UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES

Electronics and Computer Science

Automated Negotiation for Opportunistic Direct Cooperation Between Neighbouring Wireless Sensor Networks

by

Andre Ortega Alban

Thesis for the degree of Doctor of Philosophy

May 2020

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES
Electronics and Computer Science

Doctor of Philosophy

AUTOMATED NEGOTIATION FOR OPPORTUNISTIC DIRECT COOPERATION BETWEEN NEIGHBOURING WIRELESS SENSOR NETWORKS

by Andre Ortega Alban

As the Internet of Things grows, multiple wireless sensor networks (WSNs) are likely to coexist. From wearable health monitors to smart cities, WSNs will play an increasingly key role in most scenarios. In many of these applications, sensor nodes are likely to be battery-powered and hence limited in energy supply. Energy harvesting technologies have gained widespread attention to increase node lifetime. However, these exhibit spatio-temporal variations and expose a discontinuous power supply.

Visioning a future with cooperative networks, this work proposes to extend network performance optimisation to an inter-domain approach by opportunistic cooperation among WSNs that share a common area. Since WSNs are highly heterogeneous and self-interested, cooperation is not guaranteed. The cooperation problem has been addressed using a game-theoretic approach. However, assumptions as full rationality or complete knowledge are not justified in this domain. Instead, this work utilises multi-agent design methods to provide a new methodology on negotiation-based cooperation that enables suitable agreements on energy sharing.

With the aim to optimise a network's power management using the suggested approach, a node's own efficiency is computed. Thus, a self-organising algorithm capable of making optimal use of harvested energy is proposed. This power management technique is tested during every simulation presented. Such an algorithm enables self-organised nodes that can anticipate insufficient energy allocation schemes and identify the opportunity to start an energy negotiation (OEN).

Experiments show the accomplishment of energy-neutrality when networks find energy flow agreements and adopt conciliatory behaviours. The effect on the power consumption and latency of establishing OEN is also quantified and proved to be insignificant (<0.01 J, <0.1 s).

A novel partner selection method based on multi-armed bandits is also introduced to facilitate the estimation of successful negotiation agreements. The proposed model allows networks to maximise their energy allocation in the long run, while adapting to a highly dynamic and uncertain environment. The viability of the approach is measured through simulation, and results show that networks may improve their energy allocation by over 40% in the most challenging scenario considered in this thesis.

Contents

List of Figures

List of Tables

Abbreviations

AI	Artificial Intelligence
BN	Boundary Node
BSN	Body Sensor Network
CBSNs	Collaborative Body Sensor Networks
CCH	Common CHannel
DCH	Data CHannel
EHWSN	Energy Harvesting Wireless Sensor Network
FPL-UE	Follow the Perturbed Leader with Uniform Exploration
GR	Geometric Resampling
GT	Game Theory
IoT	Internet of Things
LP	Linear Programming
LSPI	Least-Squares Policy Iteration
MABs	Multi-Armed Bandits
MAB/EM	Multi-Armed Bandit based Energy Management
MAC	Medium Access Control
MAS	Multi-Agent System
MDP	Markov Decision Process
NBS	Nash Bargaining Solution
NE	Nash Equilibrium
ODI	Opportunistic Direct Interconnection
OEN	Opportunistic Energy Negotiation
OET	Opportunistic Energy Trading
OI-MAC	Opportunistic Inter-connection MAC
PVGIS	PhotoVoltaic Geographical Information System
RL	Reinforcement Learning
U	Uniform Random Distribution
VCB	Virtual Cooperation Bond
WSN	Wireless Sensor Network

Nomenclature

N	Network
I_i	Set of sensor nodes in network i
T	Period
t	t-th slot
n	Total number of time slots
L	Duration of a time slot
u	Utility
d	Disagreement utility
o^{NBS}	Offer computed with NBS
S	Feasible set of agreements
i_j	Agent of node j in network i
$B_{i,j}(t)$	Battery level of agent i_j at the beginning of slot t
$B_{i,j}^{max}$	Maximum battery level of agent i_j
e	Battery efficiency
$d(t)$	Amount of energy that is discharged from the battery at time slot t
$c(t)$	Amount of energy that is charged to the battery at time slot t
$E_{i,j}^{hrv}(t)$	Amount of energy harvested by agent i_j at slot t
G_b	Solar radiation
v	Wind speed
ρ	Air density
$E_{i,j}^{alloc}(t)$	Energy allocation for agent i_j at slot t
A	Area
f	Efficiency of solar panel or wind turbine
p	Effect of ambient perturbation
$w_{i,j}(t)$	Surplus of energy of agent i_j
V	Voltage supplied to agent i_j
D	Duty cycle of agent i_j
I^{active}	Current consumed in active mode by agent i_j
I^{sleep}	Current consumed in sleep mode by agent i_j
$E_{k,i}^{c}(t)$	Total energy consumed by agent i_j at slot t
o	Energy flow offer
R	Set of rounds

r_{max}	Maximum number of rounds
$min_{i,j}$	Minimum energy allocation for agent i_j
α^r	Time dependent function
β	Concession shape
$\Omega_{i,j}$	1-hop local neighbourhood of agent i_j
$e_i(i\_u, l\_v)$	Edge between agents u and v from networks i and l
K	Set of arms or actions
Tr	Set of trials or opportunistic negotiation encounters
tr	tr-th trial
$a(tr)$	Action of an agent at trial tr
$r_{i,j}$	Reward function of agent i_j
R_{Tr}	Weak regret over Tr
ε	Exploration factor in ε-Greedy Action Selection Strategy
$\hat{r}_k$	Estimated reward of action k
$pulls_k$	Number of times action k has been executed in ε-Greedy
$rewards_k$	Cumulative reward of action k in ε-Greedy
λ	Exploration factor in FPL-UE policy
η	Mean parameter for exponential distribution in FPL-UE
M	Maximum number of samples in FPL-UE
K_val_k	Reciprocal of probability of action k in FPL-UE
z_k	Exponential random number in FPL-UE
γ	Exploration factor in EXP3 policy
w_k	Weight value of action k in EXP3
p_k	Probability of action k in EXP3

Declaration of Authorship

I, Andre Ortega Alban , declare that the thesis entitled *Automated Negotiation for Opportunistic Direct Cooperation Between Neighbouring Wireless Sensor Networks* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as: [1]

Signed:...

Date:..

Acknowledgements

Thanks to God for giving me the strength and blessing me along the way. To my parents and family for their support not only with my studies but throughout my life. Monica, Alberto, Horten, Manuel, Rosa, Keshia, Ashley and Morita thanks for being there no matter how far we are. To my husband, Titov, for accompanying me on this journey and always being by my side when I needed him the most. For being the support and inspiration that encouraged me during this time.

My sincerest thanks to my supervisors for all their guidance and support throughout my Ph.D. To Dr Geoff, for giving me the opportunity of doing this research. His trust and mentoring made this work possible and I will be always grateful. To Dr Gopal, whose expertise in multi-agent systems and excellent feedback helped me to achieve the contributions of my work. Special thanks to Dr Long Tran-Thanh for sharing his time during several sessions and guiding me throughout the multi-armed bandits topic.

I would like to acknowledge the financial support from Secretaría de Educación Superior, Ciencia, Tecnología e Innovación (SENESCYT) and Escuela Superior Politécnica del Litoral (ESPOL) under Ph.D. studies 2016, allowing this research to be conducted. My deepest gratitude to Dr Cristina Abad, Dr Katherine Chiluiza, Dr Enrique Pelaez, Dr Xavier Ochoa for their continuos support and guidance.

Thanks: to my family in Southampton, Addy and Chris, who offered me their help since day one; to my friends from Ecuador, CPS and AIC, for sharing this experience with me one way or another and give me the push to get back on track every time I felt the pressure.

To my family, for their love and support

Chapter 1

Introduction

A wireless sensor network (WSN) is formed by multiple sensor devices equipped with limited memory, processing, power and communication capabilities. The typical application of WSNs involves the collection of data from the environment, data processing, and routing of aggregated data to a central node (or sink). For these features, WSNs are presently used in a wide range of applications that include environmental monitoring and target tracking.

The ubiquitous nature of these networks is a key component in the Internet of Things (IoT), where sensor devices are deployed to support the development of smart, dynamic and context-aware IoT applications. With the increasing popularity of IoT, the idea of multiple overlapping WSNs constructed in the same area becomes more feasible.
An example of sensor network applications deployed within small geographic vicinity is shown in Figure 1.1.

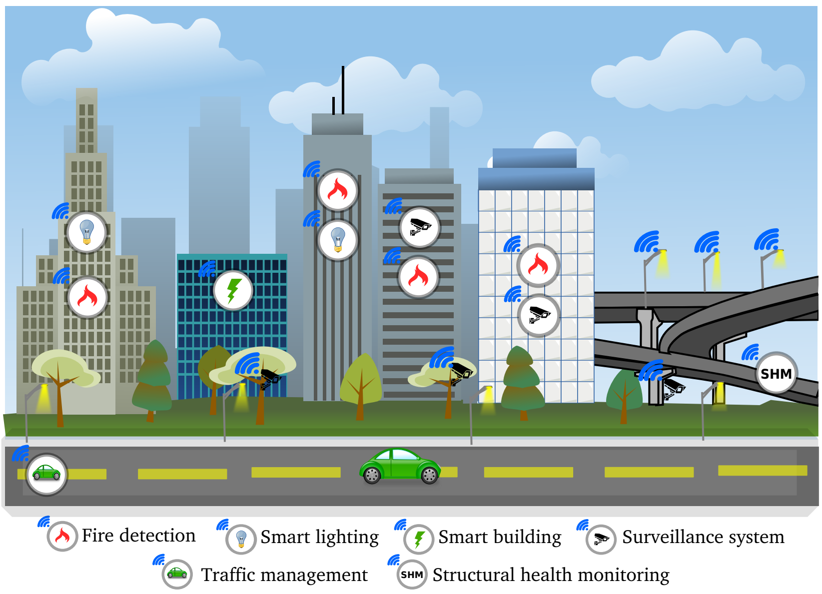

Figure 1.1: An IoT ecosystem with multiple co-located WSNs.

The constrained nature of sensor nodes and the realization of IoT encourage research on cooperation among co-located [1] devices. The cooperation among nearby sensor nodes is important to exploit their inherent diversity. The gains of such cooperation will depend on the goal or utility each device tries to maximise. Accordingly, the coordination of resources, as energy and data, for cooperation between multi-domain sensor networks could lead to many benefits, e.g., longer lifetime or higher data coverage.

In more detail, work in [2] presents the advantages of data sharing among WSNs. The operation of a WSN measuring relative humidity is optimised using the data from an external WSN that measures temperature. The mechanism may reduce the energy consumption of the network and extend its lifetime by making predictions and adapting its operation.

The exchange of information can also be useful in expanding the data coverage of a WSN. Specifically, Zia et al. [3] study the importance of data sharing between WSNs in co-located farm fields. The goal in this domain is to have an effective water quality management at a catchment scale (see Figure 1.2). Data sharing is expanded on in Section 2.2.1.

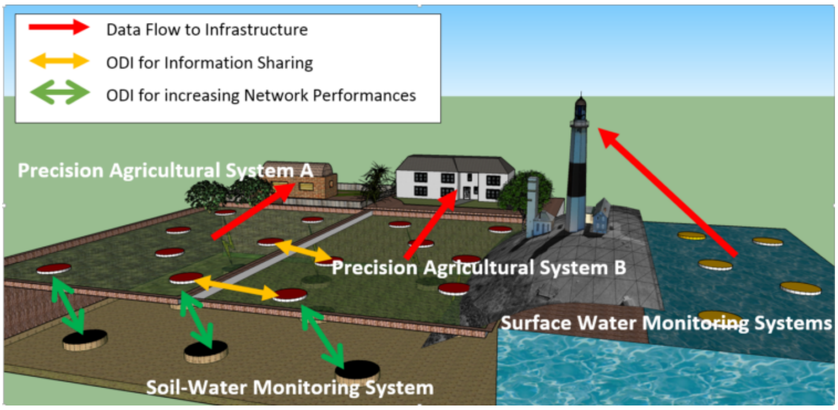

Figure 1.2: Collaborative catchment-scale monitoring of precision agricultural and water quality control systems (taken from [4]).

In [5], several use cases are described to illustrate how cooperation among sensor networks is useful for their performance optimization. The authors in [6] describe a paradigm referred to as “Symbiotic networks” for supporting the integration of different wireless networks from their design. Using symbiotic networks, they suggest many different forms of cooperation, from sharing a resource as information, or nodes for processing and routing purposes, to offering networking services to each other. For instance, work in [7] introduces a symbiotic software platform, which enables the combination of data from distinct networks to construct service composition relying on semantic web technology.

In effect, the idea of cooperation among multiple WSNs has been previously studied by several works [8–13] including game-theoretic models to investigate the impact of cooperation, cooperation strategies for packet forwarding to prolong the networks’ lifetime, resource trading

[1]The terms co-located, overlapped and neighbour are used interchangeably to denote nodes with overlapping radio range.

between neighbouring WSNs to control energy transfers, and a negotiation-based approach to measure the benefits of different networks sharing network services. By exploiting cooperation, networks can accomplish multiple benefits in terms of data-sharing or energy-sharing.

Since energy is scarce in sensor nodes, a clearly visible challenge of WSNs is the energy consumption problem. Accordingly, the reduction of power consumption through energy sharing by cooperative packet forwarding has been the most important incentive to enable cross-network cooperation [9, 11–17].

Energy harvesting technologies have attracted a lot of attention to mitigate the energy scarcity issue and become the solution for future ubiquitous smart environments. However, energy harvesting wireless sensor networks (EHWSNs) are conditioned to spatio-temporal variations of energy availability providing intermittent power supply (e.g. negligible energy at nights in the case of solar technology). An EHWSN is a WSN, but unlike battery-powered WSN, the harvested energy is used as the first source of energy for the system with the main objective of achieving an endlessly long system lifetime. This mode of operation is called energy-neutral operation: a harvesting node achieves it if the energy supply during a harvesting period is sufficient to support the amount consumed during the same time [18].

Environmental factors such as weather conditions, physical obstructions or limited availability of the ambient sources such as light or wind can have a significant impact on energy availability. Due to the uncertainty of these sources, relying solely on them can provide only enough energy to power the sensors sporadically and not continuously. Therefore, the need to develop alternative energy management techniques arises in order to achieve efficient energy allocation in rechargeable sensor networks.

By pursuing the cooperation approach, energy transfer across network boundaries can help in the goal of energy-neutrality. In [8], Teng et al. propose a novel power management strategy based on Opportunistic Energy Trading (OET) between two networks. The spatial variation of harvestable energy allows energy-neutral systems to leverage an area wider than one single network domain (see Figure 1.3). However, in this scenario, the improvement of performance is exclusive for one network only. The proposal ignores the different energy profiles of the networks and the conflicts that arise due to their heterogeneity. The work also assumes altruistic behaviour for reasons of clarity and simplicity. More details on energy sharing are available in Section 2.2.2.

This thesis adopts the idea of exploiting an area wider than the boundaries of the network, since it may represent a potential solution to use the energy sources more efficiently through opportunistic cooperation. By opportunistic, this research refers to interactions between networks that occur without having prior knowledge about the resources and characteristics of the other part. Besides, the heterogeneous nature of WSNs in their power consumption, batteries and energy harvesting profiles can be useful to complement the energy requirements of all networks involved in opportunistic cooperation.

Generally, the amount of solar energy harvested increases during morning time slots, declines during afternoon slots, and is negligible during night slots. On the other hand, for the wind source, the replenishment profile traces are very irregular. Let's assume it is windy at night and the afternoon has some periods of wind absence while it is calm at mornings. Wind energy can be sufficient to power up sensor nodes and store some energy in their batteries. Consequently, the combination of sun and wind energy may ensure every network node to exploit its source to the fullest instead of marking the excess harvested energy as waste or satisfy only a subset of load. This is the main idea of this research, to find a suitable framework so networks can share harvested energy at some points in time in return for energy at other periods of time.

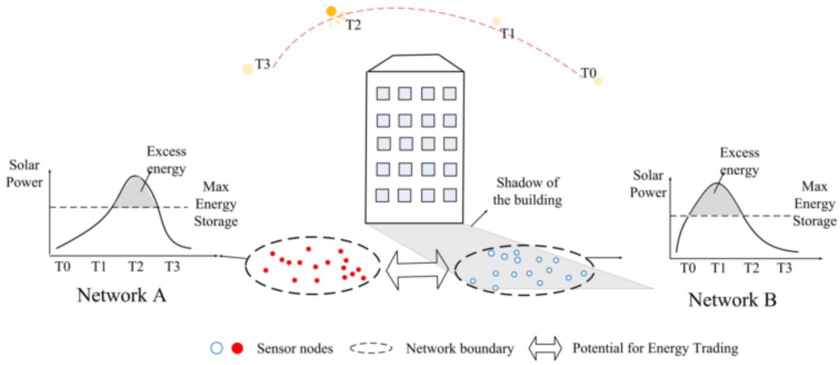

Figure 1.3: The motivation for Opportunistic Energy Trading (OET): existing power management algorithms are limited to managing resources within their network boundary, whereas OET enables the effective transfer of energy across multiple networks (taken from [8]).

Even if multiple WSNs share the same area, the authorities associated with the networks and their distinguished application-specific designs may make this sharing not straightforward. The co-located networks can be considered independent and self-interested with different optimisation goals, which suggests that cooperation without a mutually acceptable agreement might not be fair or guaranteed. Thus, networks need to decide first whether cooperation is beneficial or not.

Accordingly, this thesis presents an overview of several different methods to study both the cooperation problem and current solutions to ensure energy neutral operation. In this work, the aim is to investigate how co-located and distinct WSNs can coordinate to jointly cooperate and optimise the use of their harvested energy. If so, through what method can these networks instead of operating in isolation, reach an agreement when each network authority is only interested in its own utility? This is the question this thesis aims to address here. The research should be oriented to define an opportunistic energy cooperation scheme to consider multiple nodes and multiple networks, keeping the goal of achieving efficient energy management in the long run.

1.1 Research Justification

Energy management schemes are designed to provide networks with efficient use of harvested energy. Previous efforts have concentrated on two different perspectives on how to ensure energy-neutral sensing systems: optimal energy allocation algorithms and adaptive parameter strategies. Various algorithms have been proposed to search for optimal allocation of gathered energy [19–24]. These proposals have the important limitation that the energy harvesting considered for allocation is limited to the scope of one network domain. In this direction, even when additional energy can be harvested from an ambient energy source, the energy collected remains a limited resource due to temporal and spatial variations [25]. Therefore, the energy output is virtually insufficient during intermittent supply. Moreover, if the entire network is unable to harvest energy (due to ambient conditions or obstacles in the environment), none of these solutions is enough. The energy allocation schemes of some approaches act over infinite battery storages or ideal energy buffers [21, 23, 26–28]. These assumptions are unrealistic in practical scenarios and are not required in the model described in this thesis.

A reasonable notion of limited energy buffers is presented by several allocation algorithms [22, 29–31]. However, the models define complex optimisation problems involving Markov decision processes (MDPs) or dynamic programming. These methods include increased computational costs and significant running times, compared to linear or convex optimisation models. A distributed mechanism is desirable for the network to continue to operate properly. In a centralised approach, obtaining the global information of nodes can be very expensive or impossible, since it leads to high communication overhead (bandwidth and energy consumption). Moreover, a central processing unit can be a unique point of failure. A distributed approach prevents network bottlenecks and provides scalability. Therefore, the solution algorithms need to consider the resource-constrained nature of sensor nodes instead of assuming enhanced devices.

Other approaches such as adaptive duty cycling [32, 33] and adaptive sampling strategies [34–37] have also been explored. These proposals include models that enable a harvesting sensor node the use of future energy opportunities to adjust its network parameters accordingly to operate perennially. The algorithms typically adapt parameters such as the transceiver's duty-cycle or sensing rate [38]. These works fall into the category of adaptive parameter strategies. Other energy-neutral designs exploit the spatial variation of energy harvesters and distribute load using adaptive opportunistic routing protocols [39, 40]. While these approaches work well with the expected dynamic energy harvesting, they have the common feature that their performance is limited to the boundary of one network domain. Besides, some existing algorithms require the assistance of a centralised control station to distribute the adjustment of parameters. In addition to these limitations, these algorithms may lead to deficient data collection. Since these techniques dynamically manage a node's operation to throttle its activity when energy supply is scarce and increase it during periods of high availability, the adaptive algorithms may incur in the collection of undesirable data or in the loss of collectable information. More details of the presented approaches discussed in this section can be found in Section 2.5.

Current research has not taken into account exploiting the heterogeneity of nodes or energy harvesting technologies. The traditional view of sensor networks, where multiple sensor nodes belong to one single domain neglects the potential of efficient power management from the cooperation of distinct networks equipped with renewable energy generation and finite storage. Recently, [8] evaluated the effect of cross-boundary energy transfer between two networks. In effect, the excess energy transferred out of an energy-harvesting WSN to an energy-scarce WSN optimises the energy management of the last one. The performance of the benevolent EHWSN on packet delay and energy saving is not affected nor rewarded. However, the high heterogeneity among these networks may have a relevant impact on the result of cooperation [41]. In these situations, cooperation can incorporate artificial intelligence (AI) techniques in order to reach mutually beneficial agreements.

