, x_4, x_5\}$ represents the list of neighbors at node x_1.

In the first step of the algorithm, among the ten sets, S_{x3} has the largest number of elements not yet covered, five elements. In second step, S_{x1} has three uncovered elements $\{x_3, x_4, x_5\}$, S_{x4} has two uncovered elements $\{x_9, x_{10}\}$, S_{x5} has one uncovered element $\{x_3\}$, S_{x6} has one uncovered element $\{x_3\}$, S_{x9} has one uncovered element

$\{x_4\}$, S_{x10} has one uncovered element $\{x_4\}$ and the others sets have no uncovered elements. Thus, the second step selects S_{x1}.

In third step, S_{x4} has two uncovered elements $\{x_9, x_{10}\}$ and the others sets have no uncovered elements. Thus, S_{x4} is selected and the solution set $C = \{ \{x_1, x_2, x_6, x_7, x_8\}, \{x_3, x_4, x_5\},\{x_9, x_{10}\} \}$ contains all the elements of X. In this case, the smallest number of sets was selected and the nodes selected as relays were x_3, x_1 and x_4.

The complexity of the greedy algorithm is presented in Cormen et al. (2009), where the authors presented and proved that the greedy algorithm for set covering problem is a polynomial-time $(ln|X| + 1)$-approximation algorithm.

Algorithm 2 Greedy Algorithm.

```
procedure GreedySetCovering(X, F)
    C ← ∅
    U ← X
    while U is not empty do
        Select S ∈ F that maximizes |S ∩ U|
        U ← U − S
        C ← C ∪ {S}
    end while
    return C
end procedure
```

4.4.3 Genetic Algorithm

Genetic Algorithms may be another method to find a solution to set covering problems (ANANIASHVILI, 2015). Genetic algorithms were inspired by the theory of natural evolution, consider characteristics as survival of the fittest, reproduction and mutation of a population to get the best solution to a problem (SHAHEEN; SLEIT, 2016).

Algorithm 3 presents a pseudocode for the classical genetic algorithm used to solve the relay selection problem in the ORST technique.

This algorithm works as follows: initially, it randomly generates a set of individuals; this set is called the population. Each individual is called chromosome and each chromosome has elements which are called genes. In the problem addressed in this chapter, each gene will be a possible relay node, i.e., each chromosome is composed of a set of nodes, which according to their characteristics may qualify this chromosome as a possible solution.

Each chromosome is qualified according to a cost value. For the set covering problem addressed in this chapter, each gene of a given chromosome will calculate its objective function value (Equation 11) and the chromosome cost value will be the sum of the value resulting from the objective function of each of its genes. The lower the cost of the chromosome, the better the chromosome qualification.

After generating the initial population, the reproduction process begins. In this process, two individuals from the population are randomly selected to reproduce. This step is known as crossover. In crossover step occurs the combination of the genes of two chromosomes to generate a new individual, the child chromosome. The child chromosome may mutate according to a probability ρ. If the mutation occurs, a new random gene will be inserted into the chromosome.

After the reproduction process, the population is evaluated, if the cost of a child chromosome C_{xy} is less than the cost of the worst chromosome C_i in the population, then the child chromosome C_{xy} takes the place chromosome C_i, that is, chromosome C_i is excluded from the population. The reproduction process is repeated until a satisfactory solution is found.

Table 5 – Genetic Algorithm Setting.

Parameters	Values
Initial Population Size	100
Number of Repetitions	300
Probability of Mutation ρ	0.1

Source – Author.

The parameters presented in Table 5 were selected for the genetic algorithm used in this chapter. In the ORST problem, a satisfactory solution was found with 300 repetitions of the reproduction process. This value was used as the maximum number of repetitions in the algorithm (*max_repetitions*). Simulator tests were performed to find the number of repetitions and the initial population size. Increasing the number of repetitions and the initial population size, the obtained solution did not improve. For the initial population size, values 20, 30, 50, 100, 300 and 500 were tested. For the tests of the number of repetitions, values 20, 30, 50, 100, 200, 300, 500 and 1000 were considered.

Algorithm 3 Genetic Algorithm.

```
procedure GeneticAlgorithm
    Initialize population;
    Evaluate population;
    while repetitions < max_repetitions do
        Select two chromossomes;
        Perform crossover and mutation;
        evaluate population;
    end while
end procedure
```

Using the big-O notation, the genetic algorithm for the set covering problem, used in this chapter, has a worst-case running time of $O(N^2)$. The initialize population function also has a computational complexity of $O(N^2)$, that is, for each individual generated, it is checked in the list of individuals if there is an individual equal to the newly generated, thus avoiding repeated individuals. The crossover, mutation and evaluate population functions have an $O(N)$. In this way, the aggregate complexity of the genetic algorithm is $O(N^2)$.

4.5 SIMULATION ASSESSMENT AND DISCUSSIONS OF RESULTS

The algorithms presented in Section 4.4 have been coded in C++ language. The algorithms were tested on the same computer using the Ubuntu 18.04 operating system, with an Intel® Xeon® E3-1240 v2 (3.40 GHz) CPU and 16 GB memory.

The network simulation tool OMNeT++ and the WSN framework Castalia were used to evaluate the operation of each of the algorithms to select a set of adequate relays in a WSN. The open source Solve Library *lp_solve* (SOLVE, 2007) was used to solve the optimization problem with $B\&B$ algorithm.

Each of the algorithms was performed in five scenarios defined with 21, 41, 61, 81 and 101 nodes, being one of them the Personal Area Network (PAN) coordinator. Nodes were randomly deployed in an area of 50x50m^2, with the PAN coordinator positioned in the center (considering the WSN in a star topology). The used channel model was the free space model without time-varying. Other simulation parameters are described in Table 6.

Table 6 – Simulation Setting.

Parameters	Values
Node distribution	Random with coordinator in center
Radio	CC2420
MAC layer	IEEE 802.15.4
Number of superframe slots	145 (5 are used by the CAP)
Data rate	250 kbps
Initial energy per nodo	18720 J
TxOutputPower	0 dBm
BO	6
SO	4
β^e	1.5

Source – Author.

The simulation execution time was set to 450 seconds. The radio model used was CC2420, which is compliant with the IEEE 802.15.4 PHY standard. To reduce the statistical bias, each simulation was performed 60 times with a confidence interval of 95%.

The simulations were performed considering a dynamic topology, where only 50% of nodes were associated to the network at time zero and the remainder were subsequently associated in groups of 5 by 5 nodes. The first group at time instant 50 seconds and then, all the other groups every 30 seconds. Considering the scenario with the highest number of nodes (100 nodes not counting the coordinator), after 320 seconds all nodes were associated. Later, from the time instant 320 seconds of simulation, 20% of the nodes of the network randomly left the coverage of the coordinator node. This leaving operation was performed in groups of 4 nodes, every 10 seconds of simulation. Finally, all nodes again joined the network, in the same order they have left (groups of 4), from the time instant 350 seconds of simulation, respecting an interval of 10 seconds for each group, except for the case of the network with 100 nodes, where only 10% of the outgoing nodes returned.

The ORST technique dynamically performs the relay selection, using as a criterion, the success rate of the network. That is, if the number of missed messages increases a new relay selection is triggered. Thus, the proposed scenario was designed to force changes to the nodes associated with the coordinator and in the neighbors' list, which will consequently affect the network success rate and allow to assess the reliability of the relay selection procedure.

4.5.1 Simulation Results

The simulation assessment was performed considering the execution time of each algorithm to perform the relay selection. The following metrics were used to measure the network quality performance: success rate, the number of cooperations per node, energy consumption and the percentage of duplicate (useless) messages.

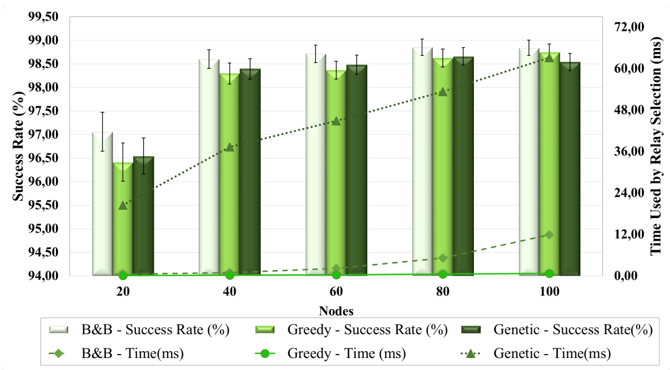
The execution time of each algorithm represents the time that each algorithm takes to perform the relay selection. This metric allows to evaluate if the selected algorithm is a viable solution to be used on real-time WSN. Considering that the solution needs to be obtained by the coordinator node and send to the network nodes in a short time to ensure proper network operation. According to the system model used by the ORST technique (LAURINDO, S. et al., 2018), the coordinator node must use only a 15.36-millisecond time slot to select the relay nodes and to send the beacon.

The success rate represents the ratio between the number of sent messages and the number of messages that actually reach the coordinator. This metric considers messages transmitted in both the transmission attempt and the retransmission attempts performed by relayers. The number of cooperations represents the average number of cooperations performed per node, i.e., it is based on the number of retransmission messages sent by each relay node. Energy consumption represents the average amount of energy spent by each node, obtained through the resource management module available in the Castalia framework. Finally, the percentage of duplicate (useless) messages represents the percentage of cooperation's messages that were not used, i.e., all messages that the relay node listened to and inserted in the cooperation message which had already arrived with success to the coordinator.

The analysis of the results for each of the metrics will allow evaluating the performance of the algorithms in a general context. That is, if the time to obtain the solution ensures the network operation and if the relay nodes are adequately selected.

Figure 32 illustrates the correlation between success rate and the time used by relay selection of the three algorithms. The time used by relay selection is the metric that presents the largest difference between the results obtained by each of the algorithms. The greedy algorithm presented the lowest value in all scenarios, using less than 0.70 milliseconds. The genetic algorithm presented the highest value, using up to 63 milliseconds, in the simulation scenario with 100 nodes. It is possible to observe that the genetic algorithm needs more time to present a similar success rate when compared to the greedy algorithm. Thus, it is evident that the genetic algorithm is an unfeasible solution exceeding the time that the coordinator node has to perform the relay selection. The *B*&*B* algorithm used less than 12 milliseconds to solve the optimization problem of the ORST technique. That is, it was able to select the relay nodes within the time allowed for the coordinator to process and send the beacon (a 15.36ms time slot) and presented a better success rate. In simulation scenarios with 80 and 100 nodes, it reaches 98.8% of the network success rate.

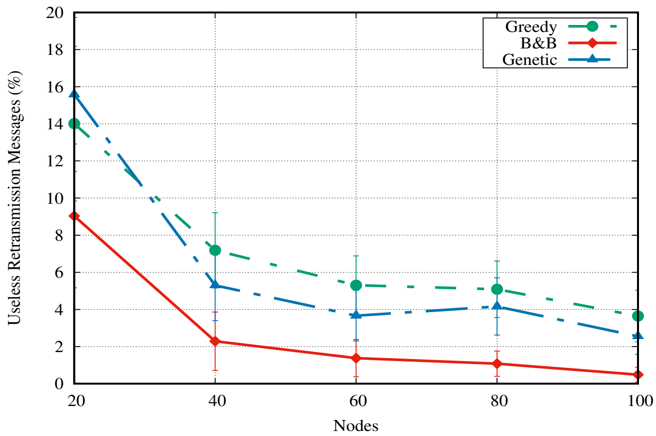
Figure 32 – Correlation between success rate and time used by relay selection.



Source – Author.

Figure 33 presents the useless retransmission messages. It is possible to observe that, when the network presents a lot of useless retransmissions messages, some unnecessary relay nodes have been selected. These nodes retransmitted messages that had already been received at the coordinator node. The *B*&*B* algorithm presented the best results. In the simulation scenario with 100 nodes, the percentage of useless messages to this algorithm was nearly zero (0.49%). Comparing just GA with greedy algorithm, GA presents better results in most scenarios (40, 60, 80 and 100 nodes). But the results as a whole show that both genetic and greedy algorithms have selected more unnecessary relay nodes than the *B&B* algorithm.

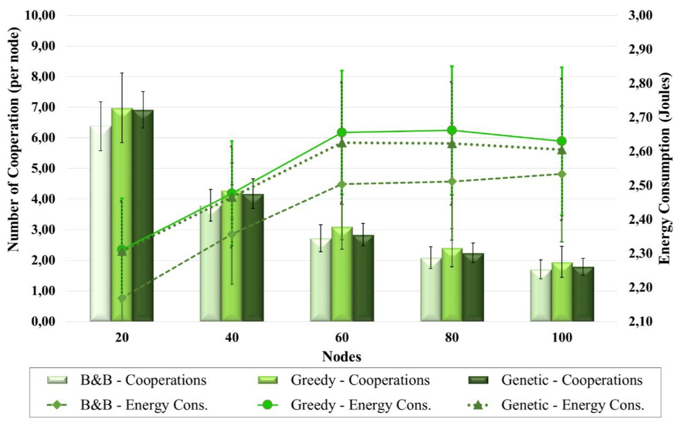
Figure 33 – Useless retransmission messages.



Source – Author.

Finally, Figure 34 illustrates the correlation between the Average number of cooperation exchanges per node and the energy consumption of the overall network, considering the different algorithms analyzed to perform the relay selection. It is possible to observe that both the energy consumption and the average number of cooperation exchanges were very similar for all algorithms. This behavior is a consequence of the smaller number of selected relays, due to the optimization problem used to model all algorithms.

Figure 34 – Correlation between Average number of cooperation exchanges and energy consumption.



Source – Author.

4.6 CONCLUSION

In this chapter, it has been assessed the performance and solution quality of three different algorithms (named genetic algorithm, greedy algorithm and *B*&*B* algorithm) for solving the optimization problem of the ORST relay selection technique.

In the obtained results, the *B*&*B* algorithm presented the best results for the metrics success rate, energy consumption, the average number of cooperation exchanges and the percentage of useless relay messages. The obtained results for the metric time used by relay selection showed that the genetic algorithm is the worst option, which is not suitable to solve the optimization problem of the ORST technique.

On the other hand, although the greedy algorithm had a higher number of useless retransmissions, which shows that some relay nodes were unnecessarily selected, it can be also considered a viable solution for the optimization problem of the ORST technique. The greedy algorithm presented features such as simplicity of implementation, extremely low execution time, and a success rate close to that obtained with the *B*&*B* algorithm. The *B*&*B* algorithm was the approach that presented the better solution quality and, it can be considered as the best solution for the relay selection when using the ORST technique.

As future work, we plan to implement the ORST technique using the *B*&*B* algorithm in a real WSN prototype.

5 RETRANSMISSION MECHANISMS COMBINED WITH THE ORST TECHNIQUE

In addition to proposing a relay selection technique, one aim of this thesis is to develop a retransmission mechanism based on a network coding technique and to evaluate the operation of a network in which both techniques are implemented. This chapter presents four new retransmission techniques, of which three are based on network coding algorithms. A study of the state-of-the-art in network coding techniques is carried out and an evaluation of the WSN operating with both ORST and the network coding technique simultaneously is presented. This study was motivated by recent research (YUE et al., 2016; VALLE, Odilson T et al., 2016; LIU, X. et al., 2014) showing that a combination of cooperative diversity and network coding techniques can improve the quality of communication over WSNs and can offer benefits such as gain in throughput, security, better use of resources, and reductions in the numbers of packets lost due to link failure (FRAGOULI et al., 2006; HO, Tracey; LUN, 2008; FRAGOULI; SOLJANIN, et al., 2007). This chapter is a transcription of a paper called "Combining network coding and retransmission techniques to improve the communication reliability of wireless sensor network", which was published in the journal *Information* (LAURINDO, S. et al., 2021). The following paragraph presents a transcript of the abstract of the cited paper.

This chapter addresses the use of network coding algorithms combined with adequate retransmission techniques to improve the communication reliability of Wireless Sensor Networks (WSN). Basically, we assess the recently proposed Optimized Relay Selection Technique (ORST) operating together with four different retransmission techniques, three of them applying network coding algorithms. The target of this assessment is to analyze the impact upon the communication reliability from each of the proposed retransmission techniques for WSN applications. In addition, this chapter presents an extensive state-of-the-art study in what concerns the use of network coding techniques in the WSN context. The initial assumption of this research work was that the ORST operating together network coding would improve the communication reliability of WNS. However, the simulation assessment highlighted that, when using the ORST technique, retransmission without network coding is the better solution.

As described above, the results showed that ORST worked better with retransmission without network coding, which was an unexpected result. An analysis showed that using a network coding technique together with ORST not only offers no benefits but can also cancel out the improvements provided by ORST in terms of communication. We therefore observed that the selection of a method for the retransmission step has an impact that is as great as the selection of a relay selection technique.

As previously mentioned, this chapter reproduces an authored paper and it is organized as follows. Section 5.1 introduces the context of network coding operations and describes the benefits of using network coding in conjunction with cooperative diversity. Section 5.2 reviews the state-of-the-art in relation to relay selection techniques for WSNs and retransmission mechanisms using network coding techniques. This section also presents a classification framework based on the ways in which network coding is performed. Section 5.3 describes the proposed relay selection technique and the related network coding technique, the aim of which is to improve the communication reliability of WSNs. Section 5.4 presents the results of simulation assessments of the relay selection and retransmission mechanisms using network coding. Finally, our conclusions are presented in Section 5.5.

5.1 INTRODUCTION

The concept of industry 4.0 refers to the evolution of production systems through the integration of industrial automation and information technologies. This evolution trend uses Industrial Internet of Things (IIoT) as a key technology. An IIoT system is integrated by both software and hardware components, where hardware refers to smart sensors and actuators and its network infrastructure, namely Wireless Sensor Network (WSN) (ADRIANO; ROSARIO, 2018).

In WSNs, the communication reliability provided by the link-layer connection depends heavily on the channel conditions. Usually, the channel conditions are different as the WSN deployment environment varies (EZ-ZAZI et al., 2017). Consequently, communications carried out in wireless industrial environments by WSNs are subject to higher reliability constraints when exchanging messages between network nodes and also on energy constraints due to battery depletion (GUNGOR, Vehbi C; HANCKE, Gerhard P, 2009; RODRIGUES et al., 2017a, 2017b). Industrial environments can be affected by electromagnetic noise and/or obstacles between the nodes that may reflect or attenuate the physical communications, preventing messages from reaching their destinations (VALLE, Odilson T et al., 2016). In this context, techniques that improve communication reliability are of major importance. Among them, cooperative communication has received attention from researchers (FRAGOULI et al., 2006; VALLE, Odilson T et al., 2018).

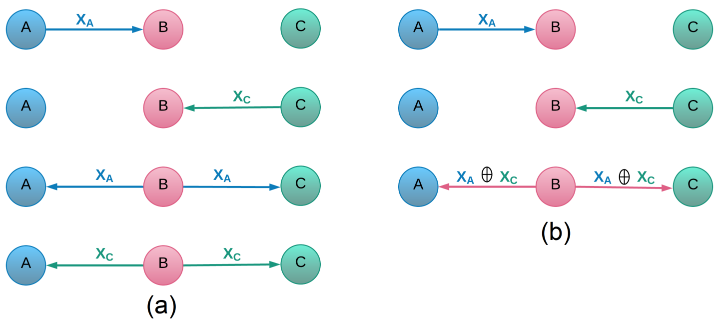
Cooperative communication techniques allow creating path diversity between source and destination nodes using intermediate nodes. Intermediate nodes share each other's antennas and may form virtual antenna arrays. That is, it allows auxiliary nodes to retransmit independent copies of messages heard from neighbor nodes (LU, X. et al., 2018; BLETSAS et al., 2006). Besides achieving diversity gain, cooperation can also increase the coverage area in wireless networks (BHUTE; RAUT, 2013).

The performance of cooperative communication schemes mainly depends on the use of efficient relay selection techniques. The performance of data transmissions can be drastically improved if relay nodes are optimally selected. Similarly, the opposite can occur; that is, if relay nodes are improperly selected, the energy consumption of the network will increase, and a large number of repeated messages will be unnecessarily sent. As a consequence, the relay selection is a decisive step for setting-up cooperative communication schemes (LAURINDO, S. et al., 2018; ASAM et al., 2019).

In addition to the relay selection technique, the retransmission protocol plays an important role in ensuring the reliability of wireless transmission and also is one of the most relevant research contents in cooperative WSNs. This type of protocol ensures that messages sent by the relay nodes correctly reach the destination node (CHEN, Ye et al., 2019; HOSSAIN et al., 2019).

According to relevant state-of-the-art research, network coding (NC) techniques have proven to be a good strategy to carry out retransmissions (HO, Tracey; LUN, 2008; FRAGOULI; SOLJANIN, et al., 2007). Network coding techniques allow relay nodes to process inputs from independent data streams, that is, to send data packets that are linear combinations of previously received packets from different sources (OSTOVARI et al., 2014). This type of technique may improve the transmission rate of the network, considering that the node will send more information in fewer data packets (HO, Tracey; LUN, 2008). An example of how the NC technique works is shown in Figure 35 (FRAGOULI; SOLJANIN, et al., 2007).

Figure 35 – Communication over a wireless network: (**a**)—without NC and (**b**)—with NC.



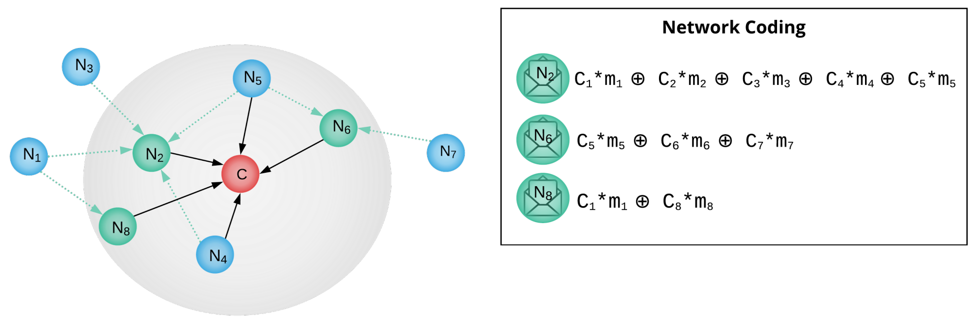
Source – Author.

Figure 35a,b show how communication occurs over wireless networks. Nodes A and C want to exchange packets between them via an intermediate node. Node A wants to send packet X_A to node C via node B. Similarly, node C wants to send a packet X_C to node A via the intermediate node B. Figure 35a presents a conventional communication in a wireless network. Figure 35b presents communication using NC techniques. In this process, only three transmission steps are required. First, node A and node C transmit packets X_A and X_C individually to node B. Then, node B receives both packets, performs an XOR operation with packets X_A and X_C, creating a new encoded packet $X_A \oplus X_C$. Finally, node B retransmits the encoded packet. Then, node A decodes $X_A \oplus (X_A \oplus X_C)$ to get packet X_C, and node C decodes $X_C \oplus (X_A \oplus$

X_C) to obtain packet X_A. In this way, NC reduces the number of packet transmissions. Network coding techniques can be performed in different ways. In the example shown in Figure 35b, binary coding based on XOR was used. In Section 5.2.2, we present the major state-of-the-art techniques to perform NC.

Consider Figure 36 to show how cooperative diversity can work together with a NC technique in the retransmission step. Figure 36 shows a network composed of ten nodes, nine final devices and one coordinator node. Nodes N_2, N_6, and N_8 are relay nodes, and nodes N_1, N_3, and N_7 are outside the coverage area of coordinator node (node C). In this network, we consider that communication takes place in two steps. In the first step, each node transmits a message to the coordinator, but not all messages successfully arrive at the coordinator; messages from nodes N_1, N_3, and N_7, due to interference, do not reach the coordinator node. Then, in a second step, each of the relay nodes applies a NC technique (Figure 36 presents a linear NC, which will be explained in Section 5.2.2) to the messages they have heard. In this case, the relay node N_2 codes messages from nodes N_1, N_3, N_4, N_5, and its own message. Relay node N_6 codes messages from nodes N_5, N_7, and its own message, and relay node N_8 codes messages from nodes N_1 and its own message. Then, each relay node retransmits its coded message to the coordinator.

Figure 36 – Communication over a wireless sensor network considering cooperative diversity and NC.



Source – Author.

Some works in state-of-the-art show benefit from using NC together with co-operative diversity. In Yue et al. (2016) and Odilson T Valle et al. (2016), the authors use relays and NC in the retransmission step, which has improved the reliability of the communications in industrial wireless networks. According to Xingcheng Liu et al. (2014) combining cooperative communication and NC can increase the packet loss-resistant capability due to the packet redundancy. In addition, the network may be able to overcome node failures via cooperative communications.

In previous work, we had studied solutions for relay selection and proposed an Optimized Relay Selection Technique (ORST) (LAURINDO, S. et al., 2018, 2019, 2020).

We also investigated the best parameters to be considered when selecting relay nodes.

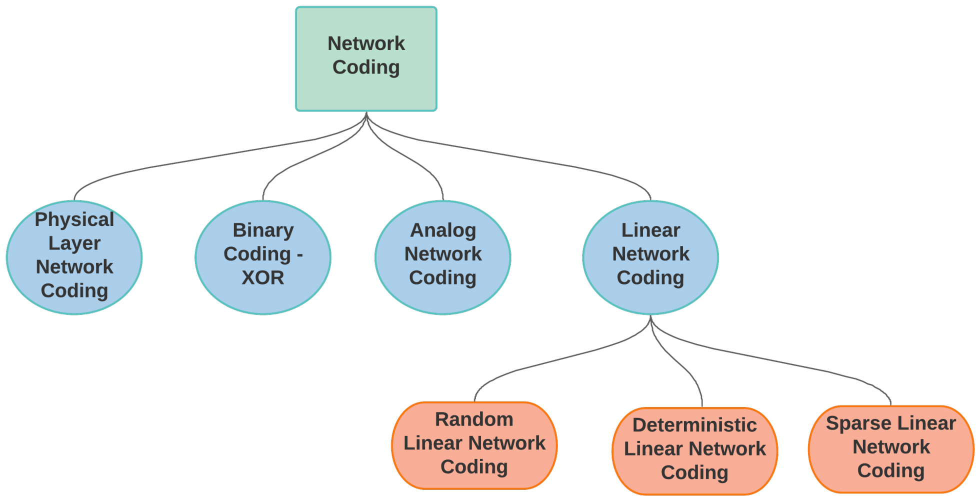
Considering the benefits of using both cooperative communication and NC techniques presented in the literature, in this chapter, we consider a holistic approach to improve the communication reliability, considering the ORST technique combined with the use of an effective retransmission mechanism, aiming to evaluate the operation of the ORST technique together with the NC approaches. Random and Sparse Linear Network Coding will be used as a retransmission mechanism, considering a scheme that allows the relay nodes and the coordinating node to combine a priori which coefficients will be used. In this way, the NC technique would be able to improve the retransmission reliability and reduce the overhead generated when sending the coding coefficients. As the main contributions in this chapter, we can mention:

- An extensive state-of-the-art study concerning relay selection and NC techniques, presenting relevant and current works;
- A simulation assessment of both proposed schemes, relay selection and the NC, working together in the communication. In addition, we will present an analysis of the advantages and drawbacks of the combined implementation of both schemes;
- A discussion about some of the negative results obtained when combining the ORST technique with specific NC approaches.

5.2 RELATED WORK

This section presents some of the most relevant state-of-the-art works related to cooperative communication and NC techniques within WSN communication context. In cooperative communication, we present relay selection techniques, which are decisive to improve communication reliability. In NC, first, we will introduce a classification of the way to perform network coding, which we divided into four categories: Physical Layer Network Coding, Analog Network Coding, Binary Coding-XOR, and Linear Network Coding, which is subdivided into Random Linear Network Coding, Deterministic Linear Network Coding, and Sparse Linear Network Coding, as presented in Figure 37. Besides, we select the most relevant state-of-the-art works with a focus on how to carry out the selection and sending of the coefficients used to encode the messages; consequently, we analyze how to reduce the overhead generated by the transport of the coefficients.

Figure 37 – Network coding categories.



Source – Author.

5.2.1 Relay Selection Related Work

Tripathi et al. (2019) proposed an Energy Balance Load Aware Relay Selection in Cooperative Routing (EBLCR) protocol. If the Packet Reception Ratio (PRR) of routing nodes is less than a respective threshold, a new relay node will be selected for doing the data transmission. The router node will broadcast a control message for relay node selection. After receiving the control message, each node checks its residual energy. If the energy is greater than a predetermined threshold, then it can work as a relay node. The possible relay nodes start their timers after receiving the control packet, and the one with the lowest timer value node will be selected as a relay. The authors compared the EBLCR protocol with just one other state-of-the-art technique, and the results show that the EBLCR improves the throughput and the energy consumption per packet.

K. Yang et al. (2019) proposed a relay selection method based on Q-learning (QL), named QL-RSA, which selects the relays using the maximum cumulative reward to obtain the maximum throughput of the cooperative networks. The authors considered that the interaction between the agent and the environment is a Markov Decision Process (MDP), which consists of a finite and discrete set of environmental states, a set of finite and discrete learner actions, scalar enhanced signals, and a learner's strategy. In each iteration, the source node (learner) perceives the state of the environment and selects actions to act upon the environment, according to the current strategy. Then, a reinforcement signal, called a reward, is generated to feedback to the source node. Based on this, the strategy is updated and the next iteration is initiated. The ultimate goal of learning is to find the best strategy for each state, aiming to maximize the expected long-term cumulative reward, and consequently, reaching the maximum throughput of the destination node after the action is performed. The reward value is obtained through the feedback channel between the source and the destination for updating the Q-matrix and guiding the future policy selection. The authors compare its proposed technique just with a random relay selection algorithm (R-RSA); the results showed that the throughput obtained by QL-RSA is better than that of R-RSA.

Mei and Y. Lu (2019) proposed and analyzed three relay selection schemes, named Random Relay Node Selection (RRNS), Best Relay Node Selection (BRNS), and All Relay Nodes Selection (ARNS). The proposed cooperative communication system considers two phases of transmission. In the first phase, the source node transmits signals to the destination node and to M relay nodes. During the second phase, depending on the given relay node selection scheme, relay nodes that correctly decoded the signals received in the previous phase are selected to relay information. Consider $D(s)$ as a set of relay nodes that can correctly decode the signals transmitted from the source node during phase I. The operation of the relay selection techniques occurs as follows. The RRNS scheme randomly selects a single relay node R_m from $D(s)$ for relaying information during the second transmission phase. In the BRNS, the best relay node B_m, defined as the relay with the highest instantaneous channel gain across the relay-destination link, is selected for relaying information. The ARNS scheme selects each node in $D(s)$ to retransmit the decoded signals to the destination node. The results presented by the authors just analyzed the outage probability. The results achieved showed that the ARNS and the BRNS performed better than RRNS. However, other metrics also need to be considered for a better assessment.

Jian Zhang et al. (2020) presented a cooperative relay selection technique for a cluster tree network. The objective is to reduce energy consumption. The authors consider that the node spends more energy to make long-distance transmissions in a single hop than if there are nodes that can cooperate with it. As a selection parameter, the authors consider the residual energy of each node and the node density. As a node density, they consider the number of neighbors of each node divided by the number of nodes within the cluster. As a result, the authors compared it to a network that considers only one hop, and their proposed technique showed lower energy consumption.

Yuhan Su et al. (2019) proposed a (Deep-Q-Net) DQN-based relay selection scheme in WSNs, named DQ-RSS. The scheme combines deep learning with Q-learning to accelerate learning for selecting the optimal relay among the relay candidates according to outage probability and channel information. A source node collects the CSI from the environment and then sends the integral system state to the DQN to evaluate the optimal policy for relay selection. Simulation results show that their relay selection scheme exceeds the Q-learning based relay selection and the random relay selection scheme in terms of lower outage probability and lower energy consumption. However, the proposed technique only works for static networks.

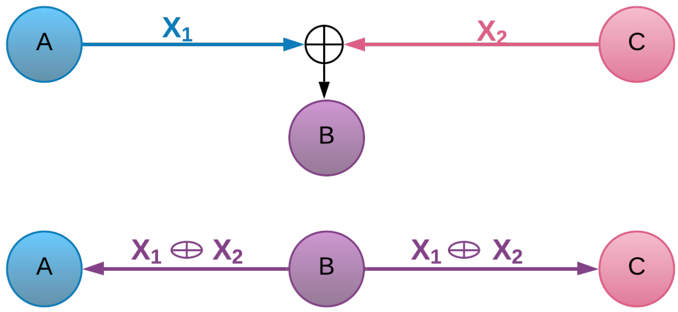
Elsamadouny et al. (2019) proposed a relay selection technique for multihop communication that allows for L relays to be selected between source and destination nodes. The authors modeled the network as a Markov chain, where each Markov chain state is parameterized by $(L-1)$ adjacent number representing the number of packets in the queue of the $(L-1)$ intermediate relay nodes. The transmission of a single packet over a specific hop will cause the system to move from one state to another. The first and the last state index represent the possibility of packet transmission from the source node to the first relay and from the last relay to the destination node, respectively. All the intermediate states represent the possibility of data transmission from an intermediate relay node to the subsequent relay node. The technique works as follows: During each time slot, the highest quality hop (best SNR) is activated for transmission as long as the corresponding relay node has packets to transmit and the corresponding receive node buffer is not full. Otherwise, the second-best hop is activated, and so on. If the selected hop has SNR below a certain threshold SNR, this event will be considered as an outage event. This threshold SNR is predetermined according to the required quality of service. The results presented by the authors showed that the outage probability of their scheme outperforms the conventional multihop scheme.

5.2.2 Brief Explanation of the Main Types of Network Coding

In this section, we present a framework for the classification of how to perform network coding, dividing it into six categories: Physical Layer Network Coding, Analog Network Coding, Binary Coding—XOR, Random Linear Network Coding, Sparse Network Coding, and Deterministic Linear Network Coding (Figure 37).

Physical Layer Network Coding exploits the overlap of electromagnetic waves that occur in wireless communication and applies the concept of NC to the physical layer. In this way, nodes A and C, as shown in Figure 38, transmit their messages simultaneously to the intermediate node, node B, which receives the overlapping signals. Then, the intermediate node extracts a linear combination from the received signal, without the need to individually obtain the messages, and proceeds similarly to the network coding technique (HUANG et al., 2017).

Figure 38 – Network coding on the physical layer.



Source – Author.

The Analog Network Coding uses the interference generated by simultaneous transmissions as an ally. The idea is the following: When two nodes *A* and *B* simultaneously transmit, the packets will collide. The signal resulting from a collision is the superposition of the different signals. Thus, node *A*, after receiving the summed signal, calculates the phase shift of node *B* by using its signal in the sum, thereby recovering the node *B* signal destinated to it; node *B* can recover the signal that it similarly expects from node *A* (HUANG et al., 2017; KATTI et al., 2007).

In the Binary Coding—XOR, a simplification assumed is to use just a basic bitwise XOR (exclusive OR) operation among messages. This basic NC technique uses a finite field $\mathbb{F}_{2^1}$, which represents a field in network coding theory with 2^1 symbol combinations, being able to encode up to 2 messages into a single message. In this way, when using binary coding, XOR operations are performed considering just two packets listened to by the intermediate node (OSTOVARI et al., 2014).

In the Linear Network Coding, there are three subcategories: Random Linear Network Coding (RLNC), Sparse Linear Network Coding (SLNC), and Deterministic Linear Network Coding (DLNC), which are described as follows.

Random Linear Network Coding (RLNC) allows for the use of a finite field higher than binary coding. In this way, it can be used in conventional network coding schemes with multiple source nodes (BASARAN; KURT, 2019). RLNC performs a random selection of the encoding coefficients from a q-element finite field denoted by $\mathbb{F}_q$ (DONG et al., 2019). The larger the finite field, the less likely it is to generate linearly dependent packages at the destinations. If all nodes systematically used the same coefficients, destinations would not decode the received packets, given the high probability of redundant packets, which would generate linearly dependent systems (HO, T. et al., 2006).