In the cooperation literature, the most important and studied criterion for inter-network collaboration has been the incentive of networks' lifetime maximisation [2, 8, 9, 11–17, 42, 43]. Existing approaches have addressed the cooperation problem using a game-theoretic framework [9, 12, 13, 16, 17], where a WSN is assumed to be rational and selfish. These works have modelled the behaviour of a network as a game to analyse the existence of strategies, looking for equilibrium among rational players that negotiate with each other to maximise their own benefit. Through simulations, they make an exhaustive search on the available space to find the best strategy for each network's authority (i.e., those that form a Nash equilibrium with the highest possible lifetimes) under different simulated parameter sets. They have studied when cooperation is feasible between networks of multiple domains without using incentives (eg. reputation, pricing) [12, 13, 44] or enforcing it through punishment [15] and dynamic pricing [45]. The feasible conditions of cooperation between heterogeneous networks have been also analysed using game theory [9, 12, 13, 16, 17]. However, energy harvesting sensor networks and efficient power management have been left out of this context.

Despite the relevance of game theory for modelling cooperation among rational nodes in WSNs, this approach in practice is usually highly complex and inefficient to implement. The most obvious drawbacks are:

- An unbiased trustable mediator is implicit, that acts to find the agreement towards the Pareto-optimal line by means of Nash axioms using complete information of the players.
- The computational complexity of this search increases significantly as the number of nodes involved grows. The set of available actions for each node needs to be fully defined as well as the possible states the system can reach.

In the absence of a powerful and central cognitive engine, a WSN would necessitate nodes making a significant effort to calculate and store not only all their possible actions at each decision point but also the ones corresponding to their counterpart. Thus, the use of game theory may demand storage and computational capacities that are not suitable in this domain. Similarly, the assumption of complete information is not accurate in opportunistic encounters between nodes.

Before cooperation can be established, networks should be able to negotiate over the set of feasible agreements. Related to this, work in [10] proposed an incentive-driven networking methodology as the calculation of the optimal set of network services to optimize their incentives. This work can be regarded as most related to the presented research, since it also aims to devise a negotiation methodology to engage networks on efficient cooperation. However, their approach uses a central trusted manager and networks integrated by an infrastructure platform using the IDRA framework [46]. Thus, this approach requires a prior backbone network, which is absent in open environments and opportunistic interconnections.

Furthermore, during the cooperation process, networks should be able to select a negotiation partner from the set of co-located and external nodes to exchange offers and find a mutually-acceptable agreement in favour of maximising their aims. The agreement in this domain will consist of an energy flow exchange that supplies mutual demands for both participants to achieve efficient energy management. Although partner selection has been previously investigated in negotiation [47–50], the existing approaches rely on historical records, or information readily available to the parties. The solution, however, must depend on the characteristics of the WSN domain. Related work on partner selection is presented in more detail in Section 2.5.

Considering the autonomy desired in networks of IoT, a mechanism for self-controlling the favours of cooperation is a requirement. Therefore, this thesis motivates the use of an automated negotiation solution to facilitate the cooperation between EHWSNs and solve coordination problems that arise by a conflict of preferences. The solution must be distributed to consider the lack of intervention of an intermediate mediator and deal with the limited capabilities of a sensor node. The main advantage of cooperation based on negotiation is that it allows the establishment of an opportunistic interaction between networks that cannot be conceived at design time about the resources of their neighbours. A negotiation-based methodology leads to a more integrated system of EHWSNs, with the goal of maximising the use of harvested energy. The term OEN is adopted in this work to describe the domain studied here and refer to the developed approach of Opportunistic Energy Negotiation.

Although the cooperation among networks can be enabled through the negotiation process in order to optimise a system-wide goal, every single node involved in the negotiation has a limited view about the state of the entire network. A node only knows how to perform its tasks and has bounded knowledge of other nodes around, either due to its location or constrained nature. Therefore, the impact that each energy flow offer will have in the performance of the entire network is decided based on a suboptimal approach where a node's local state and observations, and those of its neighbouring nodes are taken into account. To optimise the network operability, the nodes must coordinate their actions with those nodes in close proximity. In the same way, the dynamic feature intrinsic in this domain must be taken into consideration, where nodes need to adjust to topology changes, varying environmental conditions and multiple negotiation behaviours. The solution must allow each node to adapt to these variances and achieve the network objective of long-term energy allocation.

A multi-agent approach is a natural fit for this setting, where each sensor is controlled by an agent. The agent engages in communication with other agents in its area in order to achieve system-wide goals in a distributed manner. Since the combination of WSNs and Multi-Agent Systems (MAS) can bring a new generation of intelligent networks, this thesis follows the MAS approach. The use of MAS technologies in the WSN domain has raised a high interest, mostly due to their suitability for modelling autonomous self-aware sensors [51–53]. Therefore, this thesis models a sensor network as a cooperative multi-agent system, i.e. nodes in the same network coordinate their actions and act according to end-to-end goals. In this case, the main task to be performed by the sensor nodes (agents) of different networks is to bargain among them to define an effective transfer of energy that deals with the spatio-temporal profile of their energy sources. From now on, to keep the terminology consistent with the multi-agent system based WSN, the terms “sensor node”, “node”, “agent node” and “agent” may be used interchangeably.

1.2 Research Aims

The aims of the research presented in this thesis can be summarised as follows:

A1: Research and analyse existing energy allocation algorithms for rechargeable sensor networks which ensure energy neutrality to model the energy behaviour of a cooperative network.

A2: Implement and evaluate a negotiation-based cooperation approach, using experimental validation to quantify the benefits in terms of energy allocation on energy management across network boundaries. Use a benchmark to provide a comparative analysis of the technique.

A3: Measure the impact on a network’s performance of establishing an OEN through simulations.

A4: Explore current partner selection models and assess their suitability to allow the identification of the best potential partner in the domain of automated negotiation between WSNs.

1.3 Research Requirements

In order to provide a novel approach for the cooperation problem among co-located and distinct WSNs, the following research requirements need to be fulfilled by any solution. Some requirements arise from the characteristics of the networks and the type of interconnection considered in this work, while others are related to the interaction of the network managers (authorities that administrate each network) in this domain.

The following requirements for a solution to the problem of WSNs cooperation, emerge from the characteristics of the networks and their interconnection. The first is the absence of a dedicated reasoning engine, due to the inter-network communication scheme assumed in this work. The traditional way to build cooperation between independent networks is to connect them via the Internet, using additional facilities like gateways. While this design can support limited task sharing, it is not flexible for co-located networks because it only enables data sharing and depends on the availability of connectivity, which may not be possible in some situations (e.g. a disaster scenario). In a disaster scene, backbone networks like the Internet may be unavailable or unreliable. To overcome the limitations of a gateway-based interconnection architecture, the conceptual framework to construct Opportunistic Direct Interconnection (ODI) was proposed [54]. ODI allows multiple co-located networks to be discovered and identified among each other and enables interconnection between them to help integrate WSNs with native protocols in the IoT. Even if networks can communicate through wireless standards, the absence of an intermediate entity should be considered to avoid unnecessary overhead.

Taking the example of a disaster scenario, one can derive as a consequence, node dropouts and network partitions produced by degraded network lifetimes. Networks can be partitioned because some agents in a specific area can communicate with neither the base station nor any other agents that can communicate with the base station. Fortunately, under these conditions, ODI can be enabled as a contingency support towards the mitigation of the problem and to reduce the impact of failures. In such a situation, reliability can also be enhanced regarding the opportunity for reconnection with disjoint partitions using neighbouring resources. However, before networks can proceed, a negotiation process is essential to ensure effective cooperation. To take full advantage of the ODI framework, in the absence of a centralised entity, the negotiation must occur in a distributed manner where agents decide the energy flow (or services to exchange) between them by themselves. Therefore, a decentralised solution is a fundamental requirement for this problem, which leads to the following requirement.

Since the cooperation is envisaged over a finite period of time (e.g. a day), the issues over which networks negotiate include the series of energy amounts that integrate an energy flow over the corresponding time period. Therefore, this thesis assumes a vector of values negotiated simultaneously. For example, if networks expect to cooperate for 6 hours, they need to negotiate over an energy flow that must include 6 energy values, where the energy able to allocate by each amount negotiated is affected by the energy received in earlier periods, due to the battery's dynamics. Therefore, this type of domain demands a feasible solution for interdependent multi-issue negotiations, which is far more complex to manage [55], even more, when agents employ strategies that require them to learn about the opponent's model to solve the negotiation.

As a result of the interconnection architecture, the participants of the negotiation are the constrained sensor nodes controlled by agents, one agent at each group with overlapping radio range. Since nodes are not powerful devices, the negotiation protocol must be simple, make an efficient use of scarce computational resources and require a minimal amount of processing and messages between negotiators. The reaching of an agreement or disagreement between agents must have a deadline in order to minimise the negotiation time and the communication cost incurred by the submission of offers. Then the goal is to negotiate within a short period of time for a reasonable cycle, by limiting the number of offers and the computation required to calculate them. With the aim of avoiding transmission and energy overheads when multiple agents are co-located, a novel partner selection method to estimate the success of a negotiation must be included in the workflow of the proposed solution. The proposed model should allow networks to maximise in this case, their energy allocation in the long run, while adapting to a highly dynamic and uncertain environment.

Along with the need for a solution that should not depend on heavy computing and the Internet, support for dynamism is also desirable. The topology of these networks is dynamic as sensors work in different states, such as an active, sleep, or dead state, besides, new nodes may be added to the network at any time. The dynamism of the environment where these networks are deployed must also be taken into account. This dynamism may introduce uncertainty into the negotiation process because the networks will not be static on their energy availability and neither about their preferred times to cooperate. Therefore, the variable environment and topology must be considered in the system model.

The assumption of complete information is not possible in opportunistic interconnection. Since networks are not known from their design; rather they are identified after deployment, the models used to evaluate offers and generate counter-offers are hidden from one another. Privacy of information is a requirement to be treated by the negotiation model.

Besides the requirements based on the features of WSNs and the type of interconnection between them, the characteristics assumed of the network's administrators must be taken into account. Because of their natural self-interested behaviour, each individual desires to maximise his own benefit. Therefore, they will enable cooperation solely when it increases their own utility. That is, the cooperation takes place when it is beneficial to all the participants, otherwise, it will not be acceptable.

The conflict is then generated when each network prefers the offer which yields it the highest utility. However, their interests are not completely mutually exclusive due to their energy availability and demand. There are time intervals where their preferences are aligned and there will be a variety of energy flow agreements that can generate some utility to all participants. This utility may not be their preferred highest utility (i.e. energy neutrality) but more than the obtained when networks operate in isolation. Since WSNs are typically placed in remote or hostile environments, they should operate with minimal human intervention. Then, input from network administrators is not appropriate, which indicates to us that a need for automated negotiation for the domain of cooperation between networks towards autonomous IoT is evident.

As the main motivation of this cooperation problem is to optimise the use of harvested energy when networks take into account its availability and their heterogeneity in terms of resources, a solution to the negotiation should be comparable with an efficient outcome, called a Pareto-optimal solution. An outcome achieves a Pareto-optimal state when no participant can benefit from any action without reducing the utility of another. For example, if there is an available amount of energy harvested that can be useful for only one agent, then it must be allocated for its benefit (i.e. the owner of the unused energy cooperates with the tasks of the beneficiary), otherwise, this solution will not be an efficient solution.

The following negotiation characteristics can be observed from the aforementioned description.

(i). **Decentralised.** The interconnection between networks is direct, so the agents are required to negotiate directly with a selected negotiation partner from a group of agents. In this setup, there is no central decision maker. Since each neighbourhood of agents will have a negotiating agent, the protocol needs to be distributed.

(ii). **A negotiation should be kept short.** Time is a major constraint on an agent's behaviour in this domain. The negotiation rules must set a maximum number of rounds. As part of this requirement, networks must minimise their interactions and negotiate for a reasonable deadline. The more interactions it takes, the more processing, communication latency and bandwidth usage it costs.

(iii). **Multiple issues must be addressed simultaneously.** Automated multi-attribute negotiations must be supported by the solution. The negotiation framework needs to consider the interdependency between the energy values and provide offer generation methods accordingly.

(iv). **Adaptable.** From the agent's perspective and its environment, two major changes need to be considered: environmental changes and the dynamic feature of network topology. If the topology of the networks changes because of different nodes' states, such as activeness, sleepiness, and dead state, an agent must be able to adapt to the alterations. Moreover, the agent's different energy profiles in a neighbourhood of agents must be examined to select the most prominent agent to negotiate at every opportunistic interaction.

(v). **Privacy of information.** Networks are self-interested and meet opportunistically, so the disclosure of information is inappropriate. Thus, the models used to evaluate and generate offers are one of the things that negotiators try to hide from each other. The negotiation must consider incomplete information regarding the opponent's utility preferences.

(vi). **Beneficial to all.** Since networks are self-interested they will only compromise if the cooperation brings benefits [14, 16, 43].

(vii). **Automated negotiation.** The negotiation must occur without human intervation.

(viii). **Outcome comparable with Pareto-optimal.** The solution must be compared with a centralised solution where agents reveal their reservation values at the first step of the negotiation and get Pareto-efficient outcomes.

1.4 Research Contributions

In order to solve the conflicts derived by the heterogeneity of networks in terms of resources, nodes need to coordinate. Accordingly, the new methodology presented in this work incorporates a decentralised reasoning mechanism to the cooperative approach between co-located networks with direct interconnection architecture. The research undertaken to design such a mechanism and meet the requirements listed in the previous section has led to the following novel contributions:

- A novel energy allocation scheme that provides a sensor node with the ability to detect potential opportunistic energy cooperation.
  This contribution is presented in Chapter 3 and addresses aim [A1].
- A novel cooperation model based on bilateral negotiation adopting existing negotiation techniques: the alternating-offers protocol, time-dependent strategies and orthogonal counteroffers to enable an agent in the WSN domain to exchange offers and share energy-hungry services (e.g. data processing or packet routing).
  This contribution is presented in Chapter 4 and addresses aim [A2]. This contribution led to the publication of the following document:
  - Ortega, A. P., Merrett, G. V., and Ramchurn, S. D. “Automated Negotiation for Opportunistic Energy Trading Between Neighbouring Wireless Sensor Networks.” *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm).* IEEE, 2018.
- Results showing how negotiation can be delimited by a short-term deadline and end in energy-neutral agreements with a reduced cost of establishment in energy and latency, tested using a network simulator.
  This contribution is presented in Chapter 4 and addresses aim [A3].
- A novel partner selection technique based on multi-armed bandit learning (MAB). With this approach, an agent can learn the best-fixed strategy to achieve an efficient energy allocation in the long run dealing with the dynamism of the domain.
  This contribution is presented in Chapter 5 and addresses aim [A4].

In addition, the following paper has been submitted for publication:

- Ortega, A. P., Merrett, G. V., Ramchurn, S. D. and Tran-Thanh, L. ”Partner Selection of Self-Organised Wireless Sensor Networks for Opportunistic Energy Negotiation: A Multi-Armed Bandit based Approach.”

1.5 Thesis Structure

The remainder of this thesis is structured as follows:

- Chapter 2. This research studies a negotiation process for opportunistic and direct cooperation between independent and autonomous networks. To scope the viewpoint, Chapter 2 provides a literature of work relevant to the topic and identifies the state-of-the-art in the domain of WSN cooperation and energy management. Pertinent negotiation techniques are also reviewed and the existing work on partner selection is discussed. Related work on reinforcement learning for WSNs is summarised as well.

- Chapter 3. The methodology to enable cooperation between networks is described in this chapter. The energy allocation problem for node power management is presented here and a new energy allocation algorithm is derived. The system model for efficient energy allocation across network boundaries is also described. A game-theoretic approach is revised and the optimisation results with and without the cooperation of nodes are analysed.

- Chapter 4. This chapter describes the network simulation and experimental setup to evaluate the effects of establishing an OEN in the network. Key metrics are energy consumption and latency. The heuristic negotiation approach modelled for OEN is also proposed and evaluated here.

- Chapter 5. The MAB model for partner selection is introduced in this chapter. A comparison of the results obtained using three state-of-the-art MABs policies for adversarial environments is performed.

- Chapter 6. Finally, this chapter concludes on the work and provides the potential future directions of this research.

Chapter 2

Cooperation between Networks, Energy Management and Automated Negotiation

In this chapter, an overview of relevant research studies to this thesis is presented. At first, the technology that enables cross-network interconnection to facilitate the interactions between co-located WSNs (Section 2.1) is described. The communication between multi-domain networks is essential for cooperation between them and existing synergy is based on traditional indirect interconnection that depends on the Internet connectivity. However, the proposed work on Opportunistic and Direct Interconnection (ODI) allows co-located networks to discover and interconnect directly in order to trade energy resources. The definition of ODI follows the description of the principal resources that networks might share if they interact (Section 2.2). Since the nodes are often battery-powered, the most studied criterion for their cooperation is the reduction of their power consumption. In fact, the literature review in WSN cooperation is extended in the content of this incentive: the extension of network lifetime with cooperative packet forwarding, and its game-theoretic modelling is emphasized in this chapter. Existing work on bargaining (Section 2.3) as a technique that can be useful when networks want to cooperate but have conflicting interests is then discussed. The work on a heuristic approach for bilateral negotiations is described in Section 2.4. The optimisation of energy use has not been explored before as an incentive for WSNs cooperation. Consequently, a detailed survey of existing approaches to energy management are presented in Section 2.5, highlighting the existing challenges. Then, the partner selection problem is introduced and current solutions found in different domains (Section 2.6) are described. Since the problem of partner selection is addressed in this thesis using a reinforcement learning based technique, related work of its application for WSNs is reported in Section 2.7. In the final section, a summary and discussion are presented about the existing negotiation mechanisms and their adoption in the domain of opportunistic energy negotiation between distinct networks (Section 2.8).

2.1 Opportunistic Direct Interconnection (ODI) between Neighbouring WSNs

Since the application and monitored environment of a WSN have a determinant role in its design and definition of protocols, a WSN is application-dependent. Due to the large set of WSN applications, high heterogeneity in the communication protocols is inevitable. Moreover, traditional MAC protocols do not support opportunistic and direct interconnection because their design principle is to avoid interference from neighbouring networks, which means that co-located networks desist from communicating directly, even if they adopt the same communication protocols. However, the formation of interoperability among these heterogeneous technologies is desirable to assert the concept of IoT and realize the Future Internet. Consequently, the ODI framework was proposed [54, 56] to allow interconnection between independent and co-located WSNs without the need of intermediate facilities (e.g. gateways) and thus integrate them with native protocols in the IoT (see Figure 2.1).

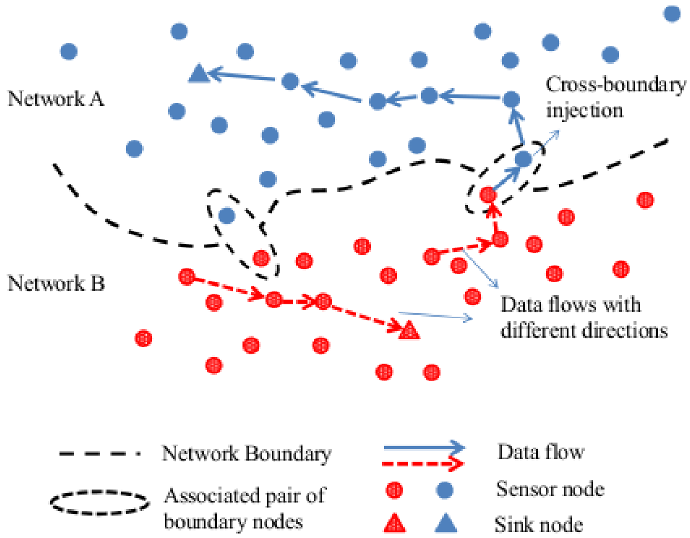

Figure 2.1: The concept of OI-MAC, showing direct opportunistic interconnection between co-located WSNs (taken from [56]).

To solve these challenges, ODI is based on the assumption that co-located WSNs adopt a common physical protocol. While the adoption of IEEE 802.15.4 provides compatibility between radio interfaces, ODI still needs to tackle the heterogeneity of higher-layers. The authors in [54] proposed OI-MAC as the link-layer protocol capable to implement direct and opportunistic interconnection. ODI solution makes the cross-domain protocol (OI-MAC) and the native MAC protocol to co-exist simultaneously. With direct interconnection, networks are not only able to share data but also network resources (packet forwarding, storage or processing) between co-located sensors. Therefore, the proposal of the ODI framework is not to replace existing WSN-Internet-WSN interconnection solutions but to behave as a complementary tool, supporting more beneficial cooperation applications among WSNs that share the same area.

ODI supports the discovery of neighbouring networks in run-time after deployment, without any previous knowledge in design. To help with this, the functions added by OI-MAC into existing MAC protocols for ODI establishment can be defined as: (1) Network discovery and (2) Cross-boundary transmission.

2.1.1 Network Discovery

OI-MAC is an asynchronous MAC protocol with multichannel communication capability. It classifies all available channels into two categories: one as a Common CHannel (CCH) and the rest as Data CHannels (DCHs). CCH is used to detect neighbouring networks and accomplish a handshake process, while DCHs are used for normal data communication. In the discovery process, the networks deployed in close proximity perform two modes of operation: active and passive. In passive discovery, each node switches periodically (time is defined by the discovery period) to the CCH to scan neighbouring networks at run-time with the broadcast of a discovery beacon that contains the network ID and DCH. While this happens, sensor nodes from neighbouring networks listen on the CCH for a discovery period (active discovery). If a discovery beacon is received, the neighbour node replies and the handshake process starts. At the handshake step, the networks exchange necessary information (e.g. network ID, the DCH frequency and wakeup period) and become associated. The node pairs that become associated are called Boundary Nodes (BNs), and constitute a bridge or gateway between neighbouring WSNs. The discovery process is shown in Figure 2.2.