This technique consists of linearly combining several messages using a randomly selected coefficient within a finite field $\mathbb{F}_{2^n}$, where n can be any positive integer (MIGABO et al., 2017). Assuming a network with a set of n_d nodes, when an intermediate node i wants to transmit k messages $(m_1, m_2, m_3, ..., m_k)$ listened from its neighbors, it first randomly selects k coefficients $(c_1^i, c_2^i, c_3^i, ..., c_k^i)$ of the finite field. Then, linearly combine the packets that it has to listen to using the Equation (18):

$$m^{cod} = c_1^i m_1 \oplus ... \oplus c_k^i m_k \tag{18}$$

Together with the linear combination, the node sends a list of the used coefficients. At destination node t, the received packets are represented by Equation (19):

$$\begin{bmatrix} m^1 \\ m^2 \\ ... \\ m^k \end{bmatrix} = \begin{bmatrix} c_1^1 & c_2^1 & ... & c_{n_d}^1 \\ c_1^2 & c_2^2 & ... & c_{n_d}^2 \\ ... & ... & ... & ... \\ c_1^k & c_2^k & ... & c_{n_d}^k \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ ... \\ m_{n_d} \end{bmatrix}$$

$$M^{(t)} = GM \tag{19}$$

where $M^{(t)}$ is a matrix whose rows are the k coded messages received at destination node t, M is a matrix, in which the rows represent the original k messages and G is a matrix in which each row represents the vector of coefficients used by the intermediate node to encode the messages. Thus, the destination node will recover the original messages by building and solving a linear system using Equation (19) (VALLE, Odilson T et al., 2016). This kind of NC requires that the node performs the coding operation to send all the coefficients used to carry out the linear combination, together with the coded message. This behavior presents, as its main drawbacks, the complexity of the decoding operation and the overhead resulting from the encoding vector.

Sparse Linear Network Coding (SLNC) is a NC technique presented as an improvement for RLNC. In the SLNC, the intermediate node does not encode all the listened to messages. It encodes only a small number of messages in each transmission. Thus, the decoding complexity is reduced on the receiver. Besides, the communication overhead generated by sending the coefficients is also reduced, considering that the number of coefficients is proportional to the number of coded messages (FEIZI et al., 2014; SEHAT; PAHLEVANI, 2019; ZAREI et al., 2018). Within this context, there are a number of approaches that exploit Low-Density-Parity-Check (LDPC) codes (BAO, X.; LI, J., 2005, 2008; CHEBLI et al., 2009). In these approaches, each relay packet includes the coefficients in a small bit-map field to reduce the overhead.

In Deterministic Linear Network Coding (DLNC), the coefficients used by the intermediate nodes to perform the NC are deterministically selected. That is, the coefficients are not randomly selected in the finite field but selected from techniques that aim to optimize the network coding process (HAN, C. et al., 2017), which means that the validity of the coding scheme is guaranteed. That is, it ensures that encoded messages are linearly independent. The disadvantage of this type of coding is that there is a control overhead to be constructed and maintain a linear coding scheme among nodes (HAO; JIN, 2009).

5.2.3 Network Coding Related Work

In Migabo et al. (2017), the authors proposed a Cooperative and Adaptive Network Coding technique for Gradient-Based Routing (GBR). The technique considers that the network density is dynamic, according to the average number of neighbor nodes, to encode interest messages. The encoding is performed utilizing linear combinations of random coefficients of a finite Galois Field (GF) of variable size $GF(2^s)$. The decoding is performed using Gaussian elimination.

When a relay node wants to transmit n accumulated data packets $(P_1, P_2, ..., P_n)$, it first randomly selects n random coefficients C_1, C_2, ..., C_n from the Galois Field of order 2^s with s being a positive integer. It then linearly combines the accumulated data packets with the randomly generated coefficients. The decoding process is performed by Gaussian elimination process in which the accumulated header data (coefficients) are grouped to form a $n \times n$ matrix $C_{n \times n}$, which is then reduced to a row-echelon form. The n encoded data packets from the transmitter node can then be decoded by solving a set of linear equations provided that the obtained equations are linearly independent from each other.

In Heide et al. (2011), a technique called Generation-based RLNC is used. This technique consists of dividing large amounts of data into smaller blocks, named generations. So, both the encoding and decoding operations are applied by generation and not on the entire data. The authors proposed a random linear network coding, in which the coefficient vector can be sent in two ways: First, the authors considered that the ratio of nonzero scalars in a coding vector is referred to as the density. If the density is low, the coding vector will be sparse and will mostly consist of 0s. Thus, the authors represent each nonzero scalar by an index-scalar pair. In this way, the coding vector is formed by index-scalar pair and, that is necessary to send together with the coding vector the number of index-scalars pairs, reducing the information to be sent. Secondly, the authors cite that the coding vector can also be represented by a bit array, that indicates which scalars are nonzero, and the values of these scalars.

Each scalar can be represented by $\log_2(q)$ bits, and as the maximal number of nonzero scalars is g, where q represents the size of the finite field and g represents the size of the generation. Besides, each index takes $\log_2(g)$ bits. In this way, the overhead generated will depend on the size of the generation and the number of nonzero scalars.

Akhtari et al. (2020) used a random linear network coding. The coefficients are selected randomly and sent in the coded packet header. The authors considered a finite field of 2^8. In addition, the authors consider that in each hop, between the source node and the destination node, the packets are recoded.

The authors check the newly arrived coded packets dependency in the destination node. For this, the destination node runs a specific algorithm. In this algorithm, M is a triangular matrix of k rows with some missing rows. For the newly received vector u packet's code, nonempty rows of M are multiplied to the corresponding coefficient and added to it. If the vector is independent of the elements, the result will not be zero. At this point, the independent vector will be added to the matrix M in the empty slot. Therefore, it is necessary to check the packet's code vector's independence to ensure that the packet is innovative.

Huangnan Wu et al. (2019) propose an algorithm to optimize the finite field size and to improve the efficiency of RLNC. They analyze the relationship between the finite field size and the completion time for the finite-buffer relay transmission scenarios. Based on the analysis, the field size is optimized via numerical search to maximize the effective data rate.

In Dong et al. (2019), the authors used RLNC and defined a new method to minimize the overhead generated by the transport of coefficients. They generated the encoding matrix using a pseudorandom generator $O(N, K)$, where the generator function uses the number of symbols (N) participating in the network coding process and state of the generator (K) as seeds. The coefficient matrix $[\alpha]$ is generated in both transmitting and receiving nodes.

The source node can send the seed to generate a random coefficient matrix at receiving nodes in two ways. The first is using the first encoded packet and the second is using a different secured channel. As soon as the encoded packet is received at destination nodes, seed encapsulated can be used to generate decoding coefficients matrix.

Ye Li et al. (2018) defined a sparse coding scheme where packets are encoded from sequentially formed random subsets of source packets called batches. The relay recodes only from the buffered packets belonging to the same batch to maintain the code sparsity. A sparse coding scheme is used to minimize the coding coefficient delivery cost. Sparse means that the number of source packets involved in generating each coded packet is much smaller than the total number of source packets. Therefore, the coding vector is sparse. Each packet only needs to carry a small number of nonzero coding coefficients (which are uniformly randomly selected from $\mathbb{F}_q$) in the header.

The authors considered that the relay has a finite buffer of size $m << M$, where M is the number of source nodes. In addition, they consider that the number of nonzero elements in each encoding vector is limited to $d << M$; d distinct source packet indexes are uniformly randomly drawn from $\{0, 1, ..., M-1\}$ with replacement to form a batch with a Sequence Number (SEQ). The corresponding d source packets are referred to be the content of the batch, and d is referred to be the batch degree. For each of the b transmissions on Source-Relay, a coded packet, which is the random linear combination of the d source packets is transmitted, where b is the Batch Transmission

Size (BTS). After b transmissions, a new batch with SEQ increased by 1 will be started. The process continues until the destination successfully decodes all the data. The SEQ and the encoding vector are delivered in the header of each coded packet.

Considering a Transmission Control Protocol (TCP) communication and based on MPTCP (Multipath TCP) standard, Changqiao Xu et al. (2017) proposed the pipeline network coding-technique (MPTCP-PNC). The authors aim to reduce encoding and decoding delays and save bandwidth by using new coding coefficient rules. The operation occurs as follows: The sender divides the original packets $P_1 \sim P_m$ with continuous Data Sequence Numbers (DSN) into N groups. Original packets within the same group are combined to form coded packets, that is, original packets included in each coded packet of a group follows the one-to-all progressive approach. For instance, considering three groups, namely $G1$, $G2$ and $G3$. In $G1$, the first original packet P_1 is encoded to C_1 with coefficient 1, and the second coded packet C_2 is a linear combination of original packets P_2 and P_1, using a random coefficient from the finite field for P_1 and coefficient 1 for P_2. The third coded packet C_3 in $G1$ contains P_1, P_2 and P_3. The m-th coded packet C_m of a group can be obtained using a linear combination of original packets $P_1 \sim P_m$.

When establishing a connection, the sender and receiver negotiate and agree to maintain three structures: a Coded Packet Coefficients Matrix (CPCM), a Redundant Packet Coefficients Matrix (RPCM), and a Mapping Rule. The mapping rule presents the following structure: MR: (S_1, S_N, DSN, $flag$) $\rightarrow$ Coding Vector (CV), where S_1 is the smallest DSN of original packets within the group, S_N is the largest DSN of original packets within the group, DSN is the data sequence numbers of the coded packet and $flag$ is an identifier that determines the use of either CPCM or RPCM.

CPCM and RPCM are coefficients matrices for coded packet and redundant coded packets, respectively. MR is a mapping rule from tuple information of a packet to its corresponding coding vector. Elements in the two matrices are generated from the finite field GF(2^8) and linear independence checks among vectors have been performed beforehand. The CPCM and RPCM are generated at the beginning of establishing a connection and will be used throughout the life of this connection. After negotiation in the connection establishment stage, the sender and receiver maintain the same CPCM, RPCM, and MR. At the sender side, the Pipeline Network Coder can use this information to select a coding vector and perform the encoding operation. At the receiver side, when determining coding coefficients for a coded packet, the Pipeline Network Decoder directly selects the coding vector from CPCM or RPCM, which is enabled by the Mapping Rule from (S_1, S_N, DSN, $flag$) to the corresponding coding vector. The tuple (S_1, S_N, DSN, $flag$) together with coded data is then able to reconstruct the coded packet.

Bin Guo et al. (2014) proposed a Decode-and-Forward Network Coding (DFNC) scheme. In this scheme, the authors do not send the coefficients used in the network coding; they just send a m-bit bit-map, signaling which packets were involved in the encoding. Thus, if the relay node receives the packet from source node s_1, the corresponding position in bit-map is set to 1; otherwise, it will be assigned to 0. The coding coefficients are generated using a pseudorandom algorithm considering three criteria: (1) the coefficient is assigned to be 0 if the relay node does not receive the corresponding packet successfully. (2) the coefficient is assigned to be 1 for its packet. (3) other coefficients are selected according to the following mapping function: $h{:}(s, r) \rightarrow GF(2^q)\backslash\{0, 1\}$, where s is the origin of the coded packet (the transmitting user ID), r indicates the sequence number of 1 in the bit-map. The destination will be equipped with the same mapping function to solve the coding coefficients according to the received bit-maps in each packet.

The authors make two considerations for recovering the original packets. First, they consider that all sent packets (by direct transmission and by retransmission) arrive at their destination successfully. Thus, the coefficient matrix is full rank and the data packets may be recovered by solving the set of equations with Gaussian Elimination Algorithm. Then, they considered that not all packets are received correctly. In this case, it will not be possible to decode all packets by solving the set of linear equations. Thus, the authors applied the decoding on the physical layer (at symbol level) to attempt to recover 'failed' received packets that may include correctly received symbols.

Xingkai Bao and Jing Li (2005, 2008) proposed a sparse linear network coding framework, named Adaptive Network Coded Cooperation (ANCC). Basically, ANCC defines two communication phases and works as follows: The relay nodes listen to and store the correctly received neighbor's messages in the broadcasting phase and, in the retransmission phase, each relay node randomly selects a predefined number of listened messages, performs an encode process (binary checksum) and retransmit to the destination. A bit-map field is included in each coded packet retransmitted by relays to inform the destination how the parity checks have been formed and can correspondingly replicate the code graph and perform the decoding.

In Chen Han et al. (2017), the authors proposed a NC technique named Weighted Vandermonde Echelon Fast Coding (WVEFC). To perform the network coding, the authors use a coding matrix F_{PC}. The F_{PC} is an $n_1 \times n_2$–order matrix, where n_1 is the number of source packets that need to be encoded and $n_2 = n_1 + k$, the value of k refers to the number of packets that requires redundant coding to improve the delivery rate of packets. In the F_{PC} matrix, the first n_1 column vectors refer an upper triangular matrix, while the rest of the sub-matrix is a Vandermonde expanding matrix.

For the network coding to work correctly, before each coding operation, the authors need to specify the row and column numbers of the generated coding matrix and specify the sequence in which packets need to be coded. The authors do not specify how the destination node obtains the coefficients used in the NC. In addition, the authors cite that the probability of linearly dependent columns appearing in the WVEFC coding matrix is lower but it still exists.

Odilson T Valle et al. (2016) proposed a communication scheme to WSN, named NetCoDer. They proposed a simple relay selection technique and a random linear network coding. The network was delimited in maximum size of 256 nodes and the finite field size used in the network coding is $\mathbb{F}_{2^8}$. The random linear network coding is performed as follows: Each node has an identification in hexadecimal format, representing its position in the slot scale, with addresses ranging from 00 to FF. The selection of the coefficient used to encode the messages, in the relay node, is based on the cited address of relay node i and the address of the neighbor t, using the following forming rule: $c_t^i = i + t \ mod \ 256$, where i is the identification of the i-th node and t is the identification of the t-th neighbor.

To inform the coordinator which messages each relay node was able to capture and encode, each node needs to forward the addresses of its neighbor nodes. Each relay node i sends a sequence of bits, which represents the presence (1) or absence (0) of the message from a node t. The coordinator is aware of this forming rule and is able to reconstruct the coefficients used in the coded messages.

5.2.4 Wrap-Up

Among the works cited in the state of the art, it can be observed that the focus is on just assessing the relay selection behavior. These works do not address all steps of communication. That is, they do not mention the scheme or protocol used to carry out the retransmission of messages heard by the relay nodes. Thus, the question remains whether retransmission mechanisms based on network coding can maximize the reception of messages at the destination when relay nodes are used.

Table 7 summarizes the network coding described works, comparing them among themselves concerning the following set of classifiers: the type of network coding used; if the coefficients used to code are sent together with the coded message; if a different strategy has been created for sending the coefficients; and whether there is a need to exchange additional messages to send the coefficients.

It is possible to observe that a number of works use random linear network coding (RLNC), which can be due to its following advantages: First, the linear system-generated has a high probability of being solvable, if all the coefficients of all the encoding vectors were randomly selected, independently, and uniformly from the finite field $\mathbb{F}_q$, considering that the finite field size is sufficiently large relative to the size of the network (HO, T. et al., 2006). Second, in this type of NC, there is no control overhead to construct and maintain a linear coding scheme among nodes (HAO; JIN, 2009), allowing the use of this type of network coding in commercial devices.

Table 7 – Network coding techniques.

State-of-the-Art	Classification	Send the Coefficients	Coefficient Submission Strategy	Exchange of Additional Messages
Migabo et al. (2017)	RLNC	√		
Heide et al. (2011)	RLNC	√		
Akhtari et al. (2020)	RLNC	√		
Huangnan Wu et al. (2019)	RLNC	√		
Dong et al. (2019)	RLNC		√	√
Ye Li et al. (2018)	SNC	√		
Changqiao Xu et al. (2017)	RLNC		√	√
Bin Guo et al. (2014)	RLNC		√	
Xingkai Bao and Jing Li (2005, 2008)	SLNC		√	
Chen Han et al. (2017)	DLNC			√
Odilson T Valle et al. (2016)	RLNC		√	

Source – Author.

In addition, it is also possible to observe among the works that do not send the coefficients together with the coded message, that just three of those works do not generate additional messages on the network. Considering that the goal is to reduce the overhead of sending the coefficients, sending extra messages to configure the coefficients is just another form of network overhead.

In this way, the methodology proposed in this chapter will consider both the communication mechanisms used in the transmission and the retransmission steps. In the transmission, we consider that the relay nodes are optimally selected and will be able to listen to the packets that the coordinator did not successfully receive, using the proposed relay selection technique, as described in Section 5.3.2. The retransmission step will consider advantageous characteristics of the following state-of-the-art works Xingkai Bao and Jing Li (2005, 2008), Bin Guo et al. (2014), and Odilson T Valle et al. (2016), which proposed new strategies for sending the coefficients without generating additional messages. Besides, we will adapt the proposed mechanism to be applied upon a sparse version of the communication network.

5.3 WSN COMMUNICATION

In this chapter, a novel communication scheme is proposed intended to improve the communication reliability in wireless sensor networks. The proposed scheme uses Optimized Relay Selection Technique (ORST) (LAURINDO, S. et al., 2018) to select the best set of relay nodes and combines it with two novel network coding approaches based on random linear network coding and sparse linear network coding to perform message retransmissions.

5.3.1 System Model

Consider a cooperative WSN communication system with n source nodes (S), one destination node (D), and m relay nodes (R). It is assumed that each node is fitted with a single antenna and signals on $S-R-D$ and $S-D$ paths use orthogonal channels through time division multiple access (TDMA). Considering the advantages of star topologies such as synchronization, latency, and energy efficiency, it is usually considered the star topology as a suitable topology for industrial usage and is used in this chapter (YANG, K. et al., 2019; CHEN, F. et al., 2008).

The IEEE 802.15.4e amendment using LLDN (Low Latency Deterministic Network) MAC operation mode is adopted for the PHY (Physical) and MAC (Medium Access Control) layers of the network. The IEEE 802.15.4e amendment has been proposed to adequately address the critical requirements of industrial IoT applications such as low latency, high reliability, and robustness of the industrial environment (SAHOO et al., 2017). In this chapter, IEEE 802.15.4e is configured to send the Group Acknowledgments (GACK) for the data received on the same superframe that they were sent.