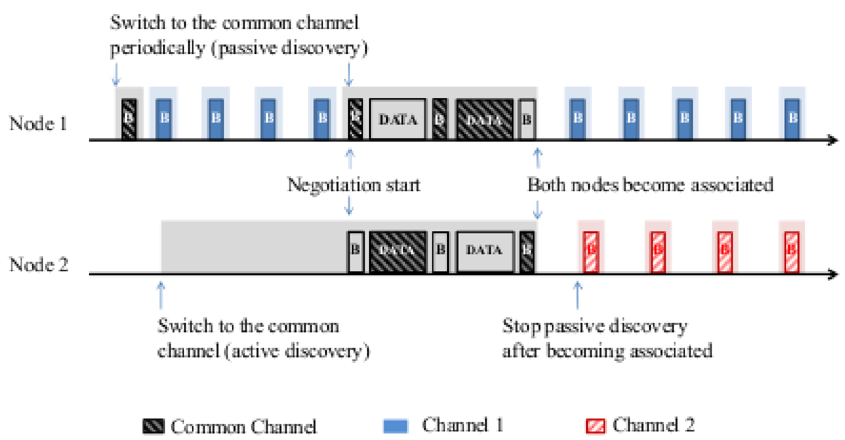

Figure 2.2: The discovery mechanism used in OI-MAC.

2.1.2 Cross-Boundary Transmission

Cross-boundary transmission is performed by nodes (i.e. sending packets into the neighbouring network) once co-located networks become associated through BNs. A node in the boundary must inform the others in its network about the discovery and interconnection. When nodes in the network have packets to transmit to neighbours, they select a boundary node to route, switch their transceivers to the corresponding DCH of the neighbouring network and transmit the packets. Thus, after successful discovery and cross-boundary transmission, ODI enables inter-network traffic to occur through boundary nodes rather than via a dedicated backbone. Results in [54] demonstrated that neighbour discovery function has a minimal effect on latency. Energy consumption also increased insignificantly compared to normal operations of each node that implemented ODI.

Since OI-MAC is a complete theoretical work that defines the details for ODI through the addition of neighbour detection and connection establishment into different types of MAC protocols, further research has been carried out based on its concept. OI-MAC was chosen by [4] to implement it on real hardware and validate it experimentally. They presented the first practical proof of ODI with two independent networks using the OI-MAC protocol, each one composed of 6 Texas Instruments eZ430-RF2500 sensors (see Figure 2.3). Evaluation results showed successful opportunistic discovery and cross-boundary communication between co-located networks. Besides, their work showed that the discovery of neighbouring networks has an irrelevant impact on energy consumption.

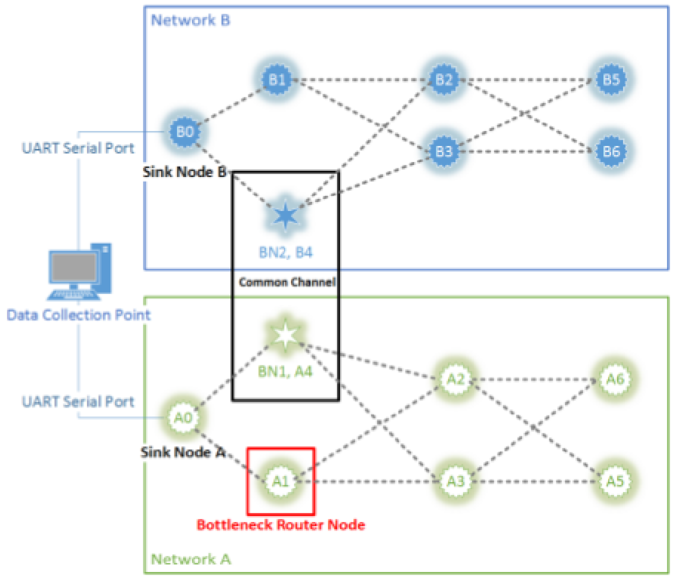

Figure 2.3: The network architecture used in the experimental evaluation of homogeneous interconnection (taken from [4]).

In [4], participating networks implemented ODI but its validation still required the adoption of OI-MAC as the only MAC protocol within the networks. Accordingly, further research was carried out in [57] to test ODI when distinct networks communicate directly using different MAC protocols, which resulted in a successful application. In this scenario, two co-located WSNs composed by the same sensor nodes from Texas Instruments use their own MAC protocol for local packet transmission but in addition, use OI-MAC as the common MAC to add functions of discovery and cross-boundary transmission (see Figure 2.4). The experimental results showed that the energy cost to maintain ODI functionality is insignificant. Mentioned approaches are known as homogeneous interconnection and heterogeneous interconnection, respectively.

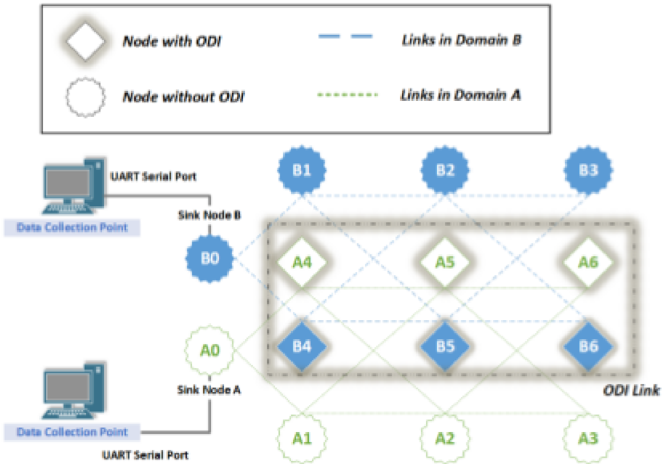

Figure 2.4: The network architecture used in the experimental evaluation of heterogeneous interconnection (taken from [57]).

Since the feasibility of ODI establishment was confirmed by previous research, further works using the OI-MAC protocol logic to study interconnection and direct cooperation between networks can continue. Thus, this work assumes the ODI interconnection architecture. As the trend towards IoT increases, more WSNs are deployed and cooperation between them is an opportunity. This can not only bring functional benefits to WSNs applications but also economic profit to the network stakeholders. One crucial factor to establish cooperation between distinct parties is the knowledge of the costs and benefits that cooperation will bring to the participating networks. Therefore, the inclusion and evaluation of a negotiation mechanism becomes an important research topic towards cognitive networks and autonomic environments in IoT.

The following section describes the potential benefits of ODI-based cooperation if networks agree to cooperate. A brief literature survey on wireless sensor network lifetime extension based on cooperative packet forwarding and its modelling using game theory is also presented.

2.2 Resource Sharing between Networks and Related Work

Now that the principles of ODI have been detailed in Section 2.1, the potential benefits and means of sharing and exchanging of resources between neighbouring networks will be covered in this section.

The most visible challenge of WSNs is the energy consumption problem. As a result, the extension of network lifetime has been the main motivation for a great deal of research on WSN. A section (Section 2.2.3) is dedicated to delving deeper into the cooperation problem in multi-domain sensor networks and its modelling using game theory when networks aim to prolong their lifetime. Indeed, the literature found about cooperation between distinct networks focuses only on that aim. However, it is necessary to consider that other incentives may exist (see Figure 2.5). In this regard, a negotiation process can be included to ensure that the performance of none of the participating WSNs is degraded after cooperation. Moreover, negotiation may facilitate their coordination or resolve conflicts derived from their heterogeneous characteristics and self-interest.

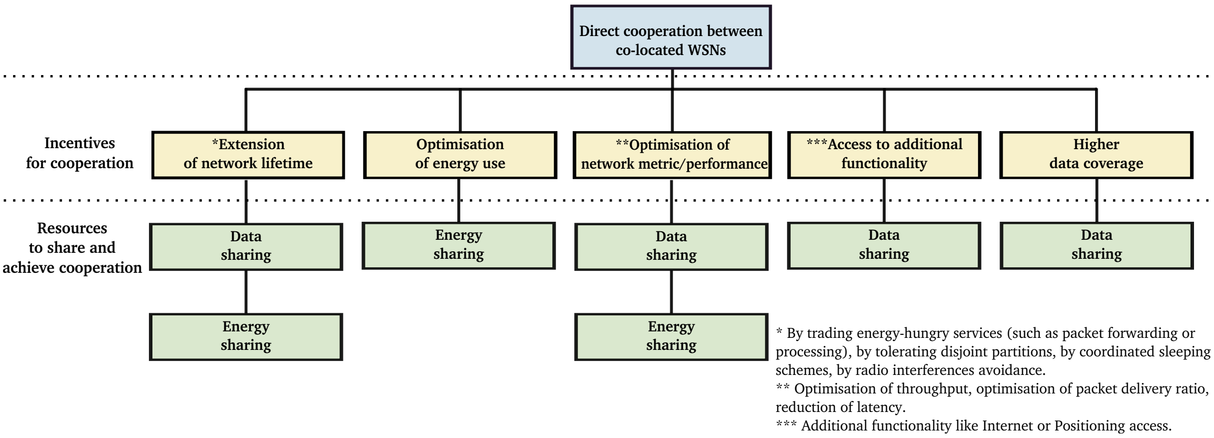

Figure 2.5: Incentives and resources to start direct cooperation between WSNs.

With reference to Figure 2.5, a shared resource, is a mean accessed by the participant WSNs to achieve a cooperation incentive. Supported by existing literature [2, 6, 8, 14, 16, 42, 43, 58], the resources that can be shared directly between co-located WSNs are: data and energy.

2.2.1 Data Sharing

The most widely adopted interconnection architecture between WSNs is the use of backbone networks like the Internet, which are originally designed for data sharing (see Figure 2.6).

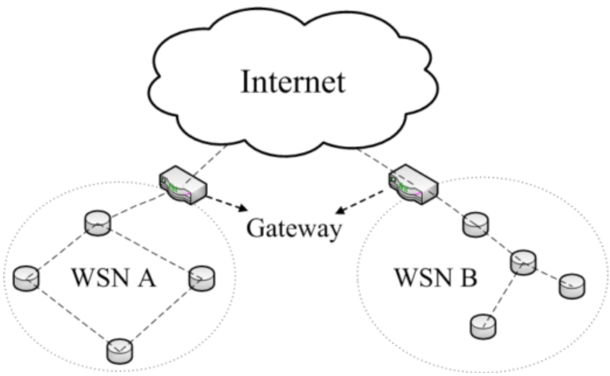

Figure 2.6: Network architecture of integrating WSNs into the Internet (taken from [54]).

The proposal of the ODI framework is not to replace existing WSN-Internet-WSN interconnection solutions but to behave as a complementary tool, based on which more beneficial cooperation applications among co-located WSNs can be developed [54]. In [2], the authors consider the architecture illustrated in Figure 2.6. Their research exploits the correlation between different types of data that sensor nodes from different networks may be able to measure. Yet, direct interconnection can offer an alternative channel for data exchange in the absence of a backbone network.

In the era of high availability of cloud data, the use of data from an external WSN can help in reducing energy consumption and improving quality of measurements made by a target WSN. A specific use-case is given in [2], where the operation of a WSN measuring relative humidity is expanded using the data obtained from a WSN measuring temperature. Periodically, the data retrieved by the nodes is transmitted to the sink and this consequently reports the values to a centralised device with enhanced capabilities (EG). The EG uses the collected data from both networks to predict changes in the future of the monitored target. Accordingly, it selects the new configuration for the WSN, which has to be applied by the sink. The new configuration includes the (de)activation of sensor nodes or changes in the sensing intervals. This application can reduce the energy consumption of the network by turning off nodes during a certain period of time. The study assumes networks are benevolent and are motivated to help each other. Such an assumption is not valid when agents are considered self-interested.

A Body Sensor Network (BSN) is a wireless network of wearable computing devices applied to the human body to monitor physiological signals. While a BSN fits under the category of a WSN, there are several differences between them [59]. However, work in [60] is under the important trend towards pervasive computing and the IoT that motivates research on cooperation between co-located networks. A framework called C-SPINE is proposed to allow data fusion from different Collaborative BSNs (CBSNs). Specifically, C-SPINE supports cooperative systems destined to co-located groups of people. CBSNs enable interaction and synchronization in collaborative applications for recognizing group activity, detecting events sensed by groups, and monitoring multiple individuals. Data shared between distinct WSNs allows each network to build a broader knowledge about its surroundings and cover wider areas to have many perspectives of the same monitored phenomena.

Data sharing can also be used between WSNs to expand data coverage. Zia et al. [3] pointed out the importance of data sharing between the WSNs in co-located farm fields to have an effective water quality management at a catchment scale. The runoff generated by agricultural irrigation in an upstream field is used by a neighbouring WSN to predict the repercussions in its monitored farm located downstream. The coverage of this additional information can enhance a control strategy by reducing irrigation water if farmers can predict the arrival of the runoff. The authors denoted the importance of direct interconnection and cooperation schemes to share network resources and assumed that good neighbours provide their data services for free. The work in [61] describes the benefits of cross-domain data sharing as the extension of network boundaries and enhancement of its scalability. However, this work is concerned with security access control and efficient data analysis. They do not take into account the individual needs of these networks, even though all these devices are very different in terms of application requirements and capabilities. The introduction of negotiation may help control conflicts between networks with different preferences.

2.2.2 Energy Sharing

In [8], opportunistic energy trading between co-located energy-harvesting and battery-powered WSNs is proposed. A EHWSN with excess energy transfers some to an energy-scarce WSN in order to extend its lifetime. The sink node in the EHWSN identifies an excess of energy based on the percentage of remaining energy informed by the nodes of its network and decides when cooperation should take place.

While it is convenient to envisage opportunistic energy negotiation as physically transferring energy across a network boundary, energy is actually logically transferred through the acceptance of energy-hungry services (i.e. in a form of accepting energy-consuming tasks such as data processing or packet forwarding) [11–13]. Let us consider a scenario where two nodes from different networks negotiate and reach an agreement on an energy flow. If they decide to participate in the energy transfer by cooperative packet forwarding, each participant node enables a path between the distinct networks to forward packets opportunistically. The energy consumption model of the networks defines the energy spent in transmissions. The nodes involved in such cooperation control the agreed energy flow by asking for/providing routing favours. In this way, nodes from different networks may participate in cross-boundary energy transfer by forwarding each other's packets. The cooperation is decided ahead of time, autonomously by negotiation, but a control policy or mechanism is required to ensure that networks respect the agreement.

The following describes the literature related to energy sharing and the services employed for its transfer.

2.2.2.1 Cooperative Data Processing

Data processing comprehends the analysis of extensive data generated in a WSN performed by each node. Such analysis is necessary to extract information that is meaningful for its consumers. Data aggregation and compression is part of this definition. Aggregation is a mechanism of a WSN to preserve energy consumption by combining data packets and eliminating redundancy from multiple sensor nodes in one data packet. In this respect, data aggregation is also known as data fusion. As a result of the elimination of redundancy in the sensory samples, the transmission cost and network overloading are also lower. Since nodes require low-energy components given their limited batteries, cooperative data processing between distinct WSNs may reduce the energy cost of this task.

In [8], Teng et al. proposed a novel power management strategy based on Opportunistic Energy Trading (OET) between co-located WSNs, where one network was composed by solar-harvesting sensor nodes and the other by battery-powered sensors. They consider a setup where a solar EHWSN harvests energy at different times because of its dependence on the movement of the sun. Since a WSN is constrained by an area, intra-network power management may be inefficient to control this spatial-variance. However, OET allows energy to be transferred through the boundaries and be used by multiple networks and applications. The design aim of the energy transmission process is to use excess harvested energy by one network to extend the lifetime of its neighbour (battery-powered network) by taking over consuming tasks (for example data processing). The battery conditions are maintained as an energy map of the entire network by the sink to identify whether or not there is an excess of energy when a threshold is reached. Once the sink realises that all batteries are above the threshold, it initiates a broadcast to inform neighbour WSNs to start the energy transfer process. Hence, the sink node decides when cooperation should occur. This approach is the one this research employs, where the requirement for cooperation is identified and established opportunistically after network deployment. However, in Teng's work, the provision of a negotiation process is excluded. Although there is currently sufficient literature on cooperative data processing [62–65], no prior study that addresses multi-domain cooperation of data processing exists.

2.2.2.2 Cooperative Packet Forwarding

Cooperative packet forwarding in multi-domain sensor networks becomes crucial for disjoint sections, where partitions can be fixed or prevented by direct cooperation between co-located WSNs [42]. Kinoshita et al. [14] proposed the introduction of shared nodes in WSNs, that can use multiple channels to relay data packets between different networks. They considered the cooperation of multiple WSNs that are deployed in the same area to address the energy hole problem, which is one of the most important issues in WSNs. Such problem can be found in nodes located around the sink, where usually many more packets are relaying than in other nodes of the network. These nodes tend to suffer from battery drought faster than other nodes since processing and communication tasks are more energy consuming than sensing [66]. Hence, the lifetime of the whole network can be prolonged with cooperative packet forwarding, by balancing the communication load at heavily loaded nodes around the sink.

The benefits of direct cooperation have also been studied by Nagata et al. [67]. Cross-network routing in spontaneous cooperation without incentives is proposed to extend the lifetime of co-located WSNs. Moreover, traffic load balancing is achieved by reducing the load on nodes around sinks in a multiple-WSN environment (see Figure 2.7). The same idea is shared with authors in [45]. However, they enforce cooperation using an incentive model based on a dynamic pricing routing algorithm.

In addition, packet routing can be further optimised with cooperation. The authors in [68] investigated the potential performance gains achieved by cooperative WSNs and concentrated on the routing performance. By exploiting hybrid nodes that allow cross-network sharing of resources, their simulation results and quantitative models showed how cooperation provides better performance for the communication of two nodes than the local routes. They analysed how the collaboration between distinct networks may reduce the average routing cost of the shortest path.

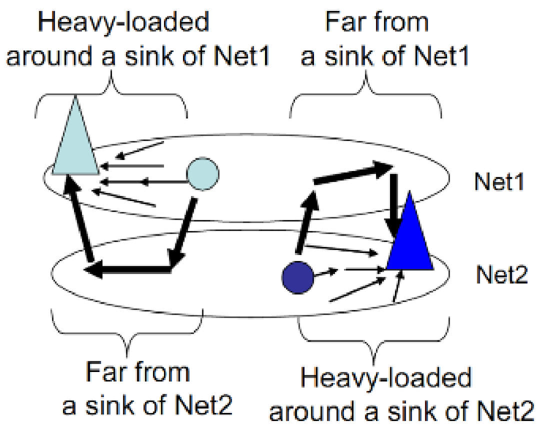

Figure 2.7: Load balancing based on the cooperation between co-located networks interconnected directly: Two WSNs have cooperative packet forwarding. One WSN uses free nodes to perform packet routing in order to reduce the traffic load around the sink (taken from [67]).

From the incentives described in this section, the idea of cooperation seems simple and straightforward, however, there are several implications that obstruct it [41]. Because of that, the problem of cooperative packet forwarding in multi-domain sensor networks has been studied using mathematical models of conflict from game theory. Previous works have searched for conditions where cooperation between networks from different domains is possible using game theory as a tool. The next section is dedicated to review the existing work in the field of energy sharing by cooperative packet forwarding and its modelling using game theory.

2.2.3 Energy Sharing and Game Theory

Game theory (GT) [69] is well used to model the behaviour of a system as a game and analyse the existence of strategies to achieve equilibrium among rational players that negotiate with each other to maximise their own benefit. By using this approach, players are assumed to have complete information about the game and also on the behaviour of opponents.

The works reviewed in this section assume networks as rational decision-makers with complete knowledge of the game settings. In reality, such assumptions may devote time and unbounded computational resources that are absent in these nodes. However, GT is an ideal platform to systematically study interactions in multi-agent systems, mathematically capturing the behaviour of players in a strategic situation [70]. In particular, game theory provides solution concepts (e.g. Nash equilibrium or dominant strategy) to predict the outcome from the correspondent interactions between agents and to analyse its properties.

In the WSN cooperation domain, the normal form of GT is represented by $G = (N, S, U)$, where G is a particular game, $N = \{n_1, n_2, ..., n_n\}$ is a finite set of networks (players) and $S = \{S_1, S_2, ..., S_n\}$ is the strategy space of the network k, for each $k \in N$. A strategy determines each action that a player takes at any stage of the game, i.e. the complete algorithm to play the game. A decision maker in a game is called a player, but the term agent is also used in the computational context.

GT in the WSN cooperation domain has been used to model strategic decision situations to find an optimal set of actions for extending the lifetime of networks. The negotiation is carried out by trading routing favours with cooperative packet forwarding [9, 13, 16, 71]. For example, in [13], network authorities must decide between two situations. In particular, networks need to define if they help each other to increase their lifetime or if they ignore potential help from others and only rely on their own nodes to send packets. Their simulation results proved that networks can have an important benefit by mutual service, using their sinks for other networks in sparse and hostile conditions. In these works, the game-theoretic modelling of cooperation shows how the objective of each player is to maximise its own payoff. This is what characterises a rational decision-maker, where each individual network has its preferences, beliefs about the context (including the other players) and looks for strategies that maximise its own gain.

Game theory is also used to determine the existence of strategies to reach a steady state; the *Nash equilibrium* [72]. The Nash equilibrium is a solution concept that represents a state which involves two or more players, where the participants can not benefit by changing their strategy if the strategies of the other players remain constant. In this context, equilibrium is found when the networks decide to not deviate from their selected strategy (e.g. providing routing favours and asking for them or not cooperating at all); otherwise, they decrease their utility. To measure an equilibrium in terms of how beneficial the resulting outcomes are to the group of players as a whole, there is a property called *Pareto-optimal*. An outcome is Pareto-optimal when no player can benefit from any action without reducing the utility of another. For example, in [16] the creation of different cooperative pairs of edges from both networks involved, makes their payoff increase. After creating all pairs of cooperative edges, the payoff of the networks is the maximum that can be achieved with their strategy, being, a Pareto Optimal solution. In this scenario, if there are nodes that are not part of cooperative edges, which are of no importance to any network but one, then the edges must be created to the benefit of that network, otherwise, this solution will not be Pareto-optimal.

Vaz de Melo et al. [16] used game theory to model the problem of cooperation between two WSNs as a repeated Prisoner's Dilemma game, where the only way to ensure cooperation is by means of a protocol. Otherwise, the networks have good incentives to deviate from cooperation. In contrast to other works, the effectiveness of cooperation to extend network lifetime is not only proven but also supported by a distributed protocol called Virtual Cooperation Bond (VCB). VCB makes different WSNs cooperate and work together to reduce the energy consumption of their communication process by the creation of cooperative edges in both participants and the control of routing favours. They use a cooperative game non-zero sum (rewards and losses are less or more than zero) related to the Iterated Prisoner's Dilemma, where the players or co-located WSNs coordinate their actions to get better payoffs.

Similar to [16], the authors in [9] propose an algorithm that enforces cooperation among distinct WSNs. The difference here is that their solution is adaptive and starts with generosity at the beginning of the game. When nodes are aware of a decrease in their battery, this level of cooperation is minimised. Their cooperation model is energy-aware based on the TIT-FOR-TAT strategy and falls into the category of non-cooperative games. This strategy is behavioural, which means that each action is based on a previous one. In their simulations, the cooperation is evaluated considering hierarchical routing and heterogeneous network topologies. The results showed that their algorithm based on TIT-FOR-TAT can increase the lifetime of a network in competition with different types of opponents.

There are different kind of games to model strategic interactions between individuals in conflict situations. To solve any game, an appropriate mathematical representation for the situation in analysis must be created. A model can include the participants, strategies, decisions, consequences, and utility functions. Then, the model is solved by computing the best/optimal strategies for all participants or for some of them.

Game theory has been used to analyse conflict situations that are generated from multi-domain networks, the strategies and conditions of their cooperation. However, to the best of our knowledge, there is no work in the design of a distributed negotiation mechanism where agents (nodes) bargain over the energy and times, which allows them to decide whether cooperation is feasible or not, considering the properties and requirements of their owner nodes. The introduction of a negotiation mechanism can help to resolve the different needs that each network may have before deciding whether to cooperate or not. This thesis presents an alternative approach, where instead of controlling how the cooperation among networks is performed, the proposal is to set an encounter where agents decide how cooperation will proceed.

When the goal is to provide an automated negotiation model on energy sharing for WSNs, computation complexity must be considered. If the techniques and results of game theory are applied, the complexity increases when the number of participants too. In these encounters, the outcome depends on the choices made by all networks and nodes in the scenario. This implies that in order for an agent to make the choice that optimises its outcome, it must reason strategically and take into account the decisions that other agents may make, plus the assumption of full rationality that implies that the others will act so as to optimise their own outcome. For WSNs, this would involve that nodes must make an effort beyond their means, considering that these are devices with limited memory space, battery capacity and processing capability.

One approach to deal with this complexity is to simplify the settings in which agents interact with each other and use heuristic methods. Game theory is only used in this work to examine situations of bargaining. The following section explains related work on bargaining that can be useful to address the problem of cooperation in the energy domain.

2.3 Bargaining

From a commercial perspective, bargaining refers to the process where a buyer and a seller debate over the price and conditions of a transaction to reach an agreement. In this setting, one agent assumes the role of the seller, while many may play the role of the buyers. In contrast, in the domain studied in this thesis, the agents are homogeneous in role. So there is not one buyer and one seller but two agents trying to reach a mutually acceptable agreement on redistribution of energy in order to optimise their power management.

Bargaining can help networks decide on whether to coordinate their actions while performing their tasks or act independently. The amounts of energy over time represent the issues of this negotiation. Each agent (together with the sensor node it controls) consumes resources such as energy, time, bandwidth of the communication channel, and computational power, and each agent endeavours are directed to utilise its resources efficiently. The full description of the negotiation domain is given in Chapter 3.

The following defines the bargaining problem and discusses related concepts.

Definition 1: The Bargaining Problem
The bargaining problem refers to a situation where two parties need to agree on an appropriate outcome from a set of possible solutions but their interests do not match. The solution set is conformed by all feasible deals or agreements [73]. The problem studies how two agents share a surplus that they can jointly generate. In the domain of WSNs, agents bargain on how to decide the amount of energy over time they both share to optimise their use of energy and deal with the spatio-temporal profile of their energy sources. The surplus is the energy and the means may be the relaying/forwarding/processing of packets.

Bargaining is bilateral when it involves two players, which is the case studied in this work. Two agents are considered for each pair of nodes (one-to-one), one from each network, that have a common interest in cooperate but have conflicting interests regarding the particular times and rates of doing so. The terms negotiation and bargaining are used indifferently in this work.

The overall order of actions during a negotiation is usually restrained by certain rules. According to the alternating-offers game [73], all of the participants around the table get a turn for making offers and counter-offers in a sequential order in rounds. The rules are set by the so-called bargaining protocol.

Definition 2: Bargaining Protocol
A bargaining protocol (also called negotiation protocol) specifies the rules that govern the interaction among negotiating participants [74].
By using the word protocol in the negotiation context, it refers to high-level protocols, it is not about communication. It determines the kinds of deals agents can make or restrictions, such as deadlines, as well as the sequence of offers and counter-offers that are allowed.

The structure of an offer is determined by the issues/attributes that are the subject of the negotiation and are defined by the domain. In a WSN cooperation, there are many different domains depending on the goals of each participant network. The cardinality of the negotiation is one of the characteristics that classifies a bargaining problem [75]. A domain can contain a single issue of conflict and require a single issue bargaining. Or involve multiple attributes and demand a multi-issue bargaining. In the scenario assumed in this thesis, the bargaining process includes an expected cooperation time that may involve different amounts of energy over time, which suggests a multi-issue bargaining.

Since bargaining theory is a sub-field of game theory, it adopts its principle concepts as individual rationality and Pareto-optimality, which define desirable properties in the outcomes of a bargaining.

Definition 3: Individually Rational
A bargaining solution is individually rational if it gives each player at least as much utility as it would get by himself in the event of no agreement [76].

Definition 4: Pareto-Optimality/Pareto-Efficient, Pareto Frontier
An outcome is Pareto efficient when no alternative outcome exists that is more preferable for at least one player without making another player worse off. The Pareto frontier is the set of all the points of Pareto-optimal outcomes, where each coordinate point corresponds to the utility of a player.

In this sense, the bargaining game can be abstracted from the negotiation model and instead, a set of properties (axioms) to be satisfied by a negotiation outcome can be specified. When the parties involved ignore the strategic aspects of bargaining, the bargaining can be categorised as cooperative or axiomatic. Axiomatic bargaining defines the criteria and axioms in the bargaining process. In contrast, when the parties consider the strategic aspects of bargaining (such as rules and opponent actions), bargaining is considered non-cooperative or strategic, which studies the strategies of players.
The existing work on these two types of bargaining is discussed in the next sections.

2.3.1 Axiomatic Bargaining

Game-theoretic analysis of bargaining can be done using one of two approaches: axiomatic or strategic. Axiomatic bargaining is focused on finding an appropriate bargaining solution through a mathematical investigation of properties but not in the process [77]. It first sets the axioms that reflect the desirable properties of the solutions and then tries to compute the outcome subjected to those properties. Another key difference with strategic bargaining is that in cooperative games there is a third-party mediator or arbitrator implicit, which is who controls the binding of agreements. By this scheme, it is possible to enforce negotiation outcomes that are mutually beneficial for the parties involved but may require reasonable compromises from an individual player's perspective [76].

Axiomatic bargaining theory originated with Nash [78]. In his work, a set of bargainers $N = \{a, b\}$ tries to come to a solution over a set of feasible agreements $\mathscr{A}$. A solution is a determination of how much each individual should gain from the situation. In case the agents fail to reach an agreement, a disagreement outcome denoted by d occurs. Nash analysed the *bargaining problem* and without modelling the negotiation process defined a result, which is the most popular solution concept known as the *Nash Bargaining Solution*. He idealised the bargaining problem under the assumption of perfect rationality, which means that the agents can compare the payoff of their possible outcomes, have equal bargaining skills and full knowledge of the preferences of each other. In this model, a utility function $u_i : \mathscr{A} \cup \{d\} \to \mathbb{R}$, for $i \in \{a, b\}$ represents the preference relation or payoff of an agent over a set of outcomes. The set of all utility pairs that result from an agreement is called the *bargaining set* $S = (u^a(z), u^b(z)) \in \mathbb{R}^2_+ : z \in \mathscr{A}$. And the pair $d = (u^a(d^a), u^b(d^b)) \in \mathbb{R}^2_+$, corresponds to the utility values of the disagreement events, represented as well as $d = (d^a, d^b)$.

Nash assumes the bargaining set to be a compact [1] and convex [2] subplane of $\mathbb{R}^2_+$, where graphically the pair of disagreement values (d^a, d^b) are a vertex in this subset. Thus, if the agents agree on $z \in \mathscr{A}$, it means that agent a gets a utility of $u^a(z)$ and agent b gets $u^b(z)$. However, if the agents do not reach an agreement, then the outcome of the negotiation is the disagreement point d where agent a gets d^a and agent b gets utility d^b. Therefore, a point (u^a, u^b) represented in the set S describes the individual shares of agents, and since agents will only cooperate if they get more utility than their disagreement values, the following $(u^a, u^b) \geq (d^a, d^b)$ must hold. Nash then defined the pair (S, d) to be a bargaining problem, where a complete theory of negotiation would ideally allow us to find some solution $f(S, d)$ in $\mathbb{R}^2_+$.

Let the set of all bargaining problems of the form (S, d) be denoted by $\mathscr{B}$, Nash defines the bargaining solution as a function $f : \mathscr{B} \to \mathbb{R}^2_+$ that computes, for every bargaining problem (S, d), a unique solution $f(S, d) \in S$ when S is a bargaining set convex and compact [79]. If x and y are the shares of agents a and b respectively, then the Nash bargaining solution is calculated by:

$$\arg\max_{x,y} \, [u^a(x) - d^a] \times [u^b(y) - d^b] \tag{2.1}$$

The solution to this equation (i.e. the values of x and y) is the shares of agents a and b respectively, i.e. the points that maximise the product of the individual utilities (see Figure 2.8).

In his work [78], Nash specifies a list of axioms that are satisfied by the Nash bargaining solution:

A1 Individual rationality. This axiom indicates that a reasonable solution must give each player at least as much utility as they get when an agreement can not be reached.

[1] A set S is compact if it is bounded.

[2] A set S is convex if for all x and y in S and all θ in the interval [0,1], the point $(1-\theta)x + \theta y$ is in S.

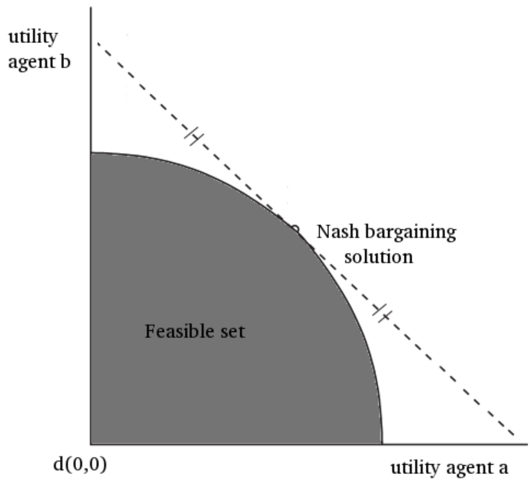

Figure 2.8: Nash Bargaining Solution. This figure shows the Pareto-efficient frontier (the solid line), the disagreement point as (d^a, d^b) = (0,0) and the Nash bargaining solution for a specific bargaining problem.

A2 Symmetry. This axiom refers that the solution should be independent of the identity of the agents (be anonymous) and rely only on their utility functions.

A3 Strong efficiency. This axiom states that the solution should be feasible and Pareto optimal. In a bargaining problem (S,d) with s, $s' \in S$ and $s'_i > s_i$ for $i = a,b$, then $f(S,d) \neq s$.

A4 Invariance. This axiom refers to the characteristic that the bargaining outcome should not change as a result of the application of linear variations to the utility of the agents, i.e. rescaling an agent's utility should not change the result of the negotiation.

A5 Independence from irrelevant alternatives. This axiom states that if some outcomes s (feasible choices which would not have been chosen) are removed from S, but s' is not, then s' is still the solution.

The unique solution concept that satisfies the above axioms is the Nash bargaining solution. Some other important axiomatic bargaining solutions are the utilitarian solution, Egalitarian solution and Kalai-Smorodinsky bargaining solution [80, 81]. The utilitarian solution is also known as the *social welfare solution* [82]. It is any function that selects an allocation for two bargainers as the maximum sum of their utilities. This solution, however, fails to be invariant to the calibration of the players' utility scales (axiom of invariance). When the utility functions between agents have a great difference, the solution benefits the agent that values the goods the most but the rest do not receive any share. From an individual's point of view, it results in agreements that are not satisfactory. On the contrary, in situations where all the agents have the same owner, the utilitarian solution gives all the goods to the agent which has the highest valuation for them and thus, it maximises the overall utility of the group. Likewise, the Egalitarian solution uses a social welfare function but it attempts to provide a social solution in terms of the utility of the individual who is worst off. The Egalitarian solution satisfies axioms 2,3,5 with the exception of 4, which leads to the same problem discussed in utilitarianism [83]. Like Nash bargaining, Kalai-Smorodinsky function determines agreements that are fair in some sense. The Kalai-Smorodinsky solution is the result of making proportional shares to the players. The solution accomplishes the axioms listed above except for the independence of irrelevant alternatives, which is replaced by a monotonicity property [81]. For a set S of individually-rational and Pareto-efficient points, let $m^i = max\{u^i\} \in S, (\text{for}\, i = a, b)$ be the maximum utility that an agent i could achieve individually. The Kalai-Smorodinsky solution corresponds to the point where it intersects the Pareto-frontier; i.e. the point that joins the disagreement point (d^a, d^b) with (m^a, m^b) (see Figure 2.9).

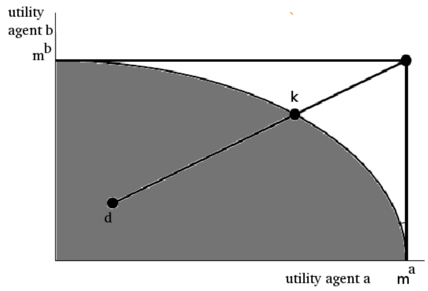

Figure 2.9: Kalai-Smorodinsky solution. The point k is the unique solution that satifies the axioms proposed by Kalai and Smorondinsky [81].

The solution concepts described above enclose mathematical mechanisms to find a bargaining solution characterised by desirable properties called axioms. The availability of information in making social decisions is implicit in this approach. For example, to compute the Nash bargaining solution, the utility functions of all the negotiating participants must be known by the agents, or they must be under the control of a centralised entity (arbitrator/mediator) that can enforce the outcome of the negotiation. In other words in a cooperative negotiation, a bargaining situation is described only by the utility functions of the agents involved and the disagreement point, which ensures the individual rationality characterisation. The approach, however, ignores the process and its elements (such as time). On the other hand, the strategic bargaining focuses on the process. More details are given in the next section.

2.3.2 Strategic Bargaining

In this game-theoretic approach, negotiation is modelled as a non-cooperative game. In contrast to the axiomatic method, the analysis here is concerned with the strategies chosen separately by the players, conducted for the best of their interests and given the strategies of the others. All the strategies involved, rules and payoffs are known in advance by the players. Complete knowledge among participants is a common assumption in these solutions.

The bargaining protocol is what sets the rules of an encounter. These protocols can be used by two bargainers to reach an agreement over a single or multiple issues. There are many different protocols in negotiation [73, 79, 84], but the most influential of these works is perhaps the *alternating-offers protocol* defined in Rubinstein's dividing pie problem [73]. In his work, the issue over which negotiation takes place is a unit-sized and continuously divisible pie, where the bargaining situation is about the conflict that two players have in order to decide an agreement on the partition of the pie. The model considers then a single-issue negotiation. Negotiation takes place over a sequence of rounds, it initiates when Agent a makes an offer of its share (portion of the pie) at time *T=0*. Agent b immediately accepts or rejects the offer. If the proposal is rejected, then Agent b makes a counter-offer at round *T=1*. In [73], time is assumed to be valuable for both players and this is represented by a discount factor or a fixed bargaining cost. The discount factor model is most popular than the fixed-costs model. Discount factors are used to model how impatient the player is [79] and the results agree with commonsense: patience in negotiation pays dividends. In [73], there is no restriction on how long negotiation can last (no deadline) and the game continues until an agreement is found (*infinite-horizon bargaining*) (see Figure 2.10).

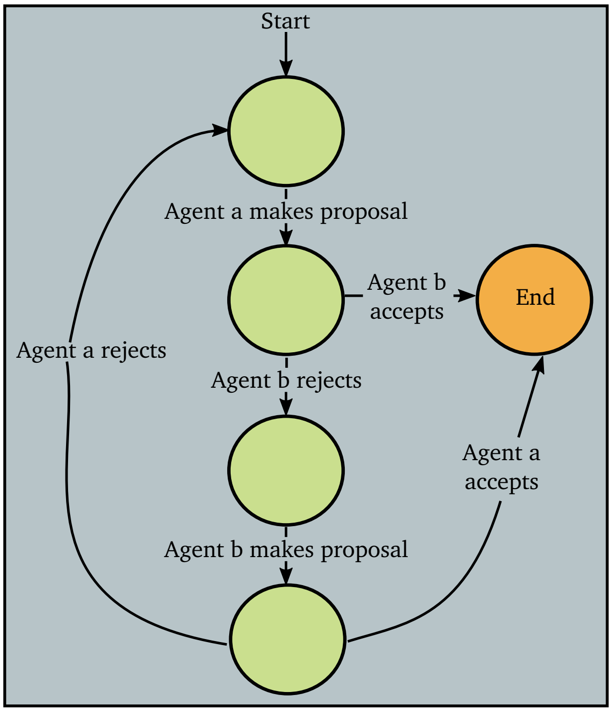

Figure 2.10: The alternating offers protocol. A process of making offers and counter-offers continues until an agreement is reached. (reproduced from [85]).

There are two main types of alternating offers protocols: *infinite horizon* and *finite horizon*. For the latter, negotiation has a definite deadline. A strategy is then a sequence of offers and replies for each stage of the game as a function of the negotiation history. Rubinstein assumes in [73] that players have common knowledge and can exchange unlimited offers. In order to analyse the protocol, he also assumes that disagreement is the worst outcome and agents pursue to maximise their utility. Since the pie is an infinitely divisible good and the protocol allows infinite rounds, any division is a Nash equilibrium and the alternating offers bargaining scheme can be an infinite set of Nash equilibrium outcomes when players are patient. However, Rubinstein showed that for an infinite game where offers are discounted, a unique solution is reachable in the first step of the protocol. The basic technique used in this analysis involves considering the last negotiation round and what players could get from that last round. With this information, a backwards induction is used to determine the optimal strategy of the player who makes the last move of the game. In this way, it determines the Nash equilibrium of each subgame (each stage of the original game). This is known as a subgame-perfect equilibrium (SPE), which is the concept that Rubinstein applied to prove that a unique SPE exists in the alternating-offers bargaining model.

Definition 5: Subgame-Perfect Equilibrium Strategies
An extensive form game is in SPE if the strategies constitute a Nash equilibrium at every decision point.

Rubinstein's work has been very influential in bargaining theory on infinite-horizon games. His analysis was further extended considering incomplete information [86]. A game with incomplete information involves players with limited knowledge. Information such as preferences, discount factors or utility functions. The consideration of incomplete information is more realistic in real-world negotiations. However, game theorists traditionally model incomplete information on other player's preferences and beliefs by describing player types. Thus, common knowledge for all players is still assumed, since the games require that participants know the probability of a player's type.

Considering that players may have bounded computational capacity, which is the case in this research, even with the above assumptions, unbounded computational resources and perfect memory are still required. These constraints limit the application of game-theoretic models in this work. The setup differs from Rubinstein's model, where the agents bargain over multiple issues (multiple time periods where the amount of energy to be exchanged in each time period can be different) instead of a single issue. For this reason, Rubinstein's model as it was originally designed, with an infinite exchange of offers is only applied in simple negotiations. In order to be applied in the WSNs cooperation domain, it would require a deadline or involve additional states and therefore, several experiments to investigate the effects of the deadline on the outcome of the negotiation.

To address some of the aforementioned limitations of the game-theoretic analysis, heuristic methods may be used to provide a reasoning mechanism that arises much less complexity.

2.4 Heuristic-Based Approaches for Automated Negotiation

Equilibrium solutions from cooperative game theory are difficult to apply in practice, especially in negotiations with incomplete information or non-linear utilities. Nash equilibrium, for example, does not consider computation cost and ignores cases where players are not aware of all aspects of the game. In contrast, research work in AI on the problem of multi-issue negotiations focuses on learning and heuristic methods to build automated negotiation models and tractable negotiation strategies. Heuristic methods take into account agents with limited computational resources and relax the requirement of processing capacity to produce good instead of optimal solutions [74, 87]. In the negotiation context, heuristics are useful for the design of agents and their ability to generate initial offers, evaluate proposals and decide counter-offers, based on computational approximations of game-theoretic techniques or computationally tractable assumptions. One example of this approach is the work presented by Faratin et al. [88–90]. The model describes heuristic decision functions for evaluating and generating offers in multi-attribute negotiation. Such heuristics have been widely used in several areas and complement multiple frameworks for multi-issue negotiation. The heuristics and methods are described below.