WSN communication occurs in two steps. In the first, called transmission, it is assumed that, in each beacon interval, each node in the network has one message to transmit and performs the transmission in its timeslot. At the end of the first step, the coordinator node sends a GACK message indicating which messages it has not received. The GACK message is a bit-map, in which if the coordinator received the message from node i, the position i of this vector will be 1 and zero otherwise. The second step is the retransmission, where the selected relay nodes will apply NC and transmit the heard messages in its retransmission slot, which is previously allocated by the coordinator node.

To perform the retransmission, the proposed relay selection technique selects the optimal set of relay nodes that can improve the diversity order and, thus, it can achieve higher throughput (YANG, K. et al., 2019), as described in Section 5.3.2.

5.3.2 The Optimized Relay Selection Technique

The Optimized Relay Selection Technique (ORST) (LAURINDO, S. et al., 2018) was designed as an optimization problem using an objective function. The objective function (Equation (20)) takes into consideration the available energy in the nodes (e). This parameter was selected among the set of available parameters because it was later demonstrated (LAURINDO, S. et al., 2019) that this was the parameter with the higher impact upon the quality of the network operation.

The objective function aims to ensure that appropriate nodes are selected as cooperating nodes. Each node x_i will calculate its objective function value W_i and this information will be sent to coordinator.

$$W_i \doteq \left(\frac{1}{e_i}\right) \tag{20}$$

where:

- $e_i = \frac{RE_i}{IE_i}$, being RE_i the remaining energy and IE_i is the initial energy of node x_i, respectively. The e_i value is the normalized remaining energy of node x_i (a real number between 0 and 1);
- the expression $\left(\frac{1}{e_i}\right)$ is used so that the node with the largest amount of energy has the lowest cost in the objective function.

In order to select the minimum number of relay nodes, ensuring at the same time every node has a reachable relay, an optimization problem is formulated as follows:

$$minimize \sum_{i=1}^{n} W_i y_i \tag{21a}$$

$$subject \quad to: \quad \mathtt{A} y \geq b \tag{21b}$$

$$\mathtt{C} y = d \tag{21c}$$

$$y_i \in \{0, 1\}$$

In the constraint presented in Equation (21b), $\mathtt{A}$ is the adjacency matrix of order nxn, where its element $a_{i,j} = 1$ if node x_i is a neighbor of node x_j and $a_{i,j} = 0$ otherwise. Matrix $\mathtt{A}$ is formed in the coordinator node based on the list of neighbors sent by each node of the network. Therefore, whenever the list of neighbors of a node x_j has not been received by the coordinator, all elements of row j of matrix $\mathtt{A}$ will be equal to zero; y is a vector of order nx1, where y_i will be equal to 1 when node x_i is selected as relay and 0 otherwise and; b is a vector whose b_i value has been defined as 1, representing the minimum number of relay nodes of each node x_i. Considering the WSN presented in Figure 36, the coordinator will build matrix $\mathtt{A}$ from the list of neighbors of nodes N_2, N_4, N_5, N_6 and N_8, which are the nodes from which the coordinator receives the message with the list of neighbors. All elements of N_1, N_3 and N_7 rows of matrix $\mathtt{A}$ will be equal to zero.

The constraint presented in Equation (21c) is determined by the coordinator node, where matrix $\mathtt{C}$ represents the set of nodes that do not have an adequate communication link with the coordinator node. Each row of matrix $\mathtt{C}$ represents a node x_i that is not able to directly communicate with the coordinator and each column represents a node that is able to hear this node. In this case, d will be equal to 1, in order to guarantee that at least one of these nodes will cooperate with node x_i.

The proposed ORST scheme aims to find a set of relays among the WSN nodes, ensuring two conditions: (1) each node x_i ($1 \leq i \leq n$) is covered by at least one relay node; (2) the sum of the weights of the relays is minimized. In this scheme, x_i is used as node identifier and n is the total number of nodes in the network. There is one node called a coordinator in the WSN (C). The ORST scheme is a resource allocation algorithm that may be reduced to the classic set-covering problem applied to WSNs (XU,

K. et al., 2005). Considering the WSN presented in Figure 36, the rows of matrix C will be filled with nodes N_1, N_3 and N_7 , which are nodes that do not communicate with the coordinator node. The columns will be filled with the nodes that listen to each of these nodes, that is, N_2 for node N_1 and N_3; N_8 for node N_1; and N_6 for node N_7. In this way, the coordinator will find the relay nodes solving the optimization problem with the mentioned constraints.

The set-covering problem seeks to find a minimum number of sets that contain all elements of all data sets. According to (CORMEN et al., 2009), the set covering problem can be formally defined as follows. An instance (X, $\mathcal{F}$) of a set covering problem consists of a finite set X and a family $\mathcal{F} = s_1, s_2, ...s_z$ of subsets of X (z is the total number of subsets in $\mathcal{F}$), such that every vertex of X belongs to at least one subset in $\mathcal{F}$:

$$X = \bigcup_{s \in \mathcal{F}} s \tag{22}$$

A subset $s \in \mathcal{F}$ covers its elements. Thus, the problem is to find a minimum-size subset $C \subset \mathcal{F}$ whose members cover all of X:

$$X = \bigcup_{s \in \mathcal{C}} s \tag{23}$$

when a subset C satisfies the Equation (23), it covers X.

The ORST problem considers a WSN composed of a set of nodes $X = \{x_1, x_2, \ldots, x_n\}$, being that every node has an associate positive weight value (W_i) and a specific communication range. We construct a directed and weighted graph $G = (X, E)$ in the following way. Each node x_i corresponds to a vertex $x_i \in X$ and two vertices x_i and x_j have an edge $e_{i,j} \in E$ if x_i is able to hear a message sent by x_j with the value of RSSI ≥ -87 dBm, as defined by Srinivasan and Levis (2006) as the minimum value for adequate communication in WSNs.

Every graph with X and E has subsets $\mathcal{F} = \{s_1, \ldots, s_k\}$, where each subset s_k is known as a set cover of the graph G. Each subset of $\mathcal{F}$ is formed by vertices that accomplish conditions (1) and (2).

The WSN problem treated in this chapter consists of finding the set-cover with minimum sum of weights. The corresponding decision problem generalizes the well-known NP-complete vertex-cover problem and is therefore also NP-hard (CORMEN et al., 2009; KARP, 1972).

Based on the $y_i \in \{0, 1\}$ variables of the ORST problem, cited in the Section 5.3.2, the minimum set cover problem was formulated as a Binary Integer Problem (BIP). In Suelen Laurindo et al. (2020) different solutions were investigated to solve the ORST problem, being the *B&B* algorithm defined as the best solution.

The *B&B* algorithm uses a tree search strategy to implicitly enumerate each of the possible solutions of a given problem (WOLSEY, 1998). The computational complexity of *B*&*B* algorithms is dependent on two factors: the branching factor b of the tree, which is the maximum number of elements (subproblems) generated at any node in the tree, and the search depth d of the tree, which is the length of the longest path from the root of T to a child element. Thus, the *B*&*B* algorithm has a worst-case running time of $O(Mb^d)$, where M is the maximum time to solve a subproblem (MORRISON et al., 2016). For further details, the reader is referred to Suelen Laurindo et al. (2020).

5.3.3 Network Coding Technique

The retransmission scheme proposed in this chapter combines the advantages of three previous methodologies proposed by Odilson T Valle et al. (2016), Xingkai Bao and Jing Li (2005, 2008), and Bin Guo et al. (2014). We use the equation proposed by Odilson T Valle et al. (2016) as a rule for forming the coefficients and the method of sending the coefficients used in Xingkai Bao and Jing Li (2005, 2008) and Bin Guo et al. (2014). The equation for the generation of the coefficients proposed in Bin Guo et al. (2014) requires that the destination node receives all messages sent in the network to be able to decode, making it impossible to use in a real network, in which message losses occur.

The coefficients are sent based on a bit-map representation. Thus, if the relay node listens to the packet from n neighbors, the corresponding position to each one of the n neighbors in bit-map is set to 1; otherwise, it will be set to 0. We consider that the relay nodes will never have a message from the coordinator to encode. In this way, we consider that the first position of the m-bit bit-map represents node 1, the second position represents node 2 and so on.

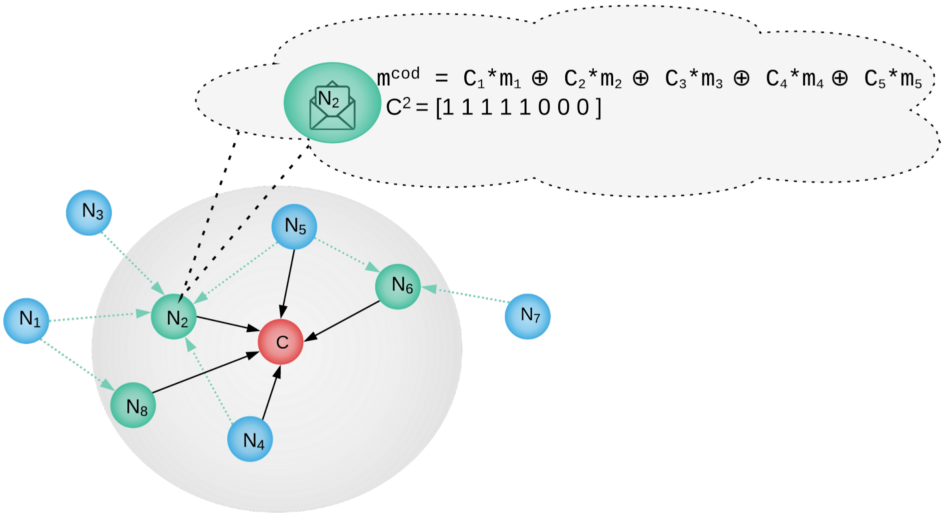
It is important to remark that sending coefficients via a bit-map technique induces a reduction in the overhead generated by sending the coefficients. A traditional RLNC technique sends a list with each of the used coefficients. If we consider that each coefficient has 8 bits and in the worst case, 255 coefficients are sent (star topology), there is an overhead of 2040 bits. Using m-bit bit-map, the overhead is reduced to m bits. Thus, in the same scenario, only 255 bits would be needed.

To use the coefficient formation rule, we modified the operation behavior of the technique proposed in Odilson T Valle et al. (2016). The authors considered that each node in the network has an identification in hexadecimal format, representing its position in the slot scale, with addresses ranging from 00 to FF. Our scheme uses the *id* assigned to the node in the formation of the network, which starts at 1 and goes up to the total number of nodes in the network, which is limited to 255 nodes, considering a star topology and that the coordinator is the node 0. Thus, the coefficients used by the relay nodes when encoding the listened messages are generated by the following formation rule:

$$c_j^i = (i + j) \quad mod \quad q \tag{24}$$

where *i* is the *id* of the *i*-th relay node, *j* is the *id* from its *j*-th neighbor, and *q* is finite field size, which was defined to $\mathbb{F}_{2^8}$ in Odilson T Valle et al. (2016). The coefficient is assigned to be 0 if the relay node does not receive one corresponding packet successfully. The coordinator node is equipped with the same formation rule to solve the coding coefficients, according to the received bit-map in each packet. Figure 39 shows an example to illustrate the structure of the bit-map and the mapping between the coding coefficients and bit-map.

Figure 39 – Node N_2 performing the network coding and creating the bit-map of the coding coefficients.



Source – Author.

Figure 39 presents a network with 8 nodes associated with the coordinator, where three of these nodes are relay nodes (N_2, N_6 and N_8). To illustrate how the structure of the bit-map works, we will consider node N_2; this relay node listened and encoded messages from neighbors N_1, N_3, N_4, N_5, and its own message. Generating the *m*-bit bit-map with the following content: 11111000. This is the bit-map that will be sent to the coordinator along with the encoded message.

When the coordinator node receives the coded message from node N_2, it checks the bit-map and applies the coefficient formation rule, knowing that node *i* sent the coded message and messages that were encoded (from which neighbors *j*), the coordinator obtains the coefficient used for each message. Then, the coordinator node has to solve the system of linear equations presented in Equation (19) to recover the original messages. According to Odilson T Valle et al. (2016), the requirement of coefficient matrix (matrix *G*, in Equation (19)) being full rank was verified and any set of coefficients that follows the coefficient formation rule presented in Equation (24) could be used as elements of the coding vector on any relay.

5.4 SIMULATION ASSESSMENTS

The network simulation tool OMNeT++ (COMMUNITY, 2011) and the WSN framework Castalia (CASTALIA, 2006) were used to assess the operation of the relay selection technique and the retransmission scheme using network coding. The open-source Solve Library *lp_solve* (SOLVE, 2007) was used to solve the optimization problem.

5.4.1 Simulation Settings

In framework Castalia, several extensions were added to the available IEEE 802.15.4e LLDN model, including the Collision Free Period (CFP), which is subdivided into guaranteed time slots (GTS) for uplink messages forwarded from the nodes to the coordinator; and the group acknowledgment (GACK) timeslot. This was necessary because Castalia still does not have a fully functional implementation of the LLDN communication mode.

The simulation assessment was performed considering networks with 21, 41, 61, 81, and 101 nodes, one of the nodes being the personal area network (PAN) coordinator. Nodes were randomly deployed in an area of 50×50 m^2, with the PAN coordinator positioned in the center. The used channel model was the free space model without time-varying. Other simulation parameters are described in Table 8.

Table 8 – Simulation Setting.

Parameters	Values
Node distribution	Random with coordinator in center
Radio	CC2420
MAC layer	IEEE 802.15.4e
Number of superframe slots	145 (5 are used by the CAP)
Data rate	250 kbps
Initial energy per nodo	18720 J
TxOutputPower	0 dBm

Source – Author.

The simulation execution time was set to 450 s, during which the coordinator is able to send up to 50 *beacons*. The radio model used was CC2420, which is compliant with the IEEE 802.15.4e PHY Standard. All the nodes use the same constant transmission power of 0 dBm. To reduce the statistical bias, each simulation was performed 60 times, reaching a confidence interval of 95%. For each simulation round, the position of the nodes around the coordinating node was randomly reorganized. That is, the distance between the coordinator and the nodes also varies in each simulation round.

Additionally, simulations were performed considering a dynamic topology, where only 50% of nodes were associated with the network at time zero and the remainder were subsequently associated in groups of 5 by 5 nodes. The first group at time instant 50 s and then all the other groups every 30 s. Considering the scenario with the highest number of nodes (100 nodes), after 320 s, all nodes were associated. Later, from the time instant 320 s of simulation, 20% of the nodes of the network randomly left the coverage of the coordinator node. This leaving operation was performed in groups of four nodes, every 10 s of simulation. Finally, all nodes again joined the network, in the same order they have left (groups of 4 in 4), from the time instant 350 s of simulation, respecting an interval of 10 s for each group, except for the case of the network with 100 nodes, where only 10% of the outgoing nodes returned.

The dynamic topology mode was designed to force the list of neighbors to undergo multiple changes during the simulation time, in order to assess the reliability of the dynamic relay selection procedure.

5.4.2 Network Coding Technique Application Scenarios

Based on the results obtained in previous works (LAURINDO, S. et al., 2018, 2020), we know that relay nodes are optimally selected. However, in those previous works, the listened messages were not really sent. Instead, it was just sent a list with the nodes listened to by each relay node. In this way, it was possible to identify if the selected relay node could hear all or almost all the nodes that the coordinator did not hear. In order to assess the delivery of messages heard by each relay node, we consider three different retransmission scenarios. Thus, it is possible to identify the impact of the retransmission step on communication when using the ORST and network coding technique together.

In all scenarios, the method of generating and sending the coefficients used in the network coding will be the method described in Section 5.3.3.

1st Scenario: The first scenario is a typical RLNC scenario. Relay nodes store all messages heard during the transmission step. In sequence, relay nodes encode all stored messages and retransmit the encoded message to the coordinator node.

2nd Scenario: The second scenario is a typical SLNC scenario. Relay nodes store all heard messages during the transmission step. However, they encode just a small number of messages among the set of listened messages. In this scenario, the network coding technique becomes sparse linear network coding. Each relay node randomly selects three messages among the listened messages; it applies the network coding technique generating a single message and retransmits it to the coordinator node.

3rd Scenario: In the third scenario, GACK was used as a resource. It contains an M-bit bitmap to indicate successful and failed transmissions in the same order as the transmissions. Thus, after all nodes carry out the transmission, the coordinator node sends a GACK message, which contains the bit map informing which messages were the ones that have failed. After receiving the GACK message, each relay node selects three messages from those that were not received by the coordinator, at random, encodes, and retransmits to the coordinator. In this scenario, we continue to apply the sparse linear network coding version. However, strategically, we only selected messages that the coordinator was unable to correctly receive in the transmission step.

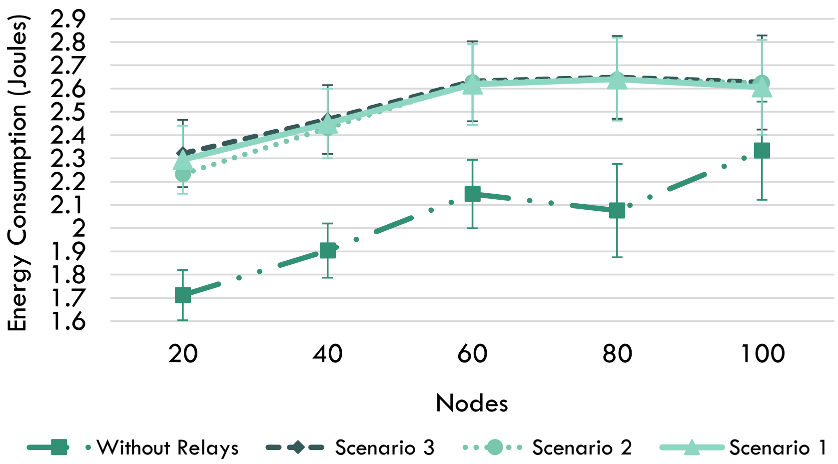
5.4.3 Simulation Assessment

The simulation assessment was performed considering the following metrics to measure the network quality performance: success rate, energy consumption, and the correlation between the average number of retransmitted messages per node and the average number of recovered messages in the decoded process.