2.4.1 Negotiation Decision Functions

Faratin et al. [90] studied strategic negotiation between autonomous computational agents and develop a formal model of reasoning to address the coordination problem. They defined a number of heuristic approaches for generating counter-offers in a two-party negotiation [88]. Such heuristics receive the name of *tactics* and use a single criterion (time, resources, behaviour, etc.) to generate new values for each issue in the negotiation set. The following families of tactics for counter-offer generation were developed:

1. **Time-dependent tactics.** These decision functions use time to decide the value of a counter-offer;
2. **Resource-dependent tactics.** These tactics generate counter-offers depending on resource levels of the agent;
3. **Behaviour-dependent or Imitative tactics.** These tactics consider the behaviour of the negotiation opponent to compute counter-offers.

In a negotiation decision function, the domain is one of the criterion listed above and the range is the set of values for the negotiation issue. But it is possible to model an offer using more than one criterion with a weighted combination of different tactics covering the set of criteria. If an agent decides to propose an offer based on the remaining time of negotiation and how a particular resource like bandwidth is being consumed, it can use two tactics: one from the time-dependent family and one from the resource-dependent family. The information considered belongs to the agent which is making the proposal. However, a tactic that takes into account the behaviour of an opponent is only applicable when the agent has sufficiently information about it. In particular, time-dependent tactics are a type of resource dependent tactics in which the only resource considered is time. In this scenario, the final value for the issue under negotiation will be the weighted combination of the two values generated by each tactic function.
The following gives a more complete description of each family of tactics.

2.4.1.1 Time-Dependent Tactics

The time elapsed in the negotiation is what conducts the values of the negotiation issues. The more time has passed, the more pressure is induced and faster concessions are possible [91]. Time-dependent functions consist of the Boulware, Linear, and Conceder tactics that determine the amount of concession for the offer depending on the remaining negotiation time. Let $j \in \{1,2,...,m\}$ be an issue under negotiation, then the value proposed by agent a at time t is giving by the following equation:

$$x^t_{a\to b}[j] = \begin{cases} \min^a_j + \alpha^a_j(t)(\max^a_j - \min^a_j) & \text{if } a\text{'s utility decreases with issue j} \\ \min^a_j + (1-\alpha^a_j(t))(\max^a_j - \min^a_j) & \text{if } a\text{'s utility increases with issue j} \end{cases} \tag{2.2}$$

where $\min^a_j$ and $\max^a_j$ define the interval of acceptable values for issue j of agent a. There are many ways of defining the function that depends on time $\alpha^a_j(t)$ such that $0 \leq \alpha^a_j(t) \leq 1$, $\alpha^a_j(0) = k^a_j$ and $\alpha^a_j(t^a_{max}) = 1$. Then, the offer will always be between the value range $(\min^a_j, \max^a_j)$ starting with the constant k^a_j at $t = 0$ and ending with the reservation value when the deadline t^a_{max} is reached. From the equations above and range $(\min^a_j, \max^a_j)$, it can be seen that if the utility function is monotonically increasing then the reservation value is $\min^a_j$; and if it is decreasing, the reservation value is $\max^a_j$. The function $\alpha^a_j(t)$ can be defined to generate different types of time-dependent tactics. For example, a polynomial function parameterised by $\beta \in \mathbb{R}^+_n$ could be used as follows:

$$\alpha^a_j(t) = k^a_j + (1-k^a_j)(\min(t, t^a_{max})/t^a_{max})^{1/\beta.} \tag{2.3}$$

This function represents an infinite number of possible tactics by varying the value of β. β determines the convexity degree (see Figure 2.11) of the curve. At a higher β, the agent is characterised by a more conceder behaviour (such tactic is called *Conceder*) and the offer rapidly changes to the reservation value, while at a lower β the agent maintains its initial proposal until it almost approaches the deadline (such tactic is known as *Boulware* [91, 92]).

2.4.1.2 Resource-Dependent Tactics

These tactics are similar to time-dependent tactics, however, unlike time, other resources may have different patterns of usage. Resource-dependent tactics are modelled by using the same functions as in time-dependent ones, but they consider one of the following variations:

1. The value of the deadline t^a_{max} is dynamic and represents a heuristic on how many resources are in the negotiation set. There is an increasing urgent of agreement with diminishing levels of resource. In the application domain of this work, a resource might

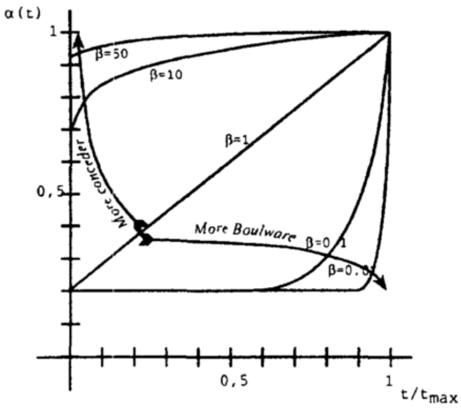

Figure 2.11: Polynomial function for the computation of $\alpha(t)$ (taken from [90]).

be energy or communication bandwidth. The greater the amount of bandwidth or energy, the lower the pressure on the agent to reach an agreement. While the decrease of both resources would add pressure to the bargaining. Then the value t^a_{max} would vary according to remaining bandwidth and energy.

2. Modelling the function α as an estimation of the amount of a particular resource. This can be modelled in a way that the agent becomes more conciliatory as the quantity of resource diminishes. The maximum value the concession can reach is the reservation value for the issue(s) under negotiation with the scarcity of resource(s). When there are enough resources, the agents assume a more Boulware behaviour. Formally, this can be modelled by the following α function:

$$\alpha^a_j(t) = k^a_j + (1 - k^a_j)e^{-resource^a(t)} \tag{2.4}$$

where the function resource$^a(t)$ returns the quantity of the resource at time t for agent a.

2.4.1.3 Behaviour-Dependent Tactics

With this type of tactics, the agent behaviour adapts to that of its opponent. There are several tactics in this type of family and the main difference between them is the type of imitation they perform. Given a sequence of offers $(\ldots, x^{t_{n-2\delta}}_{b\to a}, x^{t_{n-2\delta+1}}_{a\to b}, x^{t_{n-2\delta+2}}_{b\to a}, \ldots, x^{t_{n-2}}_{b\to a}, x^{t_{n-1}}_{a\to b}, x^{t_n}_{b\to a})$, with $\delta \geq 1$, the following families of tactics are possible:

1. **Relative Tit-For-Tat.** The type of imitation these tactics perform is proportional. The agent reproduces, in percentage terms, its opponent's behaviour $\delta \geq 1$ steps ago. The condition to apply this tactic is $n \geq 2\delta$. The following function calculates the counter-offer:

$$x^{t_{n+1}}_{a\to b}[j] = \min\left(\max\left(\frac{x^{t_{n-2\delta}}_{b\to a}[j]}{x^{t_{n-2\delta+2}}_{b\to a}[j]} x^{t_{n-1}}_{a\to b}[j], \min^a_j\right), \max^a_j\right). \tag{2.5}$$

2. **Random Absolute Tit-For-Tat.** With these tactics, the agent reproduces the exact behaviour of its opponent. If an agent proposes an offer with an increment of 2, the opponent counter-offers a value with the same increment. A component to add randomness to the behaviour by increasing or decreasing (depending on the parameter s) the value of an offer is also added. $\mathscr{M}$ is the maximum amount by which an agent can change its imitative behaviour. The condition of applicability is also $n \geq 2\delta$.

$$x_{a\to b}^{t_{n+1}}[j] = \min\left(\max\left(x_{a\to b}^{t_{n-1}}[j] + \left(x_{b\to a}^{t_{n-2\delta}}[j] - x_{b\to a}^{t_{n-2\delta+2}}[j]\right) + (-1)^s \mathscr{R}(\mathscr{M}), \min_j^a\right), \max_j^a\right). \tag{2.6}$$

where

$$s = \begin{cases} 0 & \text{if } a\text{'s utility decreases with issue j} \\ 1 & \text{if } a\text{'s utility increases with issue j.} \end{cases} \tag{2.7}$$

and $\mathscr{R}(\mathscr{M})$ is a function that returns a random integer from the interval $[0, \mathscr{M}]$.

3. **Averaged Tit-For-Tat.** With this type of tactics, an agent computes the average of percentages of changes in a window of size $\gamma \geq 1$ of its opponent's history when computing its counter-offer. When $\gamma = 1$, this is equivalent to relative Tit-For-Tat tactic with $\delta = 1$. The condition of applicability for this tactic is $n \geq 2\gamma$.

$$x_{a\to b}^{t_{n+1}}[j] = \min\left(\max\left(\frac{x_{b\to a}^{t_{n-2\gamma}}[j]}{x_{b\to a}^{t_n}[j]} x_{a\to b}^{t_{n-1}}[j], \min_j^a\right), \max_j^a\right). \tag{2.8}$$

2.4.2 Heuristic Methods for Automated Multi-issue Negotiation

Although the majority of models on non-mediated negotiation consider agents with either full information or probabilistic beliefs about the opponents, some existing approaches assume partial knowledge of the opponent's preference and avoid high computational demands through the use of tractable heuristics. This section reviews the bilateral negotiation case and focuses on existing techniques for automated multi-issue negotiation. Specifically, in a bargaining model, heuristic methods are examined here for optimising the utility of self-interested agents in a distributive negotiation setting under incomplete information. Therefore, the strategies are based on information such as the details of the domain and the opponent's characteristics.

Besides research on negotiation decision functions, work by Faratin studies a tradeoff negotiation approach as an alternative for offer generation [93]. This allows agents to exploit tradeoffs among different issues, where "Win-Win" opportunities are feasible. The addition of multiple criteria, however, may require complex computations to be performed involving multi-objective optimisation problems [94, 95]. Figure 2.12 represents a description of tradeoff approaches. Two indifferences curves I_1 and I_2 are depicted over a proposal of two issues for an agent and its opponent, respectively. Let offer P_1 be the initial proposal of one agent to another, and proposal P_2 its counter-offer. When one agent makes a tradeoff, it keeps the same utility from offer P_1 to offer P_2 but gets closer to the indifference curve of its opponent ($d_2 < d_1$). In other words, the agent selects an offer in close proximity to the preferences of its opponent without modifying its preferred utility.

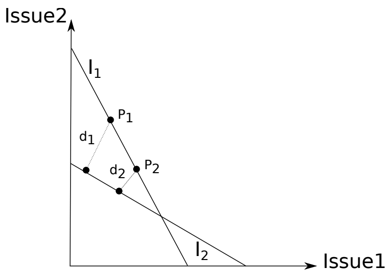

Figure 2.12: Example of tradeoff strategy.

Work in [96, 97] introduces an alternative approach with fuzzy reasoning techniques to model a framework for multi-issue negotiation. The main target in this solution is to achieve fairness in a semi-competitive environment, where agents look to maximise their own payoff. With fuzzy constraints, agents can relax their priority levels and accomplish appropriate trade-offs between them. The negotiation model in [98] considers settings with incomplete information to build concession strategies, but still assumes agents with partial knowledge about the opponent's preferences. Similarly, the strategic models in [99, 100] assume agents with prior knowledge about the structure of the opponent's utility function. The agents use bayesian rules and depth-limited combinatorial search to generate offers that approximately optimise their expected utility based on their current knowledge about the opponent's type. The rules update the agent's knowledge according to the history of negotiations and opponent's presumed strategic behaviour. The computation, however, can be expensive when the search and reasoning depths are large [101].

Yushu [102] uses a heuristic approach to design agents aware of the competitiveness and time pressure of the negotiation without any knowledge of the opponent's model. The information is employed to approximate the target utility using conservative concession strategies. Other bidding strategies for multi-issue negotiation under incomplete information are described in [103–105].

Work in [106, 107] proposes an offer generation technique for automated multi-issue negotiation with no information about the opponent's utility function using an alternating projection strategy. In this regard, several works in different domains have employed the alternating projection strategy to develop their offer generation mechanism [108–112].

The goal of this approach is to design strategies for generating offers for agents with no available information about the others and lead the negotiation process to an acceptable agreement for all the participants involved. Faratin et al. introduced the idea of choosing an offer similar to the opponent's preferences based on the existence of a fuzzy similarity function [93]. However, that approach requires a similarity function that is defined for every issue of the negotiation, which makes the mechanism domain-dependent and useful only with additive scoring functions. In contrast, works in [113–116] design an offer generation strategy, where an agent calculates offers close to the opponent's bids that match its own utility level without any additional similarity-based mechanism or information on the opponent's model. Basically, this type of strategies defines the negotiation model between two agents as follows:

- Agents compute a target utility, which represents their desired payoff to reach an agreement. A time-dependent concession strategy is usually used to determine the target utility at each round of negotiation, for example:

$$u^k = 1 + (V_{res} - 1) \times (\frac{k}{k_{max}}^{\alpha}) \tag{2.9}$$

  where k is the round of the negotiation, k_{max} is the deadline, V_{res} is the minimum acceptable utility or reserved utility value and α is the concession rate.

- Suppose an agent receives an offer that does not meet its expectations. In generating a counter-offer, the agent then has to accomplish two objectives: (1) an acceptable utility for itself, as defined in the step above and (2) an acceptable utility for its opponent in order to reach a final deal that satisfies both preferences. The agent computes the counter-offer as follows:

$$o^{k+1} = \underset{o \in C}{\arg\max} \ \{sim\{o, o_b\}\} \tag{2.10}$$

Where C is the iso-utility curve of the agent, that specifies all the offers that represent the same utility u^k. Then, the goal is to maximise the similarity between o and o_b. In the negotiation, the agents following this framework make compromises by moving their proposals towards each other. In settings with private information about the parties, each agent's proposal can be directly used to lead its opponent's counter-offer towards a final agreement. As it is described, the only information used by the agent to compute a counter-offer for its opponent is a previously received offer, o_b. In many cases, this offer o_b corresponds to the last offer made by the opponent or the best proposal exchanged by the opponent registered during the negotiation process.

The generation of offers is an important mechanism in automated negotiation, where a negotiating agent needs to select deals close to the opponent preference within the desirable benefit an agent wants to achieve. With this mechanism, agents increase their chance of reaching agreements in a finite negotiation time. The simplicity and efficiency guaranteed by these strategies are suitable for the domain of energy sharing in WSNs, where agents can have different preferences regarding the amount of energy at each time of cooperation. Basically, the proposing strategy is applicable to domains that have the following characteristics:

- There is no information available about the opponent's utility function.

- Multi-issue negotiations are considered and the domain is not limited to linearly additive utility functions.
- The issues are simultaneously negotiated.
- The domain requires simple, tractable and good solutions.

A multi-issue negotiation is more complex and challenging than a single-issue negotiation because the agreement space is n-dimensional, which also allows agents to find win-win outcomes. Every time an agent plans to concede, it needs to first decide the direction of concession. With this offer generation approach, the decision on the concession direction follows the opponent's preference, addressing the possibility of finding an acceptable deal in finite time convergence. At the same time, the agent determines how much utility to concede in each period according to a determined tactic. Moreover, this strategy has proved to approximate Pareto-efficient bargaining solutions [113–117].

All these heuristic methods make use of the opponent's model (relying on partial information), domain and offer information (independent of the opponent's knowledge) available to the agents to infer relative information that may lead to better negotiation agreements. In conclusion, heuristic mechanisms are a feasible approach for bilateral bargaining, where computationally tractable assumptions approximate the agents' decision-making during negotiation. These heuristics face the demands imposed by high computational resources and the complete availability and quality of information required for optimal negotiation strategies. The main drawback, however, of these methods is the consideration of extensive evaluation to measure the performance and get conclusions of the heuristic in the domain context [74].

2.5 Energy Management in Energy Harvesting WSNs and Related Work

An energy harvesting architecture can be classified into two types: (1) systems where energy received from the harvesting device is directly converted into electric energy to power the sensor nodes (no energy buffer is included), and (2) systems where the converted energy is managed to supply a sensor node load and use the energy storage device to save the generated energy that may not be used instantaneously. The last category is the one considered in this thesis. The work in energy management studied here includes research from the perspective of optimisation regarding energy use in rechargeable sensor networks. Figure 2.13 shows a general model for energy management methods in a sensor node with energy harvesting capability.

While a deterministic metric such as residual energy level is good enough to monitor the energy availability in the case of battery-powered WSNs, an additional source characterisation is required for a harvesting sensor node. Such characterisation corresponds to the energy input from the environment, which must be controlled while the battery dynamics are also analysed.

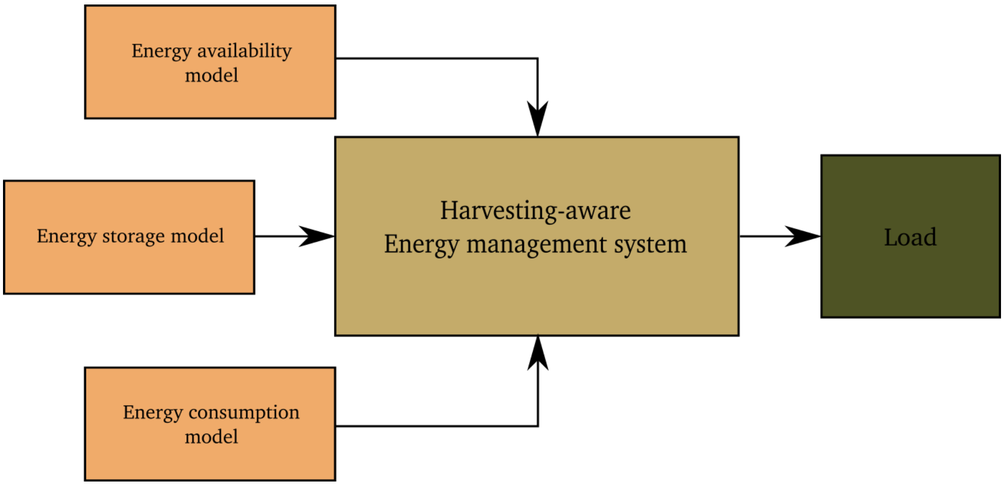

Figure 2.13: General system model for an energy harvesting wireless sensor node studied in this thesis.

Energy harvesting systems use ambient energy as the main source of power supply. With a potentially infinite amount of energy available to the sensor nodes, an energy harvesting WSN can remain powered for a long period of time. Hence, the power management objective of maximising node lifetime is not the central design issue in these systems. Instead, the optimisation of energy use to allow the system to operate perennially is the main concern in energy harvesting sensor networks. This target for energy optimisation is known as energy neutrality. For example, if the expected energy harvested matches the amount of expected energy consumption during the same time interval, it can be said that a sensor node satisfies the energy neutrality condition.

Energy harvesting devices can vary between solar cells, wind turbines and piezo-electric cells. The energy source dictates the total energy provision available for use. With the aim of ensuring energy neutrality sensing systems and control the spatio-temporal profile of energy sources, researchers in the area have proposed two different energy management schemes: efficient energy allocation and adaptive network parameter algorithms. Energy management schemes are designed to provide networks with efficient use of harvested energy. The energy management mechanisms analysed in this thesis use three design elements; energy provision, energy storage and energy consumption to harmonise a varying energy supply with a fixed demand (load).

The authors in [18] consider the problem of optimal power management for sensor nodes and present a linear programming model for the adaptation of duty cycles to provide energy-neutral networks. Similarly, work in [19] develops an energy allocation algorithm for the optimal use of energy harvested in WSNs, but the objective here is to minimise the variation in allocated energy over a period of time while satisfying the energy neutrality condition. Both algorithms focus on the optimisation of power management when nodes are harvesting-aware but differ greatly in their design. The former method adjusts a network parameter as the duty cycle to satisfy the condition of continuous operation while the latter employs the maximum amount of energy consumption to determine the amounts of energy to be allocated over a certain time span.

Various algorithms have been proposed to perform optimal energy allocation [19–21, 23, 24]. The objective function of these models should satisfy the following requirement: the energy harvested during a period of time should be fully utilised while meeting the energy-neutrality constraint, and the variation in allocated energy over time should be minimised. However, these proposals have the important limitation that the gathered energy considered for allocation is limited to the surroundings of one network domain. Consequently, even when additional energy can be harvested from an ambient energy source, the energy collected remains a limited resource due to temporal and spatial variations [25]. Moreover, some energy allocation methods [21, 23, 26–28] consider infinite energy storage or ideal mechanisms to store harvested energy. An ideal energy buffer is defined as a device that has unlimited capacity to save energy and does not have any inefficiency in charging and discharging, or energy leak over time [18].

A reasonable notion of limited energy buffers is presented by several allocation algorithms [22, 29–31]. However, these models include more complex operations based on dynamic programming or Markov decision processes (MDPs). These solutions lead to increased computational costs and significant running times, compared to linear or convex optimisation models. Such complexity requires the use of sophisticated devices to perform these techniques, therefore a centralised solver with sufficient computational capability might need to be employed.

Several adaptive duty cycling schemes [32, 33, 118] have been proposed to adjust the activity of a sensor node according to its harvesting opportunity to balance its energy consumption. Adaptive algorithms address the spatio-temporal variation of ambient energy sources by also optimising data sampling and routing in order to deliver effective power management [34, 119]. Efficient utilisation of energy scavenging by optimal energy allocation and packet routing decision is proposed in [23, 24]. Other energy-neutral designs exploit the spatial variation of energy harvesters and distribute load using adaptive opportunistic routing protocols [39]. Meanwhile, the adjustment of parameters as the sampling rate is preferred in other solutions with the same goal of modelling energy-neutral systems [34–37].