The success rate represents the ratio between the number of sent messages and the number of messages that successfully reached the coordinator. This metric considers messages transmitted in both the transmission and retransmission attempts. In the retransmission attempts, just the messages that have been successfully decoded are considered. Energy consumption represents the average amount of energy spent by each node, obtained through the resource management module available in the Castalia framework. The average number of retransmitted messages per node represents the average number of retransmissions each node performed, i.e., the average number of coded messages sent per each node. Finally, the average number of recovered messages in the decoding process represents the average number of messages that were recovered by the coordinator node solving the linear system generated by the network coding.

Figure 40 illustrates the energy consumption of the network. It is possible to observe that communication scenarios that use network coding spend more energy. This was an expected result, considering that relay nodes remain awake longer, listening to the messages in the transmission step and retransmitting the messages in the retransmission step.

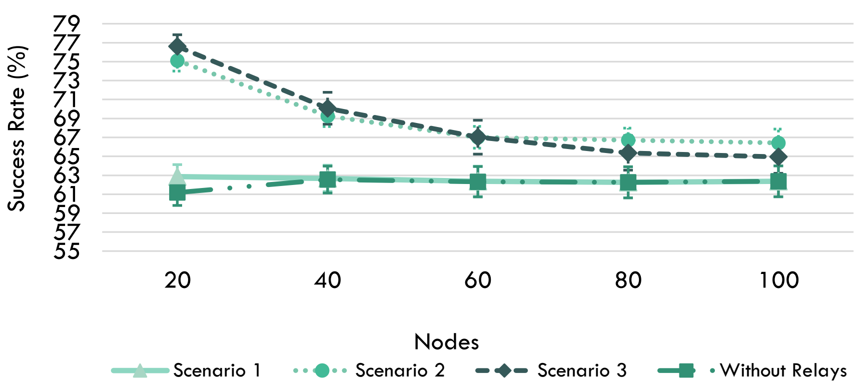
Figure 40 – Energy consumption.



Source – Author.

Figure 41 illustrates the success rate considering the three communication scenarios presented in Section 5.4.2 compared to a network without relay nodes. This result was surprisingly negative due to the small number of selected relays, as it will be shown in the following. At first sight, it was expected that the network coding linked to the ORST technique would increase the success rate of the network and consequently increase the reliability of communications. However, it is possible to observe that the behavior of Scenario 1 is similar to the network without relay nodes (where the node itself retransmits the messages for which it did not receive the ACK). Scenarios 2 and 3 show a clear improvement compared to this behavior specially for networks with less than 60 nodes.

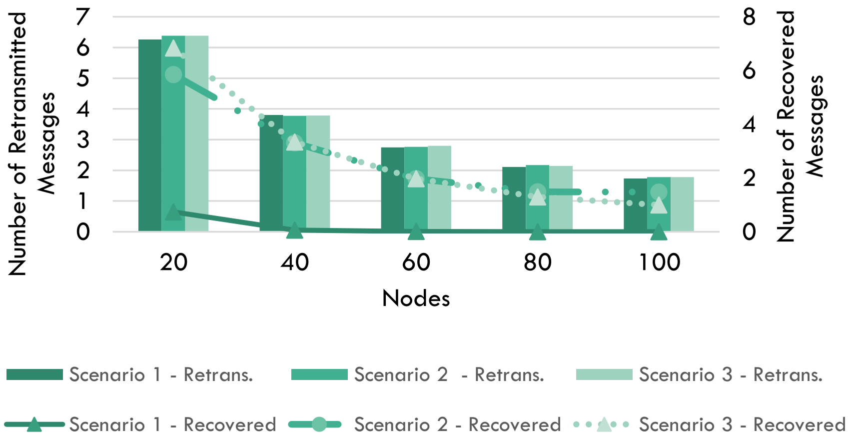
Figure 41 – Success rate.



Source – Author.

Figure 42 presents the correlation between the average number of retransmitted messages per node and the average number of recovered messages in the decoding process. It is possible to observe in Scenario 1 that the number of retransmitted messages is much greater than the number of recovered messages in the decoding process. In networks with 40, 60, 80, and 100 nodes, the average number of recovered messages was very close to zero. That is, there was almost no message recovery. In Scenarios 2 and 3, a greater number of messages was retrieved when compared to Scenario 1. However, we expected the coordinator to recover more messages, considering that each coded retransmission contains at least three messages.

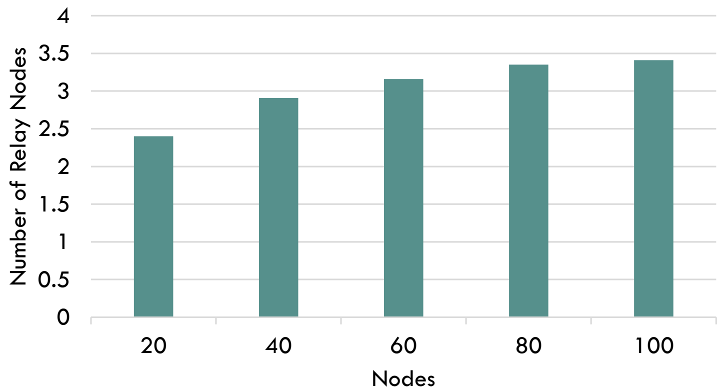
Figure 42 – Correlation between the average number of retransmitted messages per node and the average number of recovered messages in the decoding process.



Source – Author.

Investigating the obtained results, it is clear that the number of relay nodes has a direct impact upon the operation of the network coding scheme. Figure 43 presents the average number of selected relay nodes, considering all scenarios. As it can be seen, the average number of relay nodes is very small. Even considering networks with 100 nodes, the average number of relay nodes is smaller than 4. Analyzing this problem from an equation solving perspective, the coordinator node in order to be able to decode the received messages, solves a linear system that must result in a single solution. A linear system with more unknowns than equations may not have any solution or have an infinite number of solutions, but it will never have just one solution. In the context of network coding, the number of equations corresponds to the number of coded retransmissions received by the coordinator, and the number of unknowns corresponds to the number of different messages that were coded, which the coordinator did not receive successfully in the transmission step.

Figure 43 – Average number of relay nodes.



Source – Author.

This fact explains why Scenario 1 performed so poorly. In this communication scenario, each relay node codes all the heard messages. That is, there were more unknowns than equations. Thus, it was impossible to decode the received messages. This is a typical problem that often arises in network coding applications.

We have also analyzed why there was no significant improvement in the success rate of Scenarios 2 and 3. Two facts can be considered: first, the number of messages that did not successfully reach the coordinator in the transmission stage was high, around 40%, and second, the relay nodes coded only a small number of messages (three messages for each retransmission). Thus, even when the coordinator recovers some messages in the decoding process, a large number of messages that were lost were not retransmitted, which resulted in unreliable communication.

To maximize the success rate of the network, we propose a different communication approach, represented by two new scenarios, Scenarios 4 and 5. In Scenario 4, we increased the number of relay nodes and continued using SLNC. In Scenario 5, we do not use network coding. Each relay node retransmits the messages it listened to one by one, as described below.

4th Scenario: This scenario simply increases the number of relay nodes. When performing the relay selection, the coordinator node selects two auxiliary nodes for each relay node and signals this information in the beacon. Thus, the nodes that will cooperate and assist in the retransmission know that they have been selected. In this way, if there are three relay nodes, there will be six auxiliary nodes, two for each relay node. Auxiliary nodes are selected as follows: For each neighbor of the relay node that communicates with the coordinator, it is checked the list of heard nodes. The neighbors that it listens to and, at the same time, the coordinator does not listen to are counted. The two neighbors of the relay node, which communicate with the coordinator and have the largest number of neighbors that do not communicate with the coordinator, will be selected to assist the relay node.

In the retransmission step, the relay node C_i intersects its list of listened messages (L_{C_i}), with the list of messages lost by the coordinator ($L_{LostCoord}$) and with the list of messages heard from the auxiliary nodes (L_{Aux_j} and $L_{Aux_{j+1}}$), according to Equation (25).

$$I = ((L_{C_i} \cap L_{LostCoord}) \cap (L_{Aux_j} \cap L_{Aux_{j+1}})) \tag{25}$$

The result of Equation (25) are the common messages that the coordinator needs and that the relay node and the two auxiliary nodes also have. The auxiliary nodes of each relay also perform the intersection operation presented in Equation (25). The list of messages resulting from Equation (25) (I) is ordered by the *Id* of the nodes that sent them. Thus, both in the relay nodes and in the auxiliary nodes, the list I presents the same messages in the same order. When coding and retransmitting, the relay node selects a message that only it listened to and the first two messages from the list I resulting from the intersection. Retransmitting a coded message (M_{C_i}) containing three messages listened to ($M_{C_i} = m_{C_i} + m_{I_1} + m_{I2}$). The auxiliary node Aux_j selects the first two messages resulting from the intersection ($M_{Aux_j} = m_{I_1} + m_{I_2}$). The second auxiliary node Aux_{j+1} selects its own message and the second message from the list resulting from the intersection ($M_{Aux_{j+1}} = m_{Aux_{j+1}} + m_{I_2}$).

This organization in the selection of messages that will be encoded and retransmitted by the relay and auxiliaries nodes allows the coordinator to effectively solve the linear system, decoding the lost messages and recovering them.

5th Scenario: The relay nodes will retransmit each of the messages listened to in individual slots without using NC. The coordinator node is the one who will allocate slots for each relay node, according to the number of messages that each relay will retransmit.

For the correct operation of the network, there is a configuration period. This period precedes each of the relay selections in the ORST technique (LAURINDO, S. et al., 2018). It is during this period that the coordinator receives from all nodes, which communicate directly with it, the neighbor's list of each node. The neighbor's list is a bit-map, where each index of the bit-map represents a network node, and the content of the bit-map in "1" represents that the nodes are neighbors and "0" otherwise. The coordinator node uses this information from the neighborhood of each relay node together with GACK information from the previous Beacon Interval, to determine the number of slots that each relay node will receive to carry out the retransmissions.

The process to allocate slots for each relay node occurs as follows: First, the coordinator checks its GACK bit-map to identify which messages were lost. Then, the bit-map that represents the neighborhood of each relay node is updated, keeping in "1" only the positions that represent the listened neighbors and at the same time the messages not received by the coordinator by direct transmission. Up to this stage, the coordinator can identify how many slots each relay node would need, if it was to retransmit all the messages it heard, among those that the coordinator lost. The coordinator keeps this information for each relay node. In order to optimize the allocation of the slots and to prevent the relay nodes retransmiting repeated messages, the binary AND operation is performed with the neighbor's list of each relay node. Thus, it is possible to identify which relay nodes have heard the same messages, for example, in a network with five nodes, in which nodes N_1 and N_3 are relay nodes. The coordinator node lost messages from nodes N_2, N_4 and N_5. Considering the neighbors' list of each relay nodes already updated with the GACK information, the neighbors' list of node N_1 is represented by $N_1 = N_4, N_5$ and the neighbor's list of node N_3 is represented by $N_3 = N_2, N_4$ and N_5. The illustration of the AND operation performed by the coordinator, in this example, is shown below:

$$|0|0|0|1|1|AND|0|1|0|1|1| = |0|0|0|1|1|$$

The binary AND operation will result in the elements that both relay nodes heard. After identifying which messages were listened to by more than one relay, the coordinator verifies which relay node has the least number of messages to be retransmitted and selects it to be the retransmitter. The relay that will retransmit is selected considering the number of messages to be retransmitted to balance the energy consumption among the relay nodes. This is because the more messages each relay has to retransmit, the greater the energy consumption of this node will be. In the cited example, both relay nodes listened to messages from nodes N_4 and N_5. Disregarding the messages they both listened to, the node N_3 has an element in its relay list (the message from the node N_2), and the node N_1 does not have any element. Thus, the first element that the two relay nodes hear will be assigned to node N_1, and the second can be assigned to anyone since both have the same number of messages to be retransmitted.

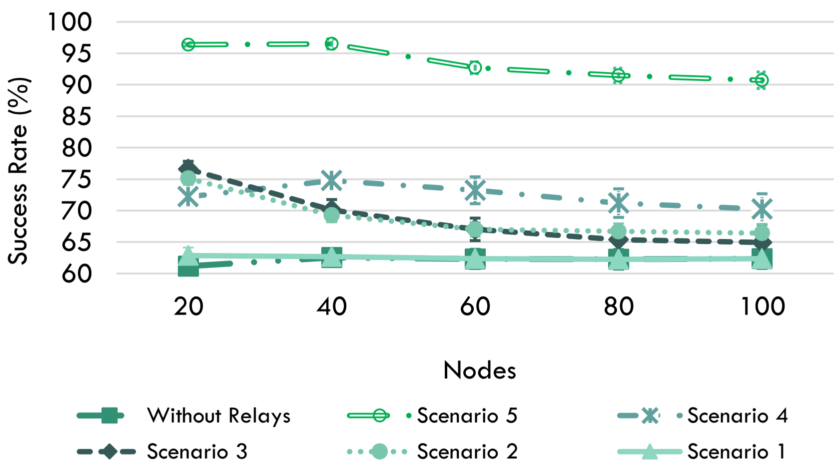
After selecting the relay node that will retransmit each message, the coordinator decreases the number of slots that would be assigned to the other relay node that had heard the same message and that was not assigned to it. If at the end of this process, it is identified that the number of messages lost is greater than the number of available slots, inevitably some messages will not be retransmitted.

We limit the total number of slots between transmissions and retransmissions to 140, as the goal is to maximize the network's success rate with the least number of messages being retransmitted. In a network that uses only Automatic Repeat Request (ARQ) protocol, each node in the network performs retransmissions whenever it does not receive an ACK. This way, the number of retransmissions can be even greater than the number of nodes in the network, considering that there are approaches that allow a node to perform the same retransmission a number x of times in case it does not receive an ACK. The objective of this scenario is that even considering a network of 100 nodes, it is possible to retransmit without needing a retransmission slot for each node in the network. Thus, the success rate will be maximized without increasing the beacon interval period between transmissions.

The information of which relay node should send the message will be sent in the next GACK message. Thus, when the relay node receives the message of GACK, it checks if, in any of the messages that the coordinator has lost, there is its own *id* signaling that it must be retransmitted. The relay knows that the messages lost by the coordinator must be retransmitted if, in the GACK, the *id* of the retransmitter is either marked as "0" or with its own *id*. If it is zero, it means that only it heard the missed message, and if it is the *id* itself, it means that more nodes listened, but he was the one selected to retransmit.

Figure 44 illustrates the success rate considering all the scenarios. It is possible to observe that Scenario 4 presents a significant improvement over Scenarios 1, 2, and 3. However, the scenario that presented the best results was Scenario 5, which in networks with 20 and 40 nodes the success rate was above 95%, and in networks with 60, 80, and 100 nodes the success rate was maintained above 90%.

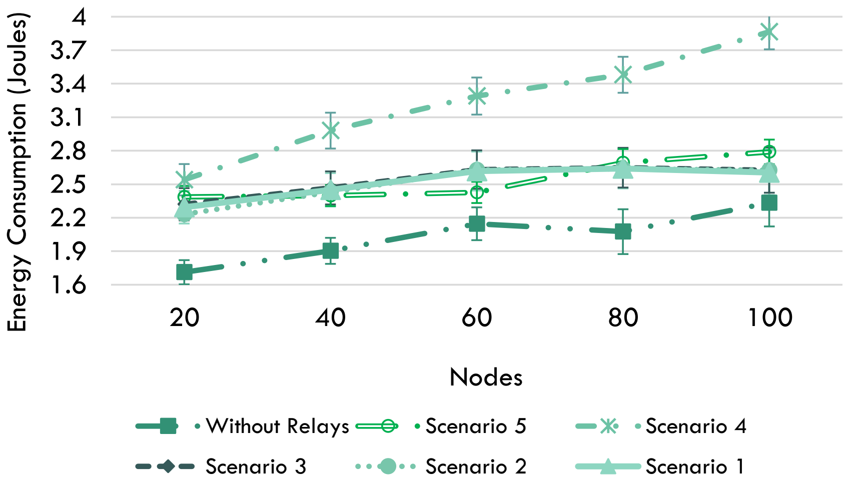
Figure 44 – Success rate.



Source – Author.

Finally, Figure 45 presents the energy consumption in all scenarios. It is possible to observe that Scenario 4, where auxiliary nodes were selected, presented a higher energy consumption. This is understandable since a greater number of nodes will have the radio on for the entire transmission stage listening to neighbors. Scenario 5 presented the energy consumption similar to Scenarios 1, 2, and 3. Thus, Scenario 5 was the one with the best results, and it can be considered the best retransmission scheme to be used linked to the ORST technique.

Figure 45 – Energy consumption.



Source – Author.

5.5 CONCLUSIONS

Nowadays, with the industry 4.0 paradigm, smart devices with sensing communicating and actuating capabilities are common in the industrial environment. Thus, sensors and actuators are integrated into the environment and communicate transparently, growing the use of WSNs. However, WSNs present challenges in maintaining reliable communication, being necessary to apply extra mechanisms to improve their performance.

Cooperative communication has been proposed to enhance the reliability of wireless communication. When applying this type of solution, it is required to determine which nodes will be the relay nodes and how they collaborate in the retransmission process. This chapter focused on combining the adequate selection of relay nodes and retransmission techniques. It was proposed the use of the ORST scheme, whose target is to adequately select relay nodes without generating overheads or excessive energy consumption, combined with four different retransmission techniques, three of them applying network coding techniques.

Differently from what we assumed when starting this study, the ORST technique working together with network coding approaches did not present interesting results. The main reason was due to the fact that ORST technique selects a small number of relay nodes. As a consequence, each relay node generates a new encoded packet with a large number of received messages, resulting in a small number of equations in the linear system and the coordinator will not be able to decode incoming messages. A key aspect of the linear network coding technique is that a node needs to receive a number of coded messages greater than or equal to the number of original messages, to successfully decode the original set of messages. However, the efficiency of the ORST technique is demonstrated when the relay nodes retransmit a set of the messages listened to in individual slots without using network coding; the success rate is greater than 90% and energy consumption is only slightly above the case without relays.

As future work, we intend to assess the implementation feasibility of the proposed schemes using available COTS (commercial off-the-shelf) WSN nodes. This implementation has been done in a centralized topology, where the PAN coordinator can be implemented as follows: as a device that has extra resources to perform the calculation or as a device with limited computational resources but which is connected to a computer that performs the processing and returns the solution to the coordinator. All the other nodes, on the other hand, can be devices with limited computational resources.

6 CONCLUSIONS

In this chapter, we review the main motivations and objectives of this thesis, present an overview of the various implementations, and discuss the main conclusions of the research work carried out in this thesis. Finally, some improvements to the current proposal are suggested for future work.

6.1 REVIEW OF MOTIVATIONS AND OBJECTIVES

The reliability of communication in WSNs is a topic that requires ongoing study, as challenges remain to be overcome that arise from interference and data loss, among other factors. Cooperative diversity techniques and network coding techniques have been used to minimize these challenges. When a cooperative diversity approach is used, the method applied to perform relay selection plays an important role, since the selected nodes will retransmit messages from nodes that do not communicate directly with the coordinator node, thus maximizing the gain from the cooperative diversity scheme.

In this context, the work in this thesis makes use of the concepts of the cooperative diversity technique, and proposes a new technique for selecting relay nodes in a WSN. In addition, a network coding technique is proposed that can act in the retransmission stage to deal with communication in a holistic way. The work presented in this thesis sought to answer the research question posed in the introduction, as follows:

– “Is it possible to increase the reliability of WSN communication by proposing a relay selection technique, considering a set of relevant criteria combined with retransmission techniques?”

6.2 THESIS OVERVIEW

This thesis began with a systematic review to identify open research issues related to relay selection techniques. The state-of-the-art proposals were analyzed and a classification framework, based on the parameters used to carry out relay selection, was proposed. Afterwards, a new approach, called ORST (optimized relay selection technique), was designed. The ORST was modeled as an optimization problem, and four parameters were considered in the objective function. These parameters were selected in view of their importance to the operation of a WSN (number of neighbors; remaining energy; Received Signal Strength Indicator (RSSI) among node x_i and its neighbor's nodes x_j; and history of successful transmission rates). The aim of the technique is to select the smallest number of relay nodes, while at the same time ensuring that each node in the network has a neighbor that is a relay node. The proposed technique was evaluated through simulations carried out in OMNeT++, and was compared with three other state-of-the-art techniques. The results showed that the proposed technique outperformed the others on all the metrics evaluated, thus demonstrating that selection of the relay nodes was optimal.

Following this, two further studies were performed. The first analyzed the impact of each of the parameters used in the objective function, and new objective functions were modeled based on each possible combination of parameters (giving a total of 14 objective functions). Each of these objective functions was assessed using the network simulation tool OMNET++ and the WSN framework Castalia. This outcome of this experiment was an optimization of the objective function in which only the amount of energy remaining in the nodes was considered as a parameter. Thus, the objective function was simplified, and the results continued to show a high probability of success in delivering messages to the destination.

The second experiment analyzed three different algorithms as possible options for solving the relay selection problem, which could be classified as an SCP. The three approaches were the branch and bound, genetic and greedy algorithms. An assessment was performed using OMNET++ and the WSN framework Castalia, and the network quality performance was measured based on the execution time required by each algorithm to perform the relay selection and other metrics such as the success rate, the number of cooperations per node, energy consumption, and the percentage of duplicate (useless) messages. The results showed that the branch and bound algorithm, additionally selecting the optimal set of relay nodes, select them into one timeslot, which ensures the proper operation of the WNS with the best results for assessed metrics.

In addition, a systematic review of the literature on retransmission mechanisms based on network coding techniques was carried out, and based on the results, a classification framework for network coding methods was proposed. The most relevant state-of-the-art works were reviewed, with a focus on how selection was carried out and how the coefficients used to encode the messages were sent. This analysis suggested ways of reducing the overhead generated by the transport of the coefficients; a new network coding technique was proposed in which only a bitmap was sent rather than sending all the coefficients used in encoding, and based on this, the destination node was able to calculate the coefficients that were used.

Four different retransmission techniques were then proposed, of which three used network coding algorithms. These four retransmission techniques were assessed together with ORST using OMNET++ and Castalia. An unexpected result was that when both techniques were applied to the WSN, the network exhibited a drop in performance, and the network success rate resembled that of a network without cooperative diversity. An analysis indicated that since ORST selects a small number of relay nodes, a small number of coded messages was sent; however, the destination node needs to receive a number of coded messages that is greater than or equal to the number of original messages in order to successfully decode the original set of messages, meaning that ORST does not work properly with a network coding technique. In contrast, the efficiency of this technique was high in a scenario where the relay nodes retransmitted each of the overheard messages in individual slots without using network coding; in this case, the success rate was greater than 90% and the energy consumption was only slightly higher than in the case without relays.

In this scenario, a WSN that uses ORST will have an optimal relay selection, which will maximize the success rate of the network without generating excessive energy consumption. In the retransmission step, the ideal case is for messages to be retransmitted individually. The configuration proposed in this thesis allows each relay to receive a number of retransmission slots and prevents the messages retransmitted by the relay nodes from being repeated, thus avoiding the sending of redundant messages and minimizing energy consumption.

6.3 LIMITATIONS OF THE RESEARCH

This thesis focused on improving the reliability of communications in WSNs through the use of relay selection and retransmission techniques based on network coding. However, it did not address all of the issues associated with these processes. The scope of this work was limited to a WSN with a star topology, in which access to the medium for sending messages was based on TDMA (Time Division Multiple Access) as defined in the LLDN standard for IEEE 802.15.4e. The use of Forward Error Correction (FEC) techniques and network coding approaches at the physical layer were outside the scope of this work.

6.4 FUTURE WORK

The problem of reliable communication is highly relevant in the context of WSNs. This thesis focused on a specific area of cooperative diversity, and proposed a relay selection technique that can be used together with a retransmission technique to maximize the reliability of communication. However, there are still several related topics that could be further studied and improved, which could not be covered in this work due to time constraints. Directions for future work could include the following:

- An analysis of the use of relay nodes in terms of the coordinator node (downlink) to verify that the relay nodes that act on the uplink would also be able to retransmit the beacon message to nodes that did not receive it due to occasional interference. If this was the case, a situation could be avoided in which nodes fail to transmit data because they did not receive the synchronization message.

- Expansion of the use of the relay selection technique to a cluster tree topology, and the introduction of the necessary changes and improvements. This would allow us to analyze the feasibility of implementing the concepts underlying ORST in other topologies, with the aim of using them in large-scale networks.
- Improving the method used in the retransmission step to determine which relay node will retransmit a certain message when more than one relay is listening to the same message. Currently, the relay with the least number of messages to retransmit is chosen when two relays hear the same message. However, other factors could be analyzed, such as the amount of energy remaining in the nodes, which would prevent the energy resources of a relay node from becoming rapidly depleted.

In addition to the suggestions put forward above, one that is already under development is the implementation of ORST using available COTS (commercial off-the-shelf) WSN nodes. WM100-Duino and Kitrfa1 boards with ATMEGA256RFR2 and ATMEGA128RFR1 (BUDKE, 2015) microcontrollers have been used, and the communication protocol used is IEEE 802.15.4e with the LLDN operating mode. This protocol has been implemented on the LWmesh stack, which enables the implementation of the IEEE 802.15.4 protocol. However, to adapt to the requirements of the LLDN, numerous changes have been made.

During this implementation, two major challenges were encountered. The first was related to the configuration of the timers to ensure the operation of the IEEE 802.15.4e protocol with LLDN, as it was found to be necessary to control all of the timers in the application layer. The second problem arose was when it was discovered that the hardware used did not support the solver needed for the relay selection problem. To address this problem, the coordinator node was connected to a computer via serial communication. In this approach, coordinator node sends the data necessary to solve the problem and the computer processes it, returning to the coordinator node the IDs of the nodes that were selected as relay nodes. During preliminary testing with 10 nodes, the entire process of sending, processing data and receiving the results took an average of 8.3 ms. A computer with the Windows 10 operating system, 8 GB of memory and a Core i5 7th generation processor was used. Most of this time (8 ms) was spent on sending data via serial communication, meaning that if a board with greater computational power was used, the time required to perform relay selection could be drastically reduced.

At present, the proposed technique has only been partially implemented, but is already able to perform relay selection. A scheme for triggering for a new relay selection dynamically, based on the success rate of the network, is still being tested. The retransmission step based on the configurations proposed in Chapter 5 is also being implemented.

The results and a discussion of the implementation of ORST using available COTS WSN nodes will be reported in a conference paper that is currently under development.

REFERENCES

ADRIANO, José D; ROSARIO, Elcio Carlos do. Wireless Sensor Networks in Industry 4.0: WirelessHART and ISA100. 11a. In: 13TH IEEE International Conference on Industry Applications (INDUSCON). São Paulo, Brasil: [s.n.], Nov. 2018. P. 924–929.

AHMAD, Ishfaq; SHAH, Khalil; ULLAH, Saif. Military applications using wireless sensor networks: A survey. **Int. J. Eng. Sci**, v. 6, n. 6, p. 7039, 2016.

AHMED, Mohammad Helal Uddin; RAZZAQUE, Md. Abdur; HONG, Choong Seon. DEC-MAC: delay- and energy-aware cooperative medium access control protocol for wireless sensor networks. **Annals of Telecommunications - Annales Des Télécommunications**, v. 68, n. 9-10, p. 485–501, 2013. Available from: http://link.springer.com/10.1007/s12243-012-0330-y.

AI, Yun; CHEFFENA, Michael. On Multi-Hop Decode-and-Forward Cooperative Relaying for Industrial Wireless Sensor Networks. **Sensors**, v. 17, n. 4, p. 695, 2017. ISSN 1424-8220. DOI: 10.3390/s17040695.

AKHTARI, Sara; MOGHIM, Neda; MAHDAVI, Mojtaba. Middleman covert channel establishment based on MORE routing protocol using network coding in ad hoc networks. **International Journal of Communication Systems**, v. 33, n. 7, p. 1–14, 2020. ISSN 10991131. DOI: 10.1002/dac.4320.

AKYILDIZ, Ian F.; SU, Weilian; SANKARASUBRAMANIAM, Yogesh; CAYIRCI, Erdal. A survey on sensor networks. **IEEE Communications Magazine**, v. 40, n. 8, p. 102–105, 2002. ISSN 01636804.

ALEMDAR, Ali; IBNKAHLA, Mohamed. A Survey of Wireless Sensor Networks: Technologies, Challenges, and Future Trends. **Adaptation and Cross Layer Design in Wireless Networks**, CRC Press, p. 84, 2018.

ALKHAYYAT, Ahmed; GAZI, Orhan; SADKHAN, Sattar B. The Role of Delay and Connectivity in Throughput Reduction of Cooperative Decentralized Wireless Networks. **Mathematical Problems in Engineering**, v. 2015, p. 10, 2015. ISSN 15635147. DOI: 10.1155/2015/294016.

AMMARI, Habib M.; GOMES, Nicholas; JACQUES, Matthew; MAXIM, Bruce R.; YOON, David H. A Survey of Sensor Network Applications and Architectural Components. **Ad Hoc & Sensor Wireless Networks**, v. 25, n. 1-2, p. 1–44, 2015.

ANANIASHVILI, Natela. Solution of Problem of Set Covering by Means of Genetic Algorithm. **Computer Science & Telecommunications**, v. 46, n. 2, p. 16–23, 2015.

ANDRE, Torsten; MARCHENKO, Nikolaj; MASOOD, Wasif; BETTSTETTER, Christian. Measurement-Based Analysis of Adaptive Relay Selection in Industrial Wireless Sensor Networks. In: MODELING & Optimization in Mobile, Ad Hoc & Wireless Networks (WiOpt). [S.l.]: 2013 11th International Symposium on IEEE, 2013. P. 113–118.

ASAM, Muhammad; HAIDER, Zeeshan; JAMAL, Tauseef; GHUMAN, Kashif; AJAZ, Aleena. Novel Relay Selection Protocol for Cooperative Networks. **CoRR**, abs/1911.07764, 2019.

AZIZ, Azlan ABD; GHANI, Hadhrami AB. Energy Efficiency in Dynamic Cluster Selection for Cooperative Wireless Sensor Networks. In: IEEE. REGION Ten Symposium (Tensymp). [S.l.: s.n.], 2019. P. 155–159.

BACCOUR, Nouha; KOUBÂA, Anis; YOUSSEF, Habib; JAMÂA, Maissa B; ROSÁRIO, Denis do; ALVES, Mário; BECKER, Leandro B. F-LQE: A Fuzzy Link Quality Estimator for Wireless Sensor Networks. In: MARRÓN, Pedro José; WHITEHOUSE, Kamin (Eds.). **Wireless Sensor Networks: 8th European Conference**. Bonn, Germany: Springer, 2010. P. 240–255.

BAO, Jianrong; WU, Jiawen; LIU, Chao; JIANG, Bin; TANG, Xianghong. Optimized power allocation and relay Location selection in cooperative Relay Networks. **Wireless Communications and Mobile Computing - Wiley**, v. 9, n. 3, p. 1–10, 2017. ISSN 1943023X.

BAO, Xingkai; LI, Jing. Adaptive Network Coded Cooperation (ANCC) for Wireless Relay Networks: Matching Code-on-Graph with Network-on-Graph. **IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS**, v. 7, n. 2, p. 574–583, 2008.

BAO, Xingkai; LI, Jing. Matching code-on-graph with network-on-graph: Adaptive network coding for wireless relay networks. In: 43RD Annual Allerton Conference on Communication, Control and Computing. [S.l.: s.n.], 2005. P. 386–395.

BARA'A, A Attea; HAMEED, Sarab M. A Genetic Algorithm for Minimum Set Covering Problem in Reliable and Efficient Wireless Sensor Networks. **Iraqi Journal of Science**, Baghdad University, v. 55, n. 1, p. 224–240, 2014.

BARCELO-ORDINAS, Jose M; CHANET, Jean-Pierre; HOU, K-M; GARCÍA-VIDAL, J. A survey of wireless sensor technologies applied to precision agriculture. In: PRECISION agriculture'13. [S.l.]: Springer, 2013. P. 801–808.

BASARAN, Semiha Tedik; KURT, Gunes Karabulut. Wireless Channel Induced Coding. **IEEE Communications Letters**, IEEE, v. 23, n. 10, p. 1688–1691, 2019.

BHUTE, Yogesh; RAUT, AR. A survey on relay selection strategies in cooperative wireless network for capacity enhancement. **International Journal of Computer Applications**, Foundation of Computer Science, v. 65, n. 25, p. 12–17, 2013.

BILAL, Nehme; GALINIER, Philippe; GUIBAULT, Francois. A New Formulation of the Set Covering Problem for Metaheuristic Approaches. **ISRN Operations Research**, Hindawi Publishing Corporation, v. 2013, 2013.

BLETSAS, Aggelos; KHISTI, Ashish; REED, David P.; LIPPMAN, Andrew. A simple cooperative diversity method based on network path selection. **IEEE Journal on Selected Areas in Communications**, v. 24, n. 3, p. 659–672, 2006. ISSN 07338716. DOI: `10.1109/JSAC.2005.862417`.

BOUABDELLAH, Kechar; NOUREDDINE, Houache; LARBI, Sekhri. Using wireless sensor networks for reliable forest fires detection. **Procedia Computer Science**, Elsevier, v. 19, p. 794–801, 2013.

BRANTE, Glauber; PERON, Guilherme De Santi; SOUZA, Richard Demo; ABRÃO, Taufik. Distributed Fuzzy Logic-Based Relay Selection Algorithm for Cooperative. **Sensors**, v. 13, n. 11, p. 4375–4386, 2013.

BUDKE, Gerson. **Wireless Module IEEE 802.15.4/Zigbee WM-100-DUINO**. [S.l.: s.n.], 2015. https://gfbudke.wordpress.com. 23 March 2021.

BURATTI, Chiara; MARTALÒ, Marco; FERRARI, Gianluigi; VERDONE, Roberto. **Sensor Networks with IEEE 802.15.4 Systems**. 1. ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. (Signals and Communication Technology). ISBN 978-3-642-17489-6. DOI: `10.1007/978-3-642-17490-2`.

CASTALIA. **Framerwork Castalia**. [S.l.: s.n.], 2006. https://github.com/boulis/Castalia. 05 Apr 2018.

CHANDU, Drona Pratap. Improved Greedy Algorithm for Set Covering Problem. **SSRG International Journal of Computer Science and Engineering**, v. X, p. 1–4, 2015.

CHEBLI, Lena; HAUSL, Christoph; ZEITLER, Georg; KOETTER, Ralf. Cooperative uplink of two mobile stations with network coding based on the WiMax LDPC code. In: IEEE. GLOBAL Telecommunications Conference (GLOBECOM). [S.l.: s.n.], 2009. P. 1–6.

CHEIKH, Mohamad El; SIMPSON, Oluyomi; SUN, Yichuang. Energy Efficient Relay Selection Method for Clustered Wireless Sensor Network. In: 23TH European Wireless Conference. Berlin, Offenbach: In European Wireless 2017, 2017. P. 92–97.

CHEN, F.; WANG, N.; GERMAN, R.; DRESSLER, F. Performance Evaluation of IEEE 802.15.4 LR-WPAN for Industrial Applications. In: FIFTH Annual Conference on Wireless on Demand Network Systems and Services. [S.l.: s.n.], 2008. P. 89–96. DOI: 10.1109/WONS.2008.4459361.

CHEN, Feng; WANG, Nan; GERMAN, Reinhard; DRESSLER, Falko. Performance Evaluation of IEEE 802.15.4 LR-WPAN for Industrial Applications. In: FIFTH Annual Conference on Wireless on Demand Network Systems and Services. Garmisch-Partenkirchen, Germany: [s.n.], 2008. P. 89–96.