While these approaches work well with high temporal variability of ambient energy, they can only optimise the network's performance in terms of the energy collected by sensors under the control of a single authority. Therefore, if the entire network is unable to harvest energy (due to ambient conditions or obstacles in the environment), no solution is enough. Some adaptive energy allocation schemes require a centralised controller (usually the sink node) to distribute the adjustment of network parameters. In addition to these limitations, these algorithms may lead to deficient data gathering. Duty cycling and sampling rate regulations are executed to control the energy consumption of a node for efficient energy expenditure. With this aim, adaptive techniques dynamically decide the node's operation (a sampling rate or duty cycle) to increase its activity when there is ample energy and speed it up when energy supply is scarce. These adjustments may result in the collection of unnecessary data or in the loss of useful information.

The heterogeneity in the characteristics of each WSN such as battery capacity, energy consumption, nodes locations and energy harvesting can be explored to jointly maximise the energy utilisation of the co-located networks. This heterogeneity should be clear to quantify the gains from cooperation when participants are different. However, there are no previous studies for this domain of cooperation. To implement the vision of WSN cooperation, this thesis needs to include an interdependent multi-issue negotiation framework to determine the distribution of gains from cooperation. With this, networks can evaluate whether or not cooperation can take place.

The specific scenario of cooperation studied here is a novel setup in the area of WSNs. In this scenario, nodes require to optimise their use of harvested energy to fulfil their energy requirements by collaboration as much as possible. From this perspective, the authors of [120] propose an approach that facilitates the energy exchange between homes equipped with renewable energy technology and storage in remote communities to achieve efficient energy management. The linear programming framework modelled in [120] for connecting agents is useful to model the cooperation between networks. Therefore, the model presented in this thesis applies the linear utility function designed in [121] to measure the bargainer's preferences (the node's preferences) for the energy flow allocation in this domain.

2.6 Selection of Negotiation Partner

The partner selection problem is connected to the nature of the WSN. When multiple nodes from different networks share the same neighbourhood, an agent seeks to provide an energy allocation from a number of potential partners. Traditionally, an agent engages in multiple concurrent bilateral negotiations for the acquisition of a good or service. In particular, the case in which a node is looking for a single node from a number of available nodes in its environment is considered here. By bargaining simultaneously with these agents and making partial agreements with them, an agent can reach good deals in an efficient manner. However, due to the memory and processing capability limitations of an agent in this domain, it cannot afford to evaluate the proposals of too many counterparts. In relatively small environments, or those with more powerful computing capabilities, a self-interested agent can reach its most preferred deal by negotiating with all agents that offer cooperation. In this situation, an agent is also capable of implementing robust negotiating strategies that result in efficient agreements even when there are dynamic changes in the environment. However, this approach may not be reasonable in domains with limited computation or restricted communication bandwidth. In an open dynamic domain, a negotiation may lead to a communication overhead when coordinated over a large number of agents, degrading a network's performance. It is always preferable to start a negotiation which is likely to succeed and reach a better agreement. Therefore, an agent should be able to anticipate the best potential partner with which to start a negotiation for the practical realisation of a solution in this domain. Indeed, the main goal for an agent is to find the most prospective negotiation partner that maximises its energy allocation.

In [47], a motivation-based mechanism maps goals and issues to motivations and uses the history of candidates' performance to select those that have the most beneficial effects in terms of current motivational needs. The selection problem of agents for negotiation has also been studied in [48]. The negotiation outcome and its equilibrium are analysed in terms of the amount of information that is known about the opponent's parameters. The results reported are useful for decision making in situations where an agent has the option to select a partner on the basis of the information state about its opponents. In [49] the authors propose a framework for automated negotiation based on negotiation profiles. Each agent gathers information during the negotiations and stores it in the associated profile: the preference profile, keeps the agent negotiation strategy, the partner cooperation profile, records the agent interaction with the other agents in the environment, and the group-of-partners negotiation profile, stores the profiles of several negotiation partners. The agent is then able to construct a set of rules which allows it to anticipate both the outcome and the best potential partner with which to start a negotiation. A central facilitator is responsible for registering new agents and informing others about it. The problem of partner selection in [50] is analysed using a possibilistic case-based decision model. Their solution provides the decision theoretical basis to predict the possibility of successful negotiation with other agents using small historical data about past negotiation behaviour and the derived qualitative expected utility for a specific situation. Accordingly, they keep a record of past negotiations to model the negotiation behaviour of the opponents and be able to predict it in the future.

As shown by previous research, the record of past negotiations is essential to choose the negotiation partner among a set of candidates. These make sense in devices equipped with advanced processors and large memory capacity. In fact, the design of automated negotiation is highly sensitive to the domain in which the interactions take place. The networks of this work are resource constrained systems that discover each other opportunistically and have no information about their neighbours. Moreover, the widespread use of WSNs predicted in the future and the increasing likelihood of different WSNs deployed in the same place demand a proper policy to aid an agent on the decision-making process of the most prospective partner. In this environment, the most promising partner for negotiation is evaluated in terms of agreements on energy cooperation, where the position of the nodes and the orientation of their energy sources strongly impact the energy harvested. For this reason, even if two nodes are geographically close, their harvesting rates may vary significantly [40, 122].

Under complete uncertainty, with no prior information about the agents in the neighbourhood, this research motivates the incorporation of some form of reinforcement learning (RL) into the partner selection problem. RL deals with decision making via interaction and feedback, or in other words, learning to achieve some goal by trial and error. Within these situations, an agent is built to explore an unknown environment and take actions to interact with it. The following section outlines relevant related work on the application of reinforcement learning in WSNs.

2.7 Reinforcement Learning as an Optimization Solution for WSNs

Reinforcement learning is a technique of machine learning that involves the perception of learning using trial-and-error movements (see Figure 2.14). It is particularly suitable for dynamic environments such as WSNs, where the state of current conditions can vary over time. RL is used to build automated agents that learn through the interaction with the world. By performing actions and adapting future decision-making based on the observed consequences of those actions, an agent can learn an optimal policy to optimise a particular objective.

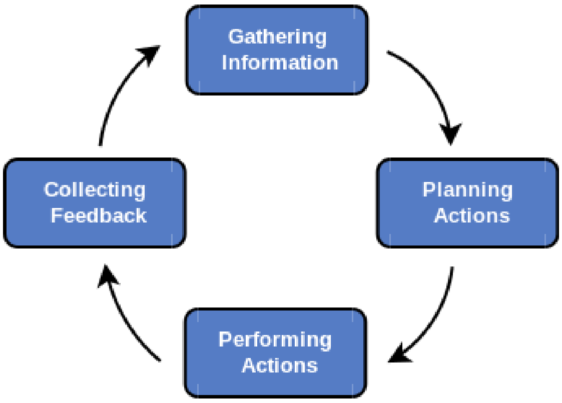

Figure 2.14: A general scheme of a four stage cognition cycle: (1) gathering information, (2) planning actions, (3) acting, (4) collecting feedback (reproduced from [123]).

Formally, the basic RL model consists of:

- a discrete set of environment states $(S = s_1, s_2, ..., s_m)$;
- a discrete set of actions $(A = a_1, a_2, ..., a_n)$ that can be chosen at each state;
- rules of transitioning between states;
- transition probabilities; and
- rules that determine the reward of a transition;

An agent choosing an action $a \in A$ causes a transition from state s to a different state s' according to a certain probability. The representation $P(s'|s,a)$ indicates the probability of entering a state s' when an action a is taken in the state s. The agent receives a reward or punishment $R(s,a,s')$ for chosing the action a in state s that lead to state s'. The rewards can be predefined values or defined by rules. The goal of RL at some time t is to find a policy that maximizes the acquired sum of rewards over time, what is called a state-action function (Q function). This is based on Bellman's principle of optimality, represented by the following equation:

$$Q(s,a) = R(s,a) + \gamma \times \sum_{s'} P(s'|s,a) \times \max_{a'} Q(s',a') \tag{2.11}$$

The Q function designates a value to each state/action pair. $R(s,a)$ is the reward gained for executing action a at state s, and the argument to the right represents the maximum expected future reward. The factor γ is the discount value used to penalise the repetition of a decision (state/action pair) over time.

With knowledge of a Q value, the agent does not need a model (reward functions and transition probabilities) to decide how to act: it can simply keep acting based on the action with the maximum Q-value in the current state, i.e. an agent is able to take expectations of Q-values using just experienced data. This is known as model-free RL. The most well-known model-free RL technique is Q-learning. In the WSN context, Q-learning has been extensively used in routing problems [124, 125]. In fact, most of the proposed reinforcement learning based approaches solely focus on solving the WSN routing problem. Solving such a problem is found to be NP-hard. However, similarly to this research, the application of reinforcement learning in routing seeks to predict the full path quality between nodes by reducing the complexity of a routing problem considering only neighbouring nodes' information [126]. Each node independently performs the routing procedures to decide the minimum cost path, which leads to near-optimal routing decisions with a very low computational complexity. In [127] the authors apply swarm intelligence-based algorithms to achieve distributed path decision making for adaptive routing.

The works in [58, 123] use RL to manage cross-network optimisation problems. The Least-Squares Policy Iteration (LSPI) algorithm is used as the reasoning method to find the optimal set of network services in each WSN node that would be beneficial for their performance when they cooperate. A central and powerful negotiation engine is assumed to continuously collect information about the system measurements and environmental states. The engine computes the configurations for each participating network so that the activation of the corresponding services positively influences the performance of each system. Along with the assumption of a centralised decision maker, their paradigm referred to as symbiotic networking contemplates the integration of different networks from their design. Similarly to their initial proposal [10], this work motivates the use of an automated solution to enable cooperation between networks. However, the cooperation develops opportunistically and is assumed to occur directly between nodes, without relying on a trusted authority. This requirement faces the challenge of dealing with constrained nodes instead of enhanced devices. Consequently, the methods require to be automated but suitable for the limited capacity of typical devices in IoT. Besides, they need to consider the rapid response required in some opportunistic cooperative domains (e.g. Emergency response).

The use of LSPI is reproduced in [128] to enable a node to learn an optimal routing scheme with multiple optimisation goals among the maximisation of its network lifetime. Similarly, work in [129] proposes a routing policy conditioned by the message importance that includes the selection of paths with the highest delivery rate learned over the previous routing experiences. The underlying approach in this case is based on Q-Learning. Although the space of options is simplified in the routing domain, these techniques need to consider the set of state-action pairs to find an optimal action-selection policy. As a result, the computational complexity of the algorithms increases as the dimensionality of the problem proportional to the state representation grows.

The drawbacks of high computation complexity and large memory requirement in comparison to more sophisticated learning algorithms are reduced with multi-armed bandit (MAB) learning. The space of options in MABs is characterised only by the set of the agent's actions. The MAB model is commonly used in the online learning literature for solving resource allocation problems. One solution in the context of WSNs is multi-armed bandit based energy management (MAB/EM) [130]. MAB/EM is a power management technique that enables an agent to adapt to the environmental changes while maximising the total amount of information collected over a period of time. In MAB/EM, the energy of an agent is intelligently allocated to the tasks of sampling, reception, and transmission of data, as the agent learns which combinations optimise its performance in long-term information collection. The allocation problem is also solved in [131] by using MAB algorithms to make efficient use of the radio spectrum and avoid collision between cognitive nodes. In the model, the nodes are not aware of the communication medium conditions, and they have to estimate the channel's availability by exploring and learning.

To balance exploration with exploitation, the partner selection problem is modelled in Chapter 5 as an adversarial MAB problem. This thesis evaluates the performance of several algorithms on partner selection through practical scenarios in WSNs. The goal is to have an accurate online estimation method of whether a particular policy will work well in practice. In this regard, there are no prior results for the reward maximisation on partner selection. The problem has been studied from different perspectives as described in Section 2.6.

2.8 Summary and Discussion

Cooperation among IoT networks is an important research topic considering the vast deployment of WSNs envisaged with the progress of IoT. Previous research has shown the benefits of having shared nodes between distinct networks and the means to enable it. A link-layer protocol called OI-MAC makes possible this sharing by supporting direct interconnection between multiple network domains. The chapter began by describing the communication technology that allows cross-boundary transmission. Along with the conceptual framework to construct ODI, practical implementations have also been explored. Since the feasibility of ODI has been confirmed, the vision of multi-domain WSN cooperation is reachable. The cooperation will not only bring benefits at the information-sharing level [3, 5] but may also be favourable in terms of optimal energy use, QoS guarantees, load balancing, interference avoidance, low latency or high reliability [6, 10, 14, 42].

Although the idea of cooperation seems straightforward, many open challenges remain, mostly due to the independent behaviour and heterogeneity of a WSN. Even if the heterogeneity in communication layers is already tackled, novel proposals in the coordination among distinct WSNs are still left out. In many applications, the networks usually represent different entities or are managed by different owners. Thus, their rational action to decide whether to cooperate or not depends on their conditions. In this scenario, on their own energy harvested, load and battery. The purpose of the negotiation is to solve this kind of problem, where different techniques have been studied to model the interactions among agents [74].

In this direction, previous research has modelled the cooperation problem using a game-theoretic framework to identify possible equilibrium strategies. However, the studies in cooperative solutions assign the control of these findings to network administrators, shared nodes with sufficiently large power supply or centralised powerful devices. This is impractical with nodes of limited bandwidth, memory and power capacities. Therefore, the objective of this thesis is to bridge the gap between constrained nodes and how negotiation is performed in this domain. For evaluation, the use of game theory is included to serve as a benchmark for the decentralised negotiation approach. In this regard, Chapter 3 describes the system model and presents an evaluation of the axiomatic solution for energy cooperation.

This thesis applies multi-agent negotiation techniques and the perspective of AI for the resolution of conflicts in WSNs. An alternative approach based on heuristics is capable of reducing the complexity of the cooperation problem and represents a novel contribution in this domain of cooperative WSNs. Moreover, the incentive for optimal energy use has not been explored using a negotiation framework. This work has presented a wide range of approaches to ensure energy-neutral sensing systems, categorised in optimal energy allocation algorithms and adaptive parameter strategies. However, none of the existing proposals allows energy-neutral algorithms to leverage an area wider than its own network domain. Based on the analysis made in this chapter, Chapter 4 describes the heuristic negotiation framework used to model the negotiation process and its performance in comparison with the centralised solution.

Automated negotiation and specifically, the bargaining scheme presented in this work, is proposed as part of a cooperation strategy among self-interest agents as an effort to solve two basic problems, when to cooperate and how to cooperate. Conversely to a fixed set of rules to define the offers made by an agent, the main idea of this proposal is to provide more flexibility in the negotiation process, such as to allow a negotiator to adopt a behaviour in the face of a limited resource. Later in Chapter 4 the limited resource is described as a number of rounds, but it can be any resource engaged in the negotiation [90].

The scenario of negotiation-based cooperation describes nodes that discover co-located devices with conflicting interests and a desire to cooperate in energy sharing. The problem of partner selection becomes natural in this context where multiple nodes share the same location. Thus, existing partner selection strategies were also discussed and analysed in this chapter. Current solutions assume knowledge about the opponents, therefore existing methods map requirements to motivations or consider the history of previous interactions to select the most promising partner for negotiation. However, in situations under uncertainty as the ones presented in opportunistic encounters, the problem is more likely to involve self-learning techniques. Chapter 5 presents the research carried out in order to address this challenge.

In conclusion, existing cooperative designs have not proposed solutions to enable energy sharing between networks, capable of performing long-term efficient energy management and keeping energy-neutrality operation. A heuristic approach attempts to overcome the limitations of the game theory mechanisms. Instead of searching for the optimal solution, agents try to find a sufficient, near-optimal outcome by reducing the search space to decrease the high computational complexity. In the absence of an optimal and central controller, results demonstrate the efficiency of a heuristic negotiation solution. Therefore, automated negotiation models based on heuristic approaches need an intensive evaluation, elaborated through comprehensive experiments and empirical analysis. With this in mind, the following chapters present the optimal solution and introduce the heuristic model to analyse the efficiency of cross-network power management using the proposed bargaining approach.

Chapter 3

Opportunistic Energy Negotiation: System Model and Cooperative Approach

Previous research about resource sharing described in Section 2.2 certifies the benefits of establishing cooperation between co-located networks. However, energy-harvesting WSNs have been left out of this context. As a result, the cooperation incentive for optimal energy use has not been previously considered. The main motivation of this work is to conduct power management by leveraging an area wider than the boundaries of one domain and enable opportunistic energy transfer across multiple networks. Designing such systems requires efficient energy management methods. To this end, first, an optimal energy allocation scheme is described in Section 3.2. Subsequently, a utility-based energy allocation algorithm is derived to address the problem of utility maximisation for a node with energy replenishment (discussed in Section 3.3). The cooperative approach is analysed in Section 3.4 by first describing the model to transfer energy between two nodes. Then, the experimental setup and results based on this model are presented in Section 3.5. Since networks are not known in design time because each is independent, a negotiation approach is suggested to deal with their different energy profiles. Accordingly, this chapter shows the system model to support cooperative energy allocation and results of the adopted bargaining solution. Typically in game theory, a solution is characterised by desirable properties like Pareto optimality or fairness. The bargaining solution used to compute the optimal energy flow satisfies the efficiency axiom. This method is used as a benchmark to evaluate the outcome obtained by the framework employed in the heuristic approach.

3.1 Opportunistic Direct Cooperation Methodology

As stated before, the negotiation-based cooperation approach aims to optimise a network's power management through negotiation based on cross-network optimisations and self-organising capabilities. Fig. 3.1 gives a general overview of the establishment of cooperation following the proposed methodology. The methodology to support cooperation between distinct networks consists of the following 5 phases, which are studied through these chapters in more detail:

1. First, a node envisages an insufficient energy allocation.
2. The node discovers nearby devices with the same interest to start a cooperation.
3. The node selects a negotiation partner to initiate a bilateral negotiation.
4. The negotiation proceeds between nodes to find a suitable agreement that best satisfies their negotiation objectives.
5. Once the deal is reached, the cooperation proceeds.

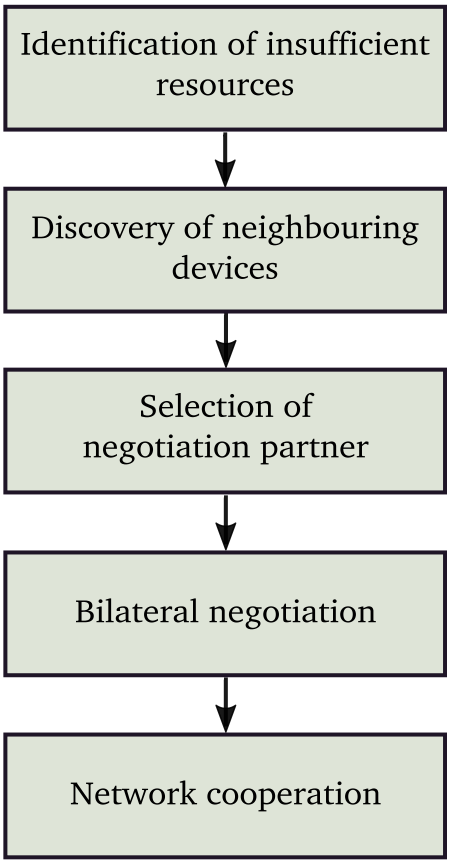

Figure 3.1: The 5 phases of the methodology proposed in this thesis to establish direct cooperation between distinct networks.

The methodology is applied here to establish opportunistic energy negotiation between co-located networks, also identified as OEN through this thesis. The next section describes the system model and considerations made for this domain. In this chapter, it is assumed that networks can reach cooperative agreements by the intervention of a mediator. The mediator finds an energy agreement that maximises the product of the participants' utilities. In order to develop all the phases, this chapter starts with the energy allocation algorithm to accomplish the first step of the methodology. Chapter 4 describes the protocols used to discover neighbouring devices and realise the negotiation process for OEN, while Chapter 5 addresses the uncertainty, dynamism and diversity intrinsic to the domain for the selection of a negotiation partner.

3.2 System Model and Problem Formulation

This section describes the model and assumptions made concerning the characteristics of the sensor networks. The network model, the energy consumption model, the energy management model for the energy source and the formulation of the optimisation problem of energy allocation are presented here. Although there are several possible WSN deployment scenarios, this chapter addresses the problem of opportunistic energy negotiation with an initial simplified setting. The empirical evaluation assumes two sensor networks sharing the same physical location, and direct interconnection can be established between any pair of sensor nodes. Thus, this chapter studies the cooperation strategy between a pair of nodes. The main motivation in investigating negotiation applied in this domain is to observe the effects of cross-boundary energy transfer for the node's power management and this setting is suitable for that purpose.

3.2.1 Network Model

A set N of m energy-harvesting wireless sensor networks $N = \{N_1, N_2 \ldots, N_m\}$ that are under the administration of distinct authorities and deployed in the same area is considered (see Figure 3.2). Each independent network N_i, $1 \leq i \leq m$ has a different type of harvesting source (e.g. wind turbine or solar panel) and is formed by a set of unique sensor nodes $I_i = \{1, \ldots, j, \ldots, z\}$ and a sink.

The networks that are studied attempt to achieve energy-neutral operation. Thus, the nodes need to satisfy as much as possible their energy allocation scheme, considering their load and expectation of energy harvested. This thesis considers that sensor nodes can forecast this last information from historical data with high certainty. This assumption is reasonable according to the experimental studies reported in [52] to forecast sub-glacial movement directly affected by climatic changes. The study considers previously forecasted data to continuously reduce the uncertainty about this data and reach zero. The hardware features of the sensor device include a 32bit ARM Processor and 2GB microSD card for storing readings [132]. As for energy sources and its availability, models for its seasonal cycles are known and can be used to predict future energy opportunities [18].