CHEN, Ye; LIU, Wei; WANG, Tian; DENG, Qingyong; LIU, Anfeng; SONG, Houbing. An adaptive retransmit mechanism for delay differentiated services in industrial WSNs. **Eurasip Journal on Wireless Communications and Networking**, EURASIP Journal on Wireless Communications and Networking, v. 2019, n. 1, 2019. ISSN 16871499. DOI: 10.1186/s13638-019-1566-2.

COMMUNITY, OpenSim. **OMNET++**. [S.l.: s.n.], 2011. https://omnetpp.org/. Online; accessed 05-Apr-2018.

CONSTANTINO, Ademir Aparecido; SANTOS, Andréia Alves dos; ARAUJO, Silvio Alexandre de. A New Genetic Algorithm for the Set Covering Problem. **Varia Scientia**, v. 10, n. 17, p. 147–162, 2010.

CORMEN, Thomas H.; LEISERSON, Charles E.; RIVEST, Ronald L.; STEIN, Clifford. **Introduction to Algorithms, Third Edition**. 3rd. [S.l.]: The MIT Press, 2009.

CRAWFORD, Broderick; SOTO, Ricardo; BERRIOS, Natalia; OLGU, Eduardo. Solving the Set Covering Problem Using the Binary Cat Swarm Optimization Metaheuristic. **World Academy of Science, Engineering and Technology**, v. 10, n. 3, p. 113–117, 2016.

DARGIE, Waltenegus; POELLABAUER, Christian. **Fundamentals of Wireless Sensor Networks: theory and practice**. 1. ed. Chichester, UK: John Wiley & Sons, Ltd, July 2010. P. 311. ISBN 9780470666388. DOI: `10.1002/9780470666388`.

DONG, Dhawa Sang; POKHREL, Yagnya Murti; GACHHADAR, Anand; MAHARJAN, Ram Krishna; QAMAR, Faizan; AMIRI, Iraj Sadegh. Resource tuned optimal random network coding for single hop multicast future 5g networks. **International Journal of Electronics and Telecommunications**, v. 65, n. 3, p. 463–469, 2019. ISSN 23001933. DOI: `10.24425/ijet.2019.129800`.

ELSAMADOUNY, Ahmed; HASNA, Mazen; KHATTAB, Tamer; ABUALSAUD, Khalid; YAACOUB, Elias. On the delay of finite buffered multi-hop relay wireless internet of things. **IEEE Vehicular Technology Conference**, IEEE, September, p. 1–7, 2019.

EREMEEV, Anton V. Restarting a Genetic Algorithm for Set Cover Problem Using Schnabel Census. In: 7TH International Conference on Optimization Problems and their Applications. Omsk, Russia: [s.n.], July 2018. P. 109–117.

ETEZADI, Farrokh; ZARIFI, Keyvan; GHRAYEB, Ali; AFFES, Sofiene. Decentralized Relay Selection Schemes in Uniformly Distributed Wireless Sensor Networks. **IEEE Transactions on Wireless Communications**, v. 11, n. 3, p. 938–951, 2012. ISSN 1536-1276. DOI: `10.1109/TWC.2012.010312.101314`.

FAHMY, Hossam Mahmoud Ahmad. Introduction. In: CONCEPTS, Applications, Experimentation and Analysis of Wireless Sensor Networks. Cham: Springer International Publishing, 2021a. P. 3–52.

FAHMY, Hossam Mahmoud Ahmad. WSN Applications. In: CONCEPTS, Applications, Experimentation and Analysis of Wireless Sensor Networks. Cham: Springer International Publishing, 2021b. P. 67–232. ISBN 978-3-030-58015-5. DOI:

10.1007/978-3-030-58015-5_3. Available from:
https://doi.org/10.1007/978-3-030-58015-5_3.

FEIZI, Soheil; LUCANI, Daniel E; SØRENSEN, Chres W; MAKHDOUMI, Ali; MÉDARD, Muriel. Tunable sparse network coding for multicast networks. In: IEEE. INTERNATIONAL Symposium on Network Coding (NetCod). [S.l.: s.n.], 2014. P. 1–6.

FERDOUSE, Lilatul; ANPALAGAN, Alagan. Relay Selection Based on Bayesian Decision Theory in Cooperative Wireless Networks Sélection de relais basée sur la théorie de décision de Bayes dans les réseaux sans fil coopératifs. **Canadian Journal of Electrical and Computer Engineering**, v. 38, n. 2, p. 116–124, 2015. DOI: 10.1109/CJECE.2014.2386698.

FRAGOULI, Christina; LE BOUDEC, Jean-Yves; WIDMER, Jörg. Network coding: an instant primer. **ACM SIGCOMM Computer Communication Review**, ACM, v. 36, n. 1, p. 63–68, 2006.

FRAGOULI, Christina; SOLJANIN, Emina. Network Coding Fundamentals. **Foundations and Trends® in Networking**, v. 2, n. 1, p. 1–133, 2006.

FRAGOULI, Christina; SOLJANIN, Emina, et al. Network coding fundamentals. **Foundations and Trends® in Networking**, Now Publishers, Inc., v. 2, n. 1, p. 1–133, 2007.

FREIRE, Patrícia de Sá. **Aumente a qualidade e a quantidade de suas publicações científicas: Manual para Elaboração de Projetos e Artigos Científicos**. 1. ed. Curitiba: CRV, 2013. P. 90. ISBN 978-85-8042-815-5.

GAIKWAD, P. P.; GABHANE, J. P.; GOLAIT, S. S. A survey based on Smart Homes system using Internet-of-Things. In: 2015 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC). [S.l.: s.n.], Apr. 2015. P. 0330–0335. DOI: 10.1109/ICCPEIC.2015.7259486.

GAO, De-Yun; ZHANG, Lin-Juan; WANG, Hwang-Cheng. Energy saving with node sleep and power control mechanisms for wireless sensor networks. **The Journal of China Universities of Posts and Telecommunications**, Elsevier, v. 18, n. 1, p. 49–59, 2011.

GOKTURK, M. Sarper; GURBUZ, Ozgur. Cooperation with multiple relays in wireless sensor networks: Optimal cooperator selection and power assignment. **Wireless Networks**, v. 20, n. 2, p. 209–225, 2014. ISSN 10220038. DOI: 10.1007/s11276-013-0585-x.

GUBBI, Jayavardhana; BUYYA, Rajkumar; MARUSIC, Slaven; PALANISWAMI, Marimuthu. Internet of Things (IoT): A vision, architectural elements, and future directions. **Future Generation Computer Systems**, Elsevier B.V., v. 29, n. 7, p. 1645–1660, 2013. ISSN 0167739X.

GUNGOR, VC; HANCKE, GP. Industrial wireless sensor networks: challenges, design principles, and technical approaches. **IEEE Transactions on Industrial**, v. 56, n. 10, p. 4258–4265, 2009.

GUNGOR, Vehbi C; HANCKE, Gerhard P. Industrial wireless sensor networks: Challenges, design principles, and technical approaches. **IEEE Transactions on industrial electronics**, IEEE, v. 56, n. 10, p. 4258–4265, 2009.

GUO, Bin; LIU, Yu; ZHOU, Chi. Exploit network coding over GF(2q) for multi-user cooperative wireless networks. **International Journal of Wireless Information Networks**, v. 21, n. 1, p. 1–14, 2014. ISSN 15728129. DOI: 10.1007/s10776-013-0233-5.

GUO, Qiang; LI, Xin. A Safety Relay Selection Method Based on Network Coding. **International Journal of Communications, Network and System Sciences**, v. 10, n. 08, p. 167–175, 2017. DOI: 10.4236/ijcns.2017.108B018.

HAN, Chen; YANG, Yuwang; HAN, Xu. A fast network coding scheme for mobile wireless sensor networks. **International Journal of Distributed Sensor Networks**, v. 13, n. 2, 2017. DOI: 10.1177/1550147717693241.

HAN, Guangjie; ZHOU, Lina; WANG, Hao; ZHANG, Wenbo; CHAN, Sammy. A source location protection protocol based on dynamic routing in WSNs for the Social Internet of Things. **Future Generation Computer Systems**, Elsevier B.V., v. 82, p. 689–697, 2018. ISSN 0167739X. DOI: 10.1016/j.future.2017.08.044.

HAO, K.; JIN, Z. A New Deterministic Linear Network Coding and Its Application on Multicast Network. In: 5TH International Conference on Wireless Communications, Networking and Mobile Computing. [S.l.: s.n.], 2009. P. 1–4. DOI: 10.1109/WICOM.2009.5302098.

HEIDE, Janus; PEDERSEN, Morten V.; FITZEK, Frank H.P.; MÉDARD, Muriel. On code parameters and coding vector representation for practical RLNC. **IEEE International Conference on Communications**, 2011. ISSN 05361486. DOI: 10.1109/icc.2011.5963013.

HIMANSHU, Katiyar; ASHUTOSH, Rastogi; RUPALI, Agarwal. Cooperative Communication: A Review. **IETE Technical Review**, v. 5, p. 1–10, 2015.

HO, T.; MEDARD, M.; KOETTER, R.; KARGER, D. R.; EFFROS, M.; SHI, J.; LEONG, B. A Random Linear Network Coding Approach to Multicast. **IEEE Transactions on Information Theory**, v. 52, n. 10, p. 4413–4430, 2006. DOI: 10.1109/TIT.2006.881746.

HO, Tracey; LUN, Desmond. **Network Coding: An Introduction**. 1. ed. New York: Cambridge University Press, 2008. v. 9780521873, p. 184.

HOSSAIN, Mohammad Asif; NOOR, Rafidah Md; YAU, Kok Lim Alvin; AHMEDY, Ismail; ANJUM, Shaik Shabana. A Survey on Simultaneous Wireless Information and Power Transfer with Cooperative Relay and Future Challenges. **IEEE Access**, v. 7, p. 19166–19198, 2019.

HUANG, Jun; GHARAVI, Hamid; YAN, Huifang; XING, Cong Cong. Network Coding in Relay-Based Device-to-Device Communications. **IEEE Network**, v. 31, n. 4, p. 102–107, 2017. ISSN 08908044. DOI: 10.1109/MNET.2017.1700063.

IEEE COMPUTER, Society. **IEEE Std 802.15.4-2015, IEEE Standard for Low-Rate Wireless Networks**. New York, USA: IEEE, 2015. v. 2015, p. 707. ISBN 9781504408462.

IKKI, Salama S; AHMED, Mohamed H. Exact Error Probability and Channel Capacity of the Best-Relay Cooperative-Diversity Networks. **IEEE Signal Processing Letters**, v. 16, n. 12, p. 1051–1054, 2009. DOI: 10.1109/LSP.2009.2030094.

IMAM, Syed Akhtar; SINGH, Manish Kumar; SACHAN, Vibhav kumar. Energy Efficient Wireless Sensor Network using DSC-MIMO. **Journal of Engineering Technology (ISSN: 0747-9964)**, v. 6, n. 2, p. 268–277, 2017.

IQBAL, Zafar; KIM, Kiseon; LEE, Heung-No. A cooperative wireless sensor network for indoor industrial monitoring. **IEEE Transactions on Industrial Informatics**, IEEE, v. 13, n. 2, p. 482–491, 2017.

JAMAL, Tauseef; MENDES, Paulo. Relay selection approaches for wireless cooperative networks. In: 6TH IEEE International Conference on Wireless and Mobile Computing, Networking and Communications. Chengdu City, China: IEEE, Oct. 2010. P. 661–668. DOI: `10.1109/WIMOB.2010.5644850`.

KARP, R. Reducibility among combinatorial problems. In: MILLER, R.; THATCHER, J. (Eds.). **Complexity of Computer Computations**. [S.l.]: Plenum Press, 1972. P. 85–103.

KATTI, Sachin; GOLLAKOTA, Shyamnath; KATABI, Dina. Embracing wireless interference: Analog network coding. **ACM SIGCOMM Computer Communication Review**, ACM New York, NY, USA, v. 37, n. 4, p. 397–408, 2007.

KHAN, Rana Azeem M; KARL, Holger. MAC Protocols for Cooperative Diversity in Wireless LANs and Wireless Sensor Networks. **IEEE Communications Surveys and Tutorials**, v. 16, n. 1, p. 46–63, 2014.

KUORILEHTO, Mauri; HÄNNIKÄINEN, Marko; HÄMÄLÄINEN, Timo D. A survey of application distribution in wireless sensor networks. **EURASIP Journal on Wireless Communications and Networking**, Springer, v. 2005, n. 5, p. 774–788, 2005.

LAURINDO, Suelen; MORAES, Ricardo; MONTEZ, Carlos. Multi-criteria Analysis to Select Relay Nodes in the ORST Technique. In: INTERNATIONAL CONFERENCE ON AD-HOC NETWORKS AND WIRELESS. Luxembourg: [s.n.], Oct. 2019. P. 167–182.

LAURINDO, Suelen; MORAES, Ricardo; MONTEZ, Carlos; VASQUES, Francisco. Assessment of Different Algorithms to Solve the Set-Covering Problem in a Relay Selection Technique. In: 25TH INTERNATIONAL CONFERENCE ON EMERGING TECHNOLOGIES AND FACTORY AUTOMATION (ETFA). Vienna, Austria: [s.n.], Sept. 2020.

LAURINDO, Suelen; MORAES, Ricardo; MONTEZ, Carlos; VASQUES, Francisco. Combining Network Coding and Retransmission Techniques to Improve the Communication Reliability of Wireless Sensor Network. **Information**, Multidisciplinary Digital Publishing Institute, v. 12, n. 5, p. 184, 2021.

LAURINDO, Suelen; MORAES, Ricardo; NASSIFFE, Riad; MONTEZ, Carlos; VASQUES, Francisco. An Optimized Relay Selection Technique to Improve the Communication Reliability in Wireless Sensor Networks. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 18, n. 10, p. 3263, 2018.

LAURINDO, Suelen; NASSIFFE, Ríad; MORAES, Ricardo; MONTEZ, Carlos; ARAUJO, Gustavo; VALLE, Odilson. Smart: Adequate Selection of Relay Nodes to Support Cooperative Communication in WSNs. In: IEEE. 13TH IEEE International Workshop on Factory Communication Systems (WFCS). Trondheim, Noruega: [s.n.], May 2017.

LAURINDO, Suelen M; MONTEZ, Carlos; VALLE, Odilson T; MORAES, Ricardo. Um Survey de Seleção de Nodos Cooperantes em Abordagens de Comunicação Cooperativa em Redes de Sensores Sem Fio. In: XXXIV SIMPóSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUíDOS (SBRC). IV Workshop de Comunicação em Sistemas Embarcados Críticos (WOCCES). Salvador: [s.n.], 2016. P. 12–21.

LAZARESCU, Mihai T. Design of a WSN platform for long-term environmental monitoring for IoT applications. **IEEE Journal on Emerging and Selected Topics in Circuits and Systems**, v. 3, n. 1, p. 45–54, 2013. ISSN 21563357.

LI, Xiong; NIU, Jianwei; KUMARI, Saru; WU, Fan; SANGAIAH, Arun Kumar; CHOO, Kim Kwang Raymond. A three-factor anonymous authentication scheme for wireless sensor networks in internet of things environments. **Journal of Network and Computer Applications**, Elsevier Ltd, v. 103, May 2017, p. 194–204, 2018. ISSN 10958592. DOI: `10.1016/j.jnca.2017.07.001`.

LI, Ye; ZHANG, Shibing; WANG, Jue; JI, Xiaodong; WU, Huangnan; BAO, Zhihua. A Low-Complexity Coded Transmission Scheme over Finite-Buffer Relay Links. **IEEE Transactions on Communications**, v. 66, n. 7, p. 2873–2887, 2018. ISSN 00906778. DOI: `10.1109/TCOMM.2018.2809627`.

LI, Yubo; YIN, Qinye; XU, Wei; WANG, Hui-ming. On the Design of Relay Selection Strategies in Regenerative Cooperative Networks with Outdated CSI. **IEEE Transactions on Wireless Communications**, v. 10, n. 9, p. 3086–3097, 2011.

LIANG, Xuedong; BALASINGHAM, Ilangko; LEUNG, Victor C M. Cooperative Communications with Relay Selection for QoS Provisioning in Wireless Sensor

Networks. In: GLOBECOM 2009 - 2009 IEEE Global Telecommunications Conference. [S.l.]: IEEE, Nov. 2009. P. 1–8. DOI: 10.1109/GLOCOM.2009.5425437.

LIU, Lingya; HUA, Cunqing; CHEN, Cailian; GUAN, Xinping. Relay Selection for Three-Stage Relaying Scheme in Clustered Wireless Networks. **IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY**, v. 64, n. 6, p. 2398–2408, 2015.

LIU, Xingcheng; GONG, Xinren; ZHENG, Yongzhao. Reliable cooperative communications based on random network coding in multi-hop relay WSNs. **IEEE sensors journal**, IEEE, v. 14, n. 8, p. 2514–2523, 2014.

LU, X.; CHENG, W.; HE, Q.; XIE, X. Cooperative communication based regular topology to achieve coverage and K-connectivity for WSNs. In: 13TH IEEE Conference on Industrial Electronics and Applications (ICIEA). Wuhan, China: [s.n.], May 2018. P. 2514–2518.

MARCHENKO, Nikolaj; ANDRE, Torsten; MASOOD, Wasif; BETTSTETTER, Christian. An Experimental Study of Selective Cooperative Relaying in Industrial Wireless Sensor Networks. **IEEE Transactions on Industrial Informatics**, v. 10, n. 3, p. 1806–1816, 2014.

MATIAS-PEREIRA, José. **Manual de Metodologia da Pesquisa Científica**. 3. ed. São Paulo: Atlas, 2012. P. 196. ISBN 978-85-224-6975-8.

MEI, Z.; LU, Y. Performance Analysis of Several Relay Node Selection Schemes for Cooperative Communications. In: 11TH International Conference on Communication Software and Networks (ICCSN). Chongqing, China: [s.n.], June 2019. P. 41–45.

MIGABO, M. E.; OLWAL, T. O.; DJOUANI, K.; KURIEN, A. M. Cooperative and Adaptive Network Coding for Gradient Based Routing in Wireless Sensor Networks with Multiple Sinks. **Journal of Computer Networks and Communications**, 2017. ISSN 2090715X. DOI: 10.1155/2017/5301462.