Time is divided into discrete time slots $T = \{1, \ldots, n\}$ of equal duration L. Each time slot t is long enough to deliver all packets to the collector and take a decision about inter-network cooperation. A node is able to perform cross-network packet transmission through direct interconnection to nodes with overlapping radio range. The scenario consists of general WSN applications with

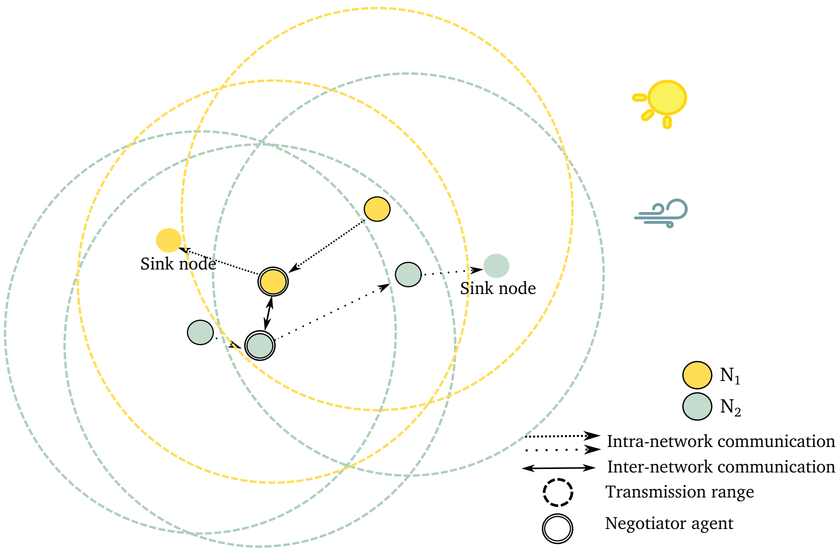

Figure 3.2: Two different WSNs (N_1 and N_2) deployed at the same area. A feasible scenario to execute Opportunistic Energy Negotiation.

randomly deployed nodes that periodically collect data from the environment and report these measurements to the sink, using multiple hops to traffic the packets.

Each network N_i is examined as a cooperative multi-agent system. Then, each node j in N_i is controlled by an agent denoted as i_j, $1 \leq i \leq m$, $j \in \mathbb{N}$. The agent has complete knowledge of all the relevant node's information, such as its neighbours, its energy profile variables: energy availability from the harvested source in each time slot t, including the availability in the future, and its energy consumption, its battery capacity and residual energy. In general, a node is an autonomous agent with advanced situational awareness of itself and its local neighbours (nodes in its own network).

3.2.2 Energy Consumption Model

The nodes in N_i usually operate unattended in a collaborative manner to perform some tasks. Such tasks include sampling, reception, processing and transmission. Although the execution of these tasks consumes a measurable amount of energy, the power used up in processing and sampling can be ignored since the communication energy for reception and transmission is a dominant factor in most sensor platforms [133]. Thus, the total energy consumed by an agent is in terms of its radio transceiver's duty cycle. Furthermore, each agent i_j consists of an energy harvester unit and a rechargeable battery. The energy management model used to derive an agent's energy profile is described in the next subsection.

In this work, a simplified model of average power consumption is adopted as it is used in [134], where a sensor node's power consumption is determined by its duty cycle. Let $E^c_{i,j}(t)$ denote the energy consumed by radio communication of agent i_j in time slot t. The vector $\boldsymbol{E}^{\boldsymbol{c}}_{i,j} = (E^c_{i,j}(1),...,E^c_{i,j}(n)) : \boldsymbol{E}^{\boldsymbol{c}}_{i,j} \in \mathbb{R}^+$ denotes the energy consumed by agent i_j in n time slots. At any given time slot t of length L, the total energy $E^c_{i,j}$ that an agent i_j consumes can be calculated as:

$$E^c_{i,j}(t) = V \times \left[D \times I^{active} + (1-D) \times I^{sleep} \right] \times L \tag{3.1}$$

Then, the maximum energy an agent can spend at each time slot t is dependent on the average power consumption ruled by the duty cycle D, supplied voltage V, active mode current I^{active}, and sleep mode current I^{sleep}. D is set by the node's application, while I^{active},I^{sleep} and V can be known in advance using datasheet information.

3.2.3 Energy Management Model

The energy management model is built on the proposal made by [18]. The power management characterisation described for a sensor node considers a harvesting system with a non-ideal energy buffer, i.e. an energy buffer is not ideal when its capacity is fixed and its charging efficiency is strictly less than 1. Figure 3.3 illustrates an agent's model for a sensor node with energy harvesting capability and limited storage.

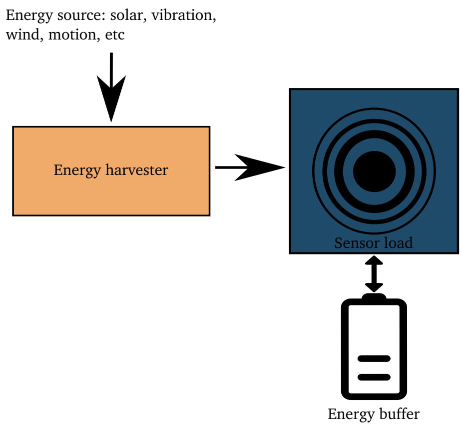

Figure 3.3: An agent-based model of a sensor node with energy harvesting capability and limited storage.

A rechargeable agent can harvest energy and reserve it in its battery for future use. However, as typically assumed, while the sink node has an unlimited power supply, each sensor node is conditioned to spatio-temporal variations of energy availability [18, 135]. Without loss of generality, it is assumed that the replenishment of energy occurs at the beginning of each time slot t and the agent stores the unused harvested energy in the battery immediately.

According to the time domain mentioned in 3.2.1, each period of energy harvesting can be divided into slots $T = (1,\dots,n)$ of equal duration L. The expected energy input during each slot t can be forecast from historical information with a high level of accuracy. Energy can then be allocated to each slot t. $E^c_{i,j}(t)$ and $E^{hrv}_{i,j}(t)$ are used to denote the energy profile variables for each time period. $E^c_{i,j}(t)$ is the maximum energy spent by a node to relay data or transmit it at time slot t defined in 3.2.2. While the amount of energy that can be generated by the harvesting unit in n time slots is defined as $\boldsymbol{E_{i,j}^{hrv}} = (E_{i,j}^{hrv}(1), ..., E_{i,j}^{hrv}(n)) : \boldsymbol{E_{i,j}^{hrv}} \in \mathbb{R}^+$. For example, if the harvesting period starts at 00:00 a.m and L is 1 hour, then $E_{i,j}^{hrv}(1)$ is the expectation for the amount of energy harvested during slot 1 (from 00:00 a.m. to 01:00 a.m.), $E_{i,j}^{hrv}(2)$ is the expectation of energy during slot 2 (01:00 a.m. to 02:00 a.m.), etc.

$B_{i,j}(t)$ is used to represent the residual battery energy at the beginning of slot t in agent i_j. Therefore, the battery energy left after the last slot of the energy harvesting period is defined as $B_{i,j}(n+1)$. Then $\boldsymbol{B}_{i,j} = (B_{i,j}(1), ..., B_{i,j}(n)) : \boldsymbol{B}_{i,j} \in \mathbb{R}^+$ denotes the battery level in n time periods. The cycle of the battery is represented as $B_{i,j}(n+1) = B_{i,j}(1)$. The battery is characterised by a limited energy capacity $B_{i,j}^{max}$ and charging efficiency e. The battery enables an agent to save and use energy over the entire period of n slots, which helps the agent to compute an energy allocation for each t. The energy allocation for each time slot t is denoted as $\boldsymbol{E_{i,j}^{alloc}} = (E_{i,j}^{alloc}(1), ..., E_{i,j}^{alloc}(n)) : \boldsymbol{E_{i,j}^{alloc}} \in \mathbb{R}^+$, which is the amount of energy allocated from the source (harvester or battery) to the energy consumed $E_{i,j}^{c}$ by the load of the sensor node.

The battery allows the agent to use the harvested energy efficiently by storing energy temporarily. When $E_{i,j}^{hrv}(t)$ is lower than $E_{i,j}^{c}(t)$, some of the energy used by the sensor node is discharged from the battery. $\boldsymbol{d} = (d(1), ..., d(n)) : \boldsymbol{d} \in \mathbb{R}^+$ represents the discharged energy amount of the battery. When $E_{i,j}^{hrv}(t)$ is higher than $E_{i,j}^{c}(t)$, all the energy used in the node is provided by the harvested source and the battery is charged with the excess as required, up to its maximum capacity. $\boldsymbol{c} = (c(1), ..., c(n)) : \boldsymbol{c} \in \mathbb{R}^+$ denotes the charge of the battery in n time slots. There are no discharge and charge amounts at the same time t. Any excess energy received at times when the energy buffer is full is discarded by the node. In other words, the energy that the node is unable to use or store at time slot t is wasted. This energy is denoted by $\boldsymbol{w}_{i,j} = (w_{i,j}(1), ..., w_{i,j}(n)) : \boldsymbol{w}_{i,j} \in \mathbb{R}^+$. Then, the energy used from the battery in any slot t can be calculated as:

$$B_{i,j}(t) - B_{i,j}(t+1) = d(t) - e \times c(t) \tag{3.2}$$

An opportunistic energy negotiation process is initiated when an agent's estimated energy level is not enough to maintain the next period. As a result, an agent foresees an insufficient energy allocation scheme given the expected energy and remaining battery level. Thus, at each period T the initial battery status $B_{i,j}(1)$ is equal to $e \times b$ where b is the energy level at $t = 1$.

A summary of notations used in this work is given in Table 3.1.

Symbol	Definition
T	The total period of energy harvesting process
t	The t-th slot, t = 1,2,...,n
n	The total number of time slots, t = 1,2,...,n
L	Duration of a time slot
i_j	Agent of node j in network i
$B_{i,j}(t)$	The battery level of agent i_j at the beginning of slot t
$B_{i,j}^{max}$	The maximum battery level of agent i_j
e	Battery efficiency
$d(t)$	The amount of energy that is discharged from the battery at time slot t
$c(t)$	The amount of energy that is charged to the battery at time slot t
$E_{i,j}^{hrv}(t)$	The amount of energy harvested by agent i_j at slot t
$E_{i,j}^{alloc}(t)$	The energy allocation for agent i_j at slot t
$w_{i,j}(t)$	The surplus energy of agent i_j
V	Voltage supplied to agent i_j
D	The duty cycle of agent i_j
I^{active}	Current consumed in active mode by agent i_j
I^{sleep}	Current consumed in sleep mode by agent i_j
$E_{i,j}^{c}(t)$	The total energy consumed by agent i_j at slot t

Table 3.1: Notations used in network model, energy consumption model and energy management model.

3.2.4 Optimal Energy Allocation Problem

The energy allocation model is based on the model developed by [121]. A linear function is modelled as [121] to design the objective function of each agent i_j. The agent first identifies its best performance given the constraints of its domain to determine the need for cooperation. If an insufficient energy allocation scheme is expected, an opportunity to initiate an opportunistic energy negotiation arises. Thus, to determine the agent's utility space, the following problem needs to be solved.

Since the current battery status only depends on the amount of energy harvested and consumed during previous slots, as shown in Equation 3.2, the energy allocation problem can be formulated as a linear program (LP). Then, the objective function is described as the total amount of energy consumption that is satisfied (i.e., energy allocation $E_{i,j}^{alloc}$) at period T. In other words, the overall utility for an agent is the sum of all satisfied loads in n time slots. Suppose the utility of an agent is represented by u, the following gives its definition:

$$Objective \quad \max u_{i,j} = \sum_{t=1}^{n} E_{i,j}^{alloc}(t) \tag{3.3}$$

That is, the aim is to maximise the total amount of energy allocated by the agent over the time interval [1, n], subjected to the following constraints:

Constraint 1: The allocated energy at time slot t, $E_{i,j}^{alloc}(t)$, is defined by the harvested energy, the charged and discharged energy from the battery and waste:

$$E_{i,j}^{alloc}(t) = E_{i,j}^{hrv}(t) - c(t) + d(t) - w_{i,j}(t) \tag{c_1}$$

Constraint 2: The following equation represents the energy balancing condition, which states that the allocated energy $E_{i,j}^{alloc}(t)$ must not exceed the maximum amount of energy $E_{i,j}^{c}(t)$ that a node can consume at time t:

$$E_{i,j}^{alloc}(t) \leq E_{i,j}^{c}(t) \tag{c_2}$$

Constraint 3: The energy used from the battery at any time t depends on the discharged $d(t)$ and charged $c(t)$ energy plus its efficiency e :

$$B_{i,j}(t) - B_{i,j}(t+1) = d(t) - e \times c(t) \tag{c_3}$$

Constraint 4: The battery level at time $t = 1$ is equal to an initial residual energy b:

$$B_{i,j}(1) = e \times b \tag{c_4}$$

Constraint 5: The cycle of the battery is represented as:

$$B_{i,j}(n+1) = B_{i,j}(1) \tag{c_5}$$

Constraint 6: The energy stored into the battery at each time t, $c(t)$, cannot be negative and must not exceed the maximum battery capacity:

$$0 \leq c(t) \leq B_{i,j}^{max} \tag{c_6}$$

Constraint 7: The energy drawn from the battery at each time t, $d(t)$, when $E_{i,j}^{hrv}(t) < E_{i,j}^{c}(t)$ starts from $E_{i,j}^{c}(t) - E_{i,j}^{hrv}(t)$. This amount must also not exceed the residual energy of the battery:

$$E_{i,j}^{c}(t) - E_{i,j}^{hrv}(t) \leq d(t) \leq B_{i,j}(t) \tag{c_7}$$

Constraint 8: At each time t, the battery must not store more energy than its capacity, also it cannot have negative values:

$$0 \leq B_{i,j}(t) \leq B_{i,j}^{max} \tag{c_8}$$

Constraint 9: Any wasted energy in t is positive and cannot exceed the energy harvested $E_{i,j}^{hrv}(t)$:

$$0 \leq w_{i,j}(t) \leq E_{i,j}^{hrv}(t) \tag{c_9}$$

The use of an LP framework allows us to characterize the energy allocation of a sensor node. The solution to the optimisation problem yields the optimal amount of energy that can be allocated to an agent in every slot t at the beginning of the harvesting period T, $E_{i,j}^{alloc}$, the evolution of residual energy in the battery over the period of n slots, described by the variables $B_{i,j}$, c and d and the energy discarded $w_{i,j}$. The model is extended in Section 3.4 to allow energy sharing between agents. Once the model considers the amount of energy offered by the agents, it can also be used to find the energy offer that maximises the agents' utilities product.

3.3 Optimal Energy Allocation Algorithm

At the stage where the agent has access to the node's information about its initial battery status $B_{i,j}(1)$, battery efficiency e, battery capacity $B_{k,i}^{max}$, detailed energy profile describing the maximum node's load $E_{i,j}^{c}$ and energy harvested $E_{i,j}^{hrv}$ available for the respective time horizon T going through from time slot 1 to n, an agent can compute the node's utility using Algorithm 1.

The algorithm meets the conditions listed in 3.2.4 that must be satisfied to optimise the objective function of energy allocation. At every time slot t, the algorithm evaluates two cases depending on the data of E^{hrv} and E^{c}: when there is enough energy harvesting supply to complete a load (Step 4) and, in the second case, when the energy availability is attempt to be supplied with the help of the battery (Step 13). The values of E^{alloc}, B, c, d, w are derived from 3.2.4 given the data of E^{hrv}, E^{c}, $B(1)$, e and B^{max}. Then, the problem can be solved for any t if $B(t-1)$ is known. Therefore, if the starting battery level $B(1)$ is given, then the algorithm works to find the agent's reserved utilities for $t=1$, and so on.

In more detail, when there is excess energy and it goes above the battery capacity (Step 6), the battery is charged with the excess as required taking into account its greatest capacity and the rest is discarded. Otherwise (Step 8), the battery is only charged with the excess. Step 14 depicts the scenario when there is not enough ambient energy to power a load. There are two cases to evaluate in this statement: when the battery cannot supply the missing energy (Step 16), and the opposite (Step 18). In every case, the values for energy allocation and discharge are depicted. At the end of the algorithm run, the resulting energy allocation scheme describes the situations (Step 16) that can be considered by agent i_j to decide if an OEN with a co-located network must be performed, i.e when an agent can not harvest enough energy for its consumption, and the difference can not be covered with the residual capacity of its battery.

The algorithm described above can be used to automatically alert the agent if a deficient energy allocation scheme is expected. Agent i_j keeps a table of its immediate surroundings or neighbourhood (the nodes that are 1-hop from the agent), which entries correspond to its local neighbours and their energy condition. Each agent piggybacks its information in broadcasting packets via routing updates. Then, agent i_j uses the information from its table to guide a cooperative OEN among its local nodes with the co-located networks. The priority assignment

Algorithm 1: Agent's utility without Opportunistic Energy Negotiation

Input : $E_{i,j}^{hrv} \in \mathbb{R}_{+}^{n}$, $E_{i,j}^{c} \in \mathbb{R}_{+}^{n}$, $B_{i,j}(1) \in \mathbb{R}^{+}$, $e \in [0,1]$, $B_{i,j}^{max} \in \mathbb{R}^{+}$, $n \in \mathbb{Z}^{+}$;

Output: $E_{i,j}^{alloc} \in \mathbb{R}_{+}^{n}$, $c \in \mathbb{R}_{+}^{n}$, $d \in \mathbb{R}_{+}^{n}$, $B_{i,j} \in \mathbb{R}_{+}^{n}$, $w_{i,j} \in \mathbb{R}_{+}^{n}$

Initialisation $(E_{i,j}^{alloc}, c, d, B, w)$;
$E_{i,j}^{alloc}(1) = E_{i,j}^{c}(1)$; $c(1) = B_{i,j}^{max} - (E_{i,j}^{hrv}(1) - E_{i,j}^{alloc}(1))$; $d(1) = 0$;
  $w_{i,j}(1) = E_{i,j}^{hrv}(1) - E_{i,j}^{c}(1) - c(1)$;
for $t \leftarrow 2$ **to** n **do**
  if $E_{i,j}^{hrv}(t) \geq E_{i,j}^{c}(t)$ **then**
    $E_{i,j}^{alloc}(t) = E_{i,j}^{c}(t)$; $d(t) = 0$;
    if $E_{i,j}^{hrv}(t) - E_{i,j}^{c}(t) > \frac{1}{e} \times (B_{i,j}^{max} - B_{i,j}(t-1))$ **then**
      $c(t) = \frac{1}{e} \times (B_{i,j}^{max} - B_{i,j}(t-1))$; $w_{i,j}(t) = E_{i,j}^{hrv}(t) - E_{i,j}^{c}(t) - c(t)$;
    else
      if $E_{i,j}^{hrv}(t) - E_{i,j}^{c}(t) \leq \frac{1}{e} \times (B_{i,j}^{max} - B_{i,j}(t-1))$ **then**
        $c(t) = E_{i,j}^{hrv}(t) - E_{i,j}^{c}(t)$; $w_{i,j}(t) = 0$;
      end
    end
  else
    if $E_{i,j}^{hrv}(t) < E_{i,j}^{c}(t)$ **then**
      $c(t) = 0$; $w_{i,j}(t) = 0$;
      if $E_{i,j}^{c}(t) > E_{i,j}^{hrv}(t) + B_{i,j}(t-1)$ **then**
        $E_{i,j}^{alloc}(t) = E_{i,j}^{hrv}(t) + B_{i,j}(t-1)$; $d(t) = B_{i,j}(t-1)$;
      else
        if $E_{i,j}^{c}(t) \leq E_{i,j}^{hrv}(t) + B_{i,j}(t-1)$ **then**
          $E_{i,j}^{alloc}(t) = E_{i,j}^{c}(t)$; $d(t) = E_{i,j}^{c}(t) - E_{i,j}^{hrv}(t)$;
        end
      end
    end
  end
  $B_{i,j}(t) = B_{i,j}(t-1) - d(t-1) + e \times c(t-1)$;
end

of agents with the same constraints is out of the scope of this thesis. The model to allow the consideration of external energy is described in the next section.

3.4 Opportunistic Energy Negotiation Between Two Agents

The idea of WSN cooperation, in general, is attractive in IoT environments, where WSNs usually have limited energy resources and heterogeneous characteristics, such as battery capacity, number of nodes, nodes locations, energy consumption. From interference avoidance to an appropriate use of energy, the variety of benefits and scenarios is wide. In order to examine each of these cooperative scenarios using a cooperative game-theoretic approach, the above problems are formulated as a two-person game. In this direction, cooperation is not straightforward since networks' authorities are independent of each other and selfish behaviour is inevitable from a rational perspective. Problems such as node's power consumption, interference avoidance for a communication medium and packet routing in cognitive Ad Hoc wireless networks have already been analysed in the form of a game. In the same manner, the majority of research in WSNs has explored the problem of networks' cooperation from a game-theoretic angle [9, 12, 13, 16, 136]. This section describes the system model to enable the study of cooperative energy management. The solution based on a game-theoretic technique described here is then used to benchmark the heuristic proposal in the next chapter.

3.4.1 System Model Supporting Opportunistic Energy Negotiation

Algorithm 1 is simple, decentralised and has a time complexity of $O(n)$, where n is the total number of time slots involved in the period of analysis. This energy management method forms the basic scheme needed for energy allocation and gives the sensor network self-organising ability. Following the allocation algorithm, an agent can identify if a cooperative effort is required. Such cooperation effort can increase the agent's utility if the agents find a suitable agreement during negotiation. This thesis considers self-interested agents which prefer the energy flow that maximises their own utility. The additional variable of energy flow must be evaluated in the proposed LP model in Subsection 3.2.4 to realise the vision and analyse the benefits of opportunistic energy negotiation.