MORRISON, David R; JACOBSON, Sheldon H; SAUPPE, Jason J; SEWELL, Edward C. Branch-and-Bound Algorithms: A Survey of Recent Advances in Searching, Branching, and Pruning. **Discrete Optimization**, Elsevier, v. 19, p. 79–102, 2016.

NHON, Tran; KIM, Dong seong. Relay Selection Scheme for Hierarchical Wireless Sensor Networks. **International Journal of Control and Automation**, v. 7, n. 3, p. 147–160, 2014.

NOSRATINIA, Aria; HUNTER, Todd E; HEDAYAT, Ahmadreza. Cooperative Communication in Wireless Networks. **IEEE Communications Magazine**, IEEE, v. 42, October, p. 74–80, 2004.

OSTOVARI, Pouya; WU, Jie; KHREISHAH, Abdallah. Network coding techniques for wireless and sensor networks. In: THE art of wireless sensor networks. [S.l.]: Springer, 2014. P. 129–162. DOI: `https://doi.org/10.1007/978-3-642-40009-4`.

OUYANG, Fengchen; GE, Jianhua; GONG, Fengkui; HOU, Jun. Random access based blind relay selection in large-scale relay networks. **IEEE Communications Letters**, v. 19, n. 2, p. 255–258, 2015. ISSN 10897798. DOI: `10.1109/LCOMM.2014.2379280`.

PHAM, Tung Linh; KIM, Dong Seong. Efficient Forwarding Protocol for Dual-Hop Relaying Wireless Networks. **Wireless Personal Communications**, Springer US, v. 89, n. 1, p. 165–180, 2016. ISSN 1572834X. DOI: `10.1007/s11277-016-3258-0`.

PHAN, Khoa T; NGUYEN, Duy HN; LE-NGOC, Tho. Joint power allocation and relay selection in cooperative networks. In: IEEE. GLOBAL Telecommunications Conference (GLOBECOM). [S.l.: s.n.], 2009. P. 1–5.

PILLONI, Virginia. How Data will Transform Industrial Processes: Crowdsensing, Crowdsourcing and Big Data as Pillars of Industry 4.0. **Future Internet**, Multidisciplinary Digital Publishing Institute, v. 10, n. 3, p. 24, 2018.

PINTO, A.R.; MONTEZ, C.; ARAÚJO, G.; VASQUES, F.; PORTUGAL, P. An approach to implement data fusion techniques in wireless sensor networks using genetic machine learning algorithms. **Information Fusion**, v. 15, p. 90–101, 2014.

POLAVARAPU, Sushma Chowdary; PANDA, Sampad Kumar. A Survey on Industrial Applications using MEMS and WSN. In: IEEE. 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC). [S.l.: s.n.], 2020. P. 982–986.

RAJANNA, B.; CHANDRASHRKAR, P.; KISHAN RAO, K. Optimized power allocation and cooperative relay selection in communication system. **Journal of Advanced Research in Dynamical and Control Systems**, v. 9, n. 3, p. 14–18, 2017. ISSN 1943023X.

RODRIGUES, Leonardo M; MONTEZ, Carlos; BUDKE, Gerson; VASQUES, Francisco; PORTUGAL, Paulo. Estimating the lifetime of wireless sensor network nodes through the use of embedded analytical battery models. **Journal of Sensor and Actuator Networks**, Multidisciplinary Digital Publishing Institute, v. 6, n. 2, p. 8, 2017a.

RODRIGUES, Leonardo M; MONTEZ, Carlos; MORAES, Ricardo; PORTUGAL, Paulo; VASQUES, Francisco. A temperature-dependent battery model for wireless sensor networks. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 17, n. 2, p. 422, 2017b.

SADIKU, Matthew NO; EZE, Kelechi G; MUSA, Sarhan M. Wireless Sensor Networks for Healthcare. **Journal of Scientific and Engineering Research**, v. 5, n. 7, p. 210–213, 2018.

SAHOO, Prasan Kumar; PATTANAIK, Sudhir Ranjan; WU, Shih-Lin. Design and analysis of a low latency deterministic network MAC for wireless sensor networks. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 17, n. 10, p. 2185, 2017.

SEHAT, Hadi; PAHLEVANI, Peyman. On the Probability of Partial Decoding in Sparse Network Coding. **arXiv preprint arXiv:1907.12051**, 2019.

SELVARAJ, M. D.; MALLIK, Ranjan K. Performance analysis of a multi-relay cooperative diversity network with decode and forward relaying. **Proceedings of 16th National Conference on Communications, NCC 2010**, 2010. DOI: 10.1109/NCC.2010.5430227.

SENANAYAKE, Rajitha; ATAPATTU, Saman; EVANS, Jamie S; SMITH, Peter J. Decentralized relay selection in multi-user multihop decode-and-forward relay networks. **IEEE Transactions on Wireless Communications**, IEEE, v. 17, n. 5, p. 3313–3326, 2018.

SHAHEEN, Ameen; SLEIT, Azzam. Comparing Between Different Approaches to Solve the 0/1 Knapsack Problem. **International Journal of Computer Science and**

Network Security (IJCSNS), International Journal of Computer Science and Network Security, v. 16, n. 7, p. 1–10, 2016.

SOLVE, lp. **lp solve Library**. [S.l.: s.n.], 2007. http://lpsolve.sourceforge.net/. 05 Apr 2018.

SONKAR, Poonam; SARKARI, Vertika; TIWARI, BB; SAUBAGYA, Nitant. A Review on Cooperative Communication with Relay. **Imperial Journal of Interdisciplinary Research**, v. 2, n. 9, 2016.

SRINIVASAN, Kannan; LEVIS, Philip. RSSI is Under Appreciated. In: PROC of the Third Workshop on Embedded Networked Sensors (EmNets). [S.l.]: Cambridge, MA, USA, 2006. P. 239–242.

SRIVASTAVA, Ankita; MISHRA, Pramod Kumar. A Survey on WSN Issues with its Heuristics and Meta-Heuristics Solutions. **Wireless Personal Communications**, Springer, p. 1–70, 2021.

SU, Weifeng; SADEK, Ahmed K.; RAY LIU, K. J. Cooperative communication protocols in wireless networks: Performance analysis and optimum power allocation. **Wireless Personal Communications**, v. 44, n. 2, p. 181–217, 2008. ISSN 09296212. DOI: 10.1007/s11277-007-9359-z.

SU, Yuhan; LU, Xiaozhen; ZHAO, Yifeng; HUANG, Lianfen; DU, Xiaojiang. Cooperative Communications with Relay Selection Based on Deep Reinforcement Learning in Wireless Sensor Networks. **IEEE Sensors Journal**, IEEE, v. 19, n. 20, p. 9561–9569, 2019.

SUDHA, G; PRAKASH, R; GANESH, A Balaji; GIRISH, Siva V. Network Coding based Real Time Wireless Sensor Network for Environmental Monitoring. **International Conference on Wireless Communications, Signal Processing and Networking**, p. 1269–1272, 2016.

SUN, Li; ZHANG, Taiyi; LU, Long; NIU, Hao. Cooperative communications with relay selection in wireless sensor networks. **IEEE Transactions on Consumer Electronics**, v. 55, n. 2, p. 513–517, 2009. ISSN 00983063. DOI: 10.1109/TCE.2009.5174415.

TOMIĆ, Ivana; MCCANN, Julie A. A survey of potential security issues in existing wireless sensor network protocols. **IEEE Internet of Things Journal**, IEEE, v. 4, n. 6, p. 1910–1923, 2017.

TOZONI-REIS, Marília Freitas de Campos. **Metodologia da Pesquisa**. 2. ed. Curitiba: IESDE Brasil S.A, 2009. P. 136.

TRIPATHI, Y.; PRAKASH, A.; TRIPATHI, R. An Energy Balanced Load Aware Relay Selection in Cooperative Routing for Wireless Sensor Network. In: 9TH Annual Information Technology, Electromechanical Engineering and Microelectronics Conference (IEMECON). Jaipur, Índia: [s.n.], 2019. P. 223–231.

TUAN, Nguyen Trong; KIM, Dong-Seong; LEE, Jae-Min. On the Performance of Cooperative Transmission Schemes in Industrial Wireless Sensor Networks. **IEEE Transactions on Industrial Informatics**, v. 3203, n. 100, p. 1–1, 2018. ISSN 1551-3203. DOI: 10.1109/TII.2018.2846671.

VALLE, O. T.; MONTEZ, C.; ARAÚJO, G. M.; MORAES, R.; VASQUES, F. A WSN data retransmission mechanism based on network coding and cooperative relayers. In: 11TH IEEE World Conference on Factory Communication Systems (WFCS). Palma de Mallorca, Spain: [s.n.], May 2015. P. 1–4.

VALLE, Odilson T; BUDKE, Gerson; MONTEZ, Carlos; MORAES, Ricardo; VASQUES, Francisco. Experimental assessment of LNC-based cooperative communication schemes using commercial off-the-shelf wireless sensor network nodes. **International Journal of Communication Systems**, Wiley Online Library, v. 31, n. 7, e3508, 2018.

VALLE, Odilson T; MILACK, André V; MONTEZ, Carlos; PORTUGAL, Paulo; VASQUES, Francisco. Experimental evaluation of multiple retransmission schemes in IEEE 802.15.4 wireless sensor networks. In: 9TH IEEE International Workshop on Factory Communication Systems (WFCS). [S.l.: s.n.], 2012. P. 201–210.

VALLE, Odilson T; MONTEZ, Carlos; ARAUJO, Gustavo de; VASQUES, Francisco; MORAES, Ricardo. NetCoDer: A Retransmission Mechanism for WSNs Based on Cooperative Relays and Network Coding. **Sensors**, v. 16, n. 6, 2016.

VASEGHI, Behrouz; POURMINA, Mohammad Ali; MOBAYEN, Saleh. Secure communication in wireless sensor networks based on chaos synchronization using adaptive sliding mode control. **Nonlinear Dynamics**, Springer Netherlands, v. 89, n. 3, p. 1689–1704, 2017. ISSN 1573269X. DOI: `10.1007/s11071-017-3543-9`.

WANG, Chin Liang; SYUE, Syue Ju. An Efficient Relay Selection Protocol for Cooperative Wireless Sensor Networks. In: IEEE Wireless Communications and Networking Conference. Budapest, Hungary: IEEE, 2009. P. 1–5.

WANG, Heng; YANG, Shizhong; LIN, Jinzhao; ZHONG, Yuanhong. Single Relay Selection With Feedback and Power Allocation in Multisource Multidestination Cooperative Networks. **IEEE Signal Processing Letters**, v. 17, n. 12, p. 997–1000, 2010.

WANG, Jun; WU, Qilin. Relay selection for maximizing the number of successive transmission in cooperative networks. **International Journal of Parallel, Emergent and Distributed Systems**, Taylor & Francis, v. 32, n. 6, p. 632–646, 2017.

WILLIG, Andreas; MATUSOVSKY, Yakir; KIND, Adriel. Retransmission scheduling in 802.15. 4e LLDN-a reinforcement learning approach with relayers. In: 26TH International Telecommunication Networks and Applications Conference (ITNAC). [S.l.: s.n.], 2016. P. 63–69.

WILLIG, Andreas; UHLEMANN, Elisabeth. On relaying for wireless industrial communications: Is careful placement of relayers strictly necessary? In: 9TH IEEE International Workshop on Factory Communication Systems (WFCS). [S.l.: s.n.], 2012. P. 191–200.

WOLSEY, Laurence A. **Integer programming**. 1. ed. New York: Wiley, 1998. ISBN 0471283665.

WU, Huangnan; LI, Ye; HU, Yingdong; TANG, Bin; BAO, Zhihua. On Optimizing Effective Rate for Random Linear Network Coding over Burst-Erasure Relay Links. **IEEE Wireless Communications Letters**, v. 8, n. 2, p. 588–591, 2019. ISSN 21622345. DOI: `10.1109/LWC.2018.2881157`.

XU, Changqiao; WANG, Peng; XIONG, Chunshan; WEI, Xinpeng; MUNTEAN, Gabriel Miro. Pipeline network coding-based multipath data transfer in heterogeneous wireless networks. **IEEE Transactions on Broadcasting**, v. 63, n. 2, p. 376–390, 2017. ISSN 00189316. DOI: `10.1109/TBC.2016.2590819`.

XU, Kenan; WANG, Quanhong; HASSANEIN, Hossam; TAKAHARA, Glen. Optimal wireless sensor networks (WSNs) deployment: minimum cost with lifetime constraint. In: IEEE. WIRELESS and Mobile Computing, Networking and Communications, 2005.(WiMob'2005), IEEE International Conference on. [S.l.: s.n.], 2005. P. 454–461.

YANG, K.; ZHU, S.; DAN, Z.; TANG, X.; WU, X.; OUYANG, J. Relay Selection for Wireless Cooperative Networks using Adaptive Q-learning Approach. In: 2019 Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC). Taiyuan, China.: [s.n.], July 2019. P. 1–5.

YANG, Qi; NOFSINGER, Adam; MCPEEK, Jamie; PHINNEY, Joel; KNUESEL, Ryan. A Complete Solution to the Set Covering Problem. In: INTERNATIONAL Conference on Scientific Computing (CSC). Las Vegas, Nevada, USA: [s.n.], July 2015. P. 36–41.

YETGIN, Halil; CHEUNG, Kent Tsz Kan; EL-HAJJAR, Mohammed; HANZO, Lajos Hanzo. A survey of network lifetime maximization techniques in wireless sensor networks. **IEEE Communications Surveys & Tutorials**, IEEE, v. 19, n. 2, p. 828–854, 2017.

YICK, Jennifer; MUKHERJEE, Biswanath; GHOSAL, Dipak. Wireless sensor network survey. **Computer Networks**, v. 52, n. 12, p. 2292–2330, 2008.

YUE, Jing; GIRS, Svetlana; AKERBERG, Johan; BJORKMAN, Mats; LIN, Zihuai. Network coding for reliable downlink transmissions in industrial wireless sensor and actuator networks. **Proceedings of the International Symposium on Wireless Communication Systems**, IEEE, 2016-Octob, p. 560–565, 2016. ISSN 21540225. DOI: 10.1109/ISWCS.2016.7600967.

ZAREI, Amir; PAHLEVANI, Peyman; DAVOODI, Mansoor. On the partial decoding delay of sparse network coding. **IEEE Communications Letters**, IEEE, v. 22, n. 8, p. 1668–1671, 2018.

EZ-ZAZI, Imad; ARIOUA, Mounir; EL OUALKADI, Ahmed; LORENZ, Pascal. A hybrid adaptive coding and decoding scheme for multi-hop wireless sensor networks. **Wireless Personal Communications**, Springer, v. 94, n. 4, p. 3017–3033, 2017.

ZHANG, Caiming; CHEN, Yong. A review of research relevant to the emerging industry trends: Industry 4.0, IoT, blockchain, and business analytics. **Journal of Industrial Integration and Management**, World Scientific, v. 5, n. 01, p. 165–180, 2020.

ZHANG, Jian; TANG, Jian; WANG, Feng. Cooperative relay selection for load balancing with mobility in hierarchical WSNs: A multi-armed bandit approach. **IEEE Access**, v. 8, p. 18110–18122, 2020.

ZHANG, Wei; LIN, Bin; YU, Cunqian; WANG, Ying; GAO, Changxuan. Dimensioning and Layout Planning of Integrated VANET-WSN for Intelligent Transportation. In: LIANG, Qilian; LIU, Xin; NA, Zhenyu; WANG, Wei; MU, Jiasong; ZHANG, Baoju (Eds.). **Communications, Signal Processing, and Systems**. Singapore: Springer Singapore, 2019. P. 614–622.

ZHANG, Yu; ZHANG, Bing; ZHANG, Shi. A lifetime maximization relay selection scheme in wireless body area networks. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 17, n. 6, p. 1267, 2017.

ZHAO, Gang. Wireless Sensor Networks for Industrial Process Monitoring and Control: A Survey. **Network Protocols & Algorithms**, v. 3, n. 1, p. 46–63, 2011.

ZHU, Kaiyan; WANG, Fei; LI, Songsong; JIANG, Fengjiao; CAO, Lijie. Relay Selection for Cooperative Relaying in Wireless Energy Harvesting Networks. In: 5, p. 9.

APPENDIX A – LIST OF PUBLICATIONS

A list of publications produced during this PhD is presented below.

A.1 JOURNAL PUBLICATIONS

- LAURINDO, Suelen; MORAES, Ricardo; MONTEZ, Carlos; VASQUES, Francisco.Combining Network Coding and Retransmission Techniques to Improve theCommunication Reliability of Wireless Sensor Network. **Information** (CiteScore 2020: 3.0), Multidisciplinary Digital Publishing Institute, v. 12, n. 5, p. 184, 2021.

- LAURINDO, Suelen; MORAES, Ricardo; NASSIFFE, Riad; MONTEZ, Carlos; VASQUES, Francisco. An Optimized Relay Selection Technique to Improve the Communication Reliability in Wireless Sensor Networks. **Sensors** (JCR 2020: 3.576), Multidisciplinary Digital Publishing Institute, v. 18, n. 10, p. 3263, 2018.

A.2 CONFERENCE PUBLICATIONS

- LAURINDO, Suelen; MORAES, Ricardo; MONTEZ, Carlos. Multi-criteria Analysis toSelect Relay Nodes in the ORST Technique. In: INTERNATIONAL CONFERENCE ON AD-HOC NETWORKS AND WIRELESS. Luxembourg, Oct. 2019. P. 167–182.

- LAURINDO, Suelen; MORAES, Ricardo; MONTEZ, Carlos. Cooperative Communication Mechanisms Applied to Wireless Sensor Network. In: Doctoral Conference on Computing, Electrical and Industrial Systems. Springer, Cham, 2020. p. 121-128.

- LAURINDO, Suelen; MORAES, Ricardo; MONTEZ, Carlos; VASQUES, Francisco.Assessment of Different Algorithms to Solve the Set-Covering Problem in a RelaySelection Technique. In: 25TH INTERNATIONAL CONFERENCE ON EMERGINGTECHNOLOGIES AND FACTORY AUTOMATION (ETFA). Vienna, Austria, Sept.2020.