An opportunistic energy negotiation is triggered when a node's energy level $B_{i,j}(1)$ has dropped below a threshold, and the foreseen energy $E_{i,j}^{hrv}$ is not enough to maintain the next period. During negotiation, agent i_j considers the amount of energy to receive/give from the cooperation at each time t, which is defined by $\boldsymbol{o} = (o(1),...,o(n)) : \boldsymbol{o} \in \mathbb{R}$. o represents the offer of energy at each time slot, i.e. **The issues over which the negotiation takes place**. These offers are called *energy flow offers*. A valid energy flow offer must include the energy values for the predetermined time of cooperation, e.g. If the networks expect to cooperate for 24 hours, then the energy flow must include 24 values. $\boldsymbol{o} = (o(1),...,o(n))$ can be also referred to as the value vector of the negotiation's attributes. As soon as the offers $o(t)$ appear in the model, this is no longer an LP with a straightforwardly predictable outcome as a solution. Offers make agents commit some energy outside of the case described in Step 16 (see Algorithm 1), in order to gain energy at such stage, overall leading to a possible surplus. Moreover, agents can have excess energy during intervals of less activity and share this benefit with their opponents.

The direction of the energy flow is denoted by a positive or negative sign. If positive, the amount is an offer of energy from the agent to its opponent, otherwise, it represents the energy to be received from the opponent. For example, if two agents are willing to cooperate with each other for a period of 2 hours and L is set to 30 minutes, then an offer of energy from agent i_j to the opponent can be $o = (-1.88, -0.7, 18, -4)$; where -1.88 mWh, -0.7 mWh and -4 mWh represent the energy savings of agent i_j from the opponent's cooperation (e.g. by packet routing) at time slots 1,2 and 4 respectively. While value 18 indicates that agent i_j is willing to compromise 18 mWh of energy through a collaborative effort to its opponent at time slot 3. Thus, the energy flow offer affects the utility value of the agent, i.e. the amount of energy allocated for a sensor node. By this condition, an agent has to evaluate this additional variable at every time slot since the energy available in the cooperative scenario is affected now by the energy flow offered at each time. The constraint (c_1) can now be replaced in the original optimisation problem with (c_{10}) to include the energy flow offer as follows:

$$E_{i,j}^{alloc}(t) = E_{i,j}^{hrv}(t) - c(t) + d(t) - w_{i,j}(t) - o(t) \qquad (c_{10})$$

Basically, when o is null the problem describes the energy allocation for an agent without cooperation.

In this cooperative model of energy allocation, an additional condition must be addressed. Since energy is logically transferred between networks by accepting energy-consuming tasks as data processing or packet forwarding [8, 12, 13], a change is required in constraint (c_6) when there are offers involved:

$$0 \leq c(t) \leq E_{k,i}^{hrv}(t) \qquad (c_6)$$

The result is that the battery will be charged immediately with the energy harvested by the agent while the energy supply received from the opponent's offer will be used to satisfy the agent's load.

Following the model described, an agent can compute the optimal energy flow that benefits both agents (i.e. all energy flows that give a higher utility to each agent than those provided without cooperation) but this requires complete information (energy profiles of both agents and battery information) and high computation capabilities since the set of all feasible agreements is exponential in the number of time slots. Cooperative approaches must ideally result in Pareto-efficient outcomes. By this, the resulting energy flow across inter-network nodes must be on the boundary of the feasible solutions. On this boundary, one agent cannot be better off without making the other agent worse off. As described in Section 2.3.1, a cooperation strategy based on Nash Bargaining Solution (NBS) is the solution that maximises the product of agents' payoffs over the set of all feasible agreements. With NBS, the agents will cooperatively work and each will share a certain fraction of its energy surplus for optimal energy management. The next section presents the cooperative solution that finds an agreed energy flow between agents and satisfies the property of efficiency.

3.4.2 A Cooperative Bargaining Solution for Opportunistic Energy Negotiation

Referring to the cooperative game theory approach, the problem of cooperative energy management for multi-domain rechargeable nodes is formulated as a two-agent bargaining game. Then, a cooperative energy allocation strategy based on the Nash Bargaining Solution (NBS) is presented, where each agent can share a certain amount of its energy generated for performing node's tasks (processing, sensing, transmission) in a collaborative form. This section shows how NBS (described in Section 2.3.1) can be used as an axiomatic method to compute an energy flow offer between agents. Experimental results show the NBS energy sharing is fair for both agents in a way that a node will only cooperate with a node from a different domain if this cooperation contributes to solve its energy deficiency caused by the spatio-temporal pattern of its energy source. The degree of cooperation of a node depends on how much contribution its opponent can make to its continuous operation.

In the domain of EHWSNs and energy negotiation, the set of all feasible agreements is the set of all possible energy flows which give both agents more utility than their reservation values (energy allocation when there is no negotiation). Let S denote the set of all feasible energy agreements that the players can get if they work together. Considering two negotiation agents 1_1 and 2_1 with disagreement/reserved utilities $d_{1,1}$ and $d_{2,1}$, respectively, which correspond to the maximum utilities that agents can achieve when they do not agree on an energy flow offer ($o = \{\}$), and o^{NBS} as the axiomatic solution. A point $(u_{1,1}, u_{2,1})$ represented in the bargaining set S describes the utility of the agents on energy allocation. Since agents will only cooperate if they get more utility than their disagreement values or load satisfied without cooperation, the following must hold: $(u_{1,1}, u_{2,1}) \geq (d_{1,1}, d_{2,1})$. Then, a possible solution o^{NBS} to the cooperation problem between EHWSNs nodes' can be computed as:

$$o^{NBS} = \underset{o \in S}{\arg\max}\, [u_{1,1}(o) - d_{1,1}] \times [u_{2,1}(o) - d_{2,1}] \tag{3.4}$$

The product of the two excess utilities is referred to as the Nash product. The equation 3.4 is subjected to the conditions defined in Subsection 3.2.4 and the respective constraints modifications of Subsection 3.4.1 for both agents. Normally if S is convex and compact, o^{NBS} is unique. However, since there are multiple interdependent issues involved, there are several possible solutions for o^{NBS} and any of them defines the cooperative agreement that maximises the product of the individual utilities. This is reported below.

Table 3.2 shows the reserved utilities from agents 1_1 and 2_1 with the parameters of energy harvesting E^{hrv}, load E^{c}, initial battery level $B(1)$, battery efficiency e set to 0.7 and NiMH battery capacity B^{max} of 708 mWh. The corresponding values for agent 1_1 are [0;0;4], [2;2;2] and 0 mWh. The initial battery status represents any range or threshold identified by the agent that supports an envisioned insufficient energy allocation. For agent 2_1, the values are [8;0;0], [4;4;4] and 0 mWh, respectively. The values of E^{alloc}, c, d, w and consecutive B are returned by Algorithm 1.

Tables 3.3 indicates the allocations based on NBS. The solution is computed using equation 3.4 and conditions listed in Subsection 3.2.4 with the respective constraints changed in Subsection 3.4.1 for both agents. This work uses a nonlinear bound-constrained optimisation solver in MATLAB for the computation of NBS. From agent's 1_1 perspective, the energy values accompanied by a positive sign represent an offer of energy from 1_1 to its opponent 2_1, while negative values represent the energy to be received from the agent node 2_1. The results give an illustration of the multiple potential cooperative flows explained above. Both NBS offers o^{NBS} represent the same agent's utilities.

	Agent 1_1						**Agent** 2_1					
t	$E^{hrv}_{1,1}$	$E^{c}_{1,1}$	$E^{alloc}_{1,1}$	c/d	$B_{1,1}$	$w_{1,1}$	$E^{hrv}_{2,1}$	$E^{c}_{2,1}$	$E^{alloc}_{2,1}$	c/d	$B_{2,1}$	$w_{2,1}$
1	0.00	2.00	0.00	0.00	0.00	0.00	8.00	4.00	4.00	4.00	0.00	0.00
2	0.00	2.00	0.00	0.00	0.00	0.00	0.00	4.00	3.60	-3.60	4.00	0.00
3	4.00	2.00	2.00	2.00	0.00	0.00	0.00	4.00	0.00	0.00	0.00	0.00
	Reserved utility (d): 2						**Reserved utility (d): 7.6**					

Table 3.2: Agent's reserved utilities.

	Agent 1_1							**Agent** 2_1					
t	$E^{hrv}_{1,1}$	$E^{c}_{1,1}$	$E^{alloc}_{1,1}$	c/d	$B_{1,1}$	$w_{1,1}$	o^{NBS}	$E^{hrv}_{2,1}$	$E^{c}_{2,1}$	$E^{alloc}_{2,1}$	c/d	$B_{2,1}$	$w_{2,1}$
1	0.00	2.00	2.00	0.00	0.00	0.00	-2	8.00	4.00	4.00	2.00	0.00	0.00
2	0.00	2.00	0.38	0.00	0.00	0.00	-0.38	0.00	4.00	1.42	-1.80	2.00	0.00
3	4.00	2.00	0.72	0.00	0.00	0	3.28	0.00	4.00	3.28	0.00	0.00	0.00
	Utility with NBS: 3.10							**Utility with NBS: 8.7**					

	Agent 1_1							**Agent** 2_1					
t	$E^{hrv}_{1,1}$	$E^{c}_{1,1}$	$E^{alloc}_{1,1}$	c/d	$B_{1,1}$	$w_{1,1}$	o^{NBS}	$E^{hrv}_{2,1}$	$E^{c}_{2,1}$	$E^{alloc}_{2,1}$	c/d	$B_{2,1}$	$w_{2,1}$
1	0.00	2.00	2.00	0.00	0.00	0.00	-2	8.00	4.00	4.00	2.00	0.00	0.00
2	0.00	2.00	0.55	0.00	0.00	0.00	-0.55	0.00	4.00	1.25	-1.80	2.00	0.00
3	4.00	2.00	0.55	0.00	0.00	0.00	3.45	0.00	4.00	3.45	0.00	0.00	0.00
	Utility with NBS: 3.10							**Utility with NBS: 8.7**					

Table 3.3: Agent's utilities with NBS.

The resulting energy allocation is referred to as social optimal, or Pareto optimal, which means that one agent can not improve its own utility without disadvantaging the other node's performance. The interest of game theory is to model interactions between selfish nodes and determine the cooperation strategies that could lead to socially optimal resource allocation. However, the computation of a solution as NBS implies some considerations that are infeasible in the opportunistic and direct node-to-node negotiation setup. Such considerations include the knowledge of complete information (reservation values, energy consumption models, and battery information) at the first step of the negotiation or the presence of a central and trustable mediator that collects the information about the agents and computes the solution.

The performance of the heuristic method is compared in terms of Pareto optimality with this axiomatic solution where agents declare their true reservation values and utility function. The next section presents the simulation setup and datasets used to analyse the utilities of two agents with and without energy sharing.

3.5 Experimental Validation

This section provides a detailed scenario to evaluate the performance of the optimisation algorithm based on LP and its use in the domain of cooperative energy management. All the results are obtained using MATLAB.

The problem of cooperation is studied in a simplified scenario with a pair of nodes (each from a different network); negotiating agents 1_1 and 2_1 of N_1 and N_2, respectively. It is assumed that the agents observe that their residual energy level has dropped below the threshold set, e.g. $B(1)$ has reached a value in the range [0,1]. Besides, the energy supply from the environment is not enough to feed the agent's load. Agents have appropriately synchronised times and plan to cooperate for the next 24 hours. The time period starts at 00:00 and ends at 23:00 local time with L=1 hr, i.e. Agents will have to negotiate an energy flow of 24 values. Then T corresponds to the same period and time slots $T = (1,...,24)$. The nodes' specification and energy profiles: available energy (harvested energy) plus load required to make power management decisions are described below.

3.5.1 Nodes

In this simulation model, the agents' parameters are set using empirical measurements and information presented by MEMSIC datasheets [137].

The energy generation profile is simulated with two different types of renewable energy: solar and wind. Agent 1_1 controls a sensor node with a solar panel while agent 2_1 manipulates a sensor node with wind turbines.
Agent 1_1 controls a Memsic eKo mote, which contains a 3.3 cm $\times$ 6.35 cm photovoltaic solar panel assumed to be 10% efficient to recharge a 600 mAh NiMH battery [138]. The efficiency $e = 0.7$ is considered for the simulations, which is a typical value for NiMH batteries [139].
Agent 2_1 controls a Memsic MICAz node, with a micro-wind turbine to recharge a 600 mAh NiMH battery.
Table 3.4 presents a summary of the simulation parameters.

	Agent 1_1	Agent 2_1
Active current	20 mA	19 mA
Sleep current	5 µA	5 µA
Voltage	3 V	3 V
Duty cycle	1%	5%
Battery capacity	600 *mAh*	600 *mAh*
Battery efficiency	0.7	0.7

Table 3.4: Agent's parameter values for power usage and battery storage model.

3.5.2 Energy Profiles

The energy model in 3.2.2 is used to evaluate the energy consumption of both agents, using parameters obtained from empirical measurements and datasheets [137].
A realistic scenario is considered where an eKo node normally operates at a 1% duty cycle, and its average power consumption is 0.615 mW. For the MICAz mote, an average load of 2.86 mW at 5% duty cycle of operation is desired. Thus, agent 1_1 and agent 2_1 demand 0.615 mWh and 2.86 mWh of energy in each time slot, respectively.

This thesis uses hourly wind speed collected at Weather Underground [140] and solar radiation from PVGIS [141] for a period of one year (2017) corresponding to the area of Southampton, in the United Kingdom (50.8997°N, -1.3955°W, Elevation 35m). These datasets are used to compute the energy generation profiles of the nodes.

The solar power harvesting profile is obtained from the solar radiation data of April 2017. The values of direct solar irradiance are used to estimate the hourly power output of a photovoltaic system for a day. The solar power is directly proportional to the value of solar radiation (G_b) [W/m^2], the panel dimension (0.033 m × 0.0635 m) and its efficiency (0.1) [25]:

$$E_{1,1}^{hrv} = G_b \times 0.033 \times 0.0635 \times 0.1 \times 1000 \text{ [mW]} \tag{3.5}$$

Hence, the hourly power generation of a 3.3 cm × 6.35 cm photovoltaic solar panel for a regular spring day in the area of Southampton city centre can be estimated. The estimated hourly energy output for a day is shown in Figure 3.4. As can be seen from the figure, the energy generation exhibits a temporal variation that favours time slots 6-19 which corresponds to times 05:00-18:00, where most of the energy is produced between time slots 10-15 (from 9:00 to 14:00). The total energy generated in a day is 451.5 mWh.

Raw daily data for April 2017 collected at Weather Underground is used. The API call is found in the references [142]. These records are employed to estimate the hourly average wind speed for a day.

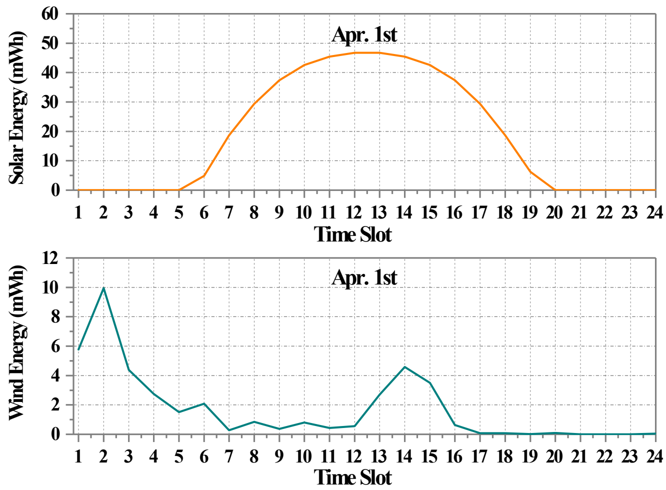

Figure 3.4: Solar (agent 1_1) and wind energy (agent 2_1) harvested throughout a day.

The power from the wind source can be calculated from the measured wind speed v [m/s] as in [27, 143]:

$$E_{2,1}^{hrv} = 0.5 \times \rho \times A \times v^3 \times f \tag{3.6}$$

where ρ is equal to the air density (1.22 $\mathrm{kg/m^3}$), A is the swept area of the wind turbine set to 0.0025 $\mathrm{m^2}$ in this scenario and f is the efficiency of the windmill (f = 1). From April data, April 1st is chosen. In order to find a feasible solution with NBS, there must be a feasible space of agreements. Thus, the diversity between generation times in solar and wind is suitable to create an opportunity for energy negotiation. The values found with the equation are scaled to get the hourly power output of a highly efficient micro-turbine [144] (Figure 3.4). In contrast to solar energy, it can be observed from the graphic how the wind energy has a very irregular pattern over a day. There are peaks in the early morning and initially in the afternoon (time periods 1-4 and 13-15) while some intervals exhibit very low or no energy (e.g. time periods 16-24). The total energy generated in a day is 41.4 mWh.

3.5.3 Energy Allocation Without Opportunistic Energy Negotiation

With the node's information and energy profiles described above, agents can compute their utilities (without cooperation) using Algorithm 1 presented in Section 3.3 when the offer o is null.

The result for agent 1_1 is illustrated in Figure 3.5. Figure 3.5(a) shows how the energy allocation (E^{alloc}) which maximises the utility of agent 1_1 is insufficient to power the sensor node and its demanded load at each time slot (0.615 mWh). This optimal allocation employs 2.6% of the energy supplied by the energy source (11.68 mWh out of 451.5 mWh) when the sensor node

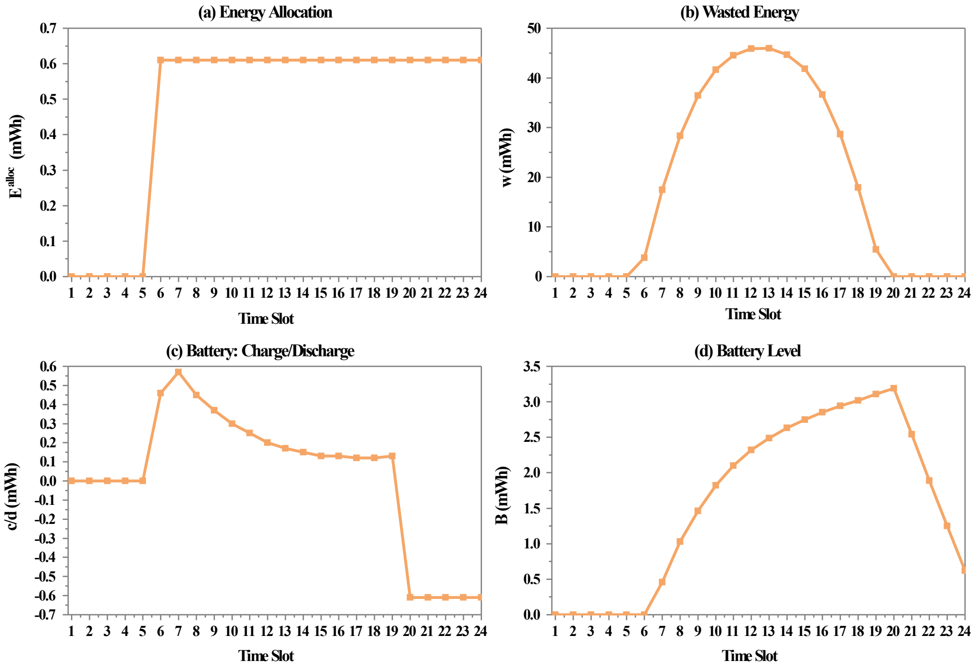

Figure 3.5: Agent 1_1: Results of utility maximisation without energy negotiation.

controlled by the agent 1_1 has a fixed duty cycle of 1%. Therefore, there exists a high energy waste (Figure 3.5(b)) due to the irregular pattern of the energy harvested from the environment, which is sufficiently large at every time interval except during time slots 1-5 (00:00 - 04:00). The waste corresponds to the excess of energy generated that is not utilized by the sensor node during time slots 6-19. Figures 3.5(c) and 3.5(d) show the dynamic of the battery during the day with charging and discharging amounts and its residual state, respectively. The residual energy at each point in time matches the dynamics of the charging and discharging flows and it does not exceed the maximum battery capacity. The positive values in Figure 3.5(c) represent the amount of energy charged to the battery, while the negative values correspond to the amount of energy discharged from the battery. It can be seen in Figure 3.5(c) that the initial battery state as well as the available energy in time slots 1-5, indicates a null level of energy. Without energy, a sensor node is useless and cannot add utility to the network as a whole. Consequently, this is anticipated by the sensor node when its resources are still enough to participate in the bargaining process and is able to seek the cooperation of a neighbour sensor node.

As observed, the battery is charged while the harvester source provides energy (time slots 6-19). This is charged with small amounts of energy, as required (e.g 0.46 mWh at time slot 6). Figure 3.5(c) shows the energy flows into the battery until time slot 20, then it starts to flow out of the battery since the energy generated by the harvester source is null. The energy discharged from the battery matches the energy requirements of the sensor node during the time when the energy generation is null. This is the reason why the charging/discharging amount is very low.

Figure 3.5(d) shows the residual energy level or battery status of the sensor node over a day. It shows how the battery level matches the energy flow of figure 3.5(c), The battery level increases until time slot 20, when it starts to decrease its level since all the energy used to feed the node during time slots 19-24 come from the battery. It can be seen how the battery level drops when the battery is discharged (time slots 19-24). The largest amount of energy stored in the battery at any time is 3.19 mWh which is far below its maximum capacity of 708 mWh. This confirms the storage capacity is sufficiently large for the demanding load.

As a result, the ratio of satisfied load to the total energy consumption of the node is 0.8 (i.e. 11.68/14.76) via optimisation without cooperation, i.e. the agent can allocate a maximum of 11.68 mWh from its harvested energy over 14.76 mWh desired using its bounded energy allocation scheme. Then, agent 1_1 can achieve a utility of 0.8 when it depends only on itself.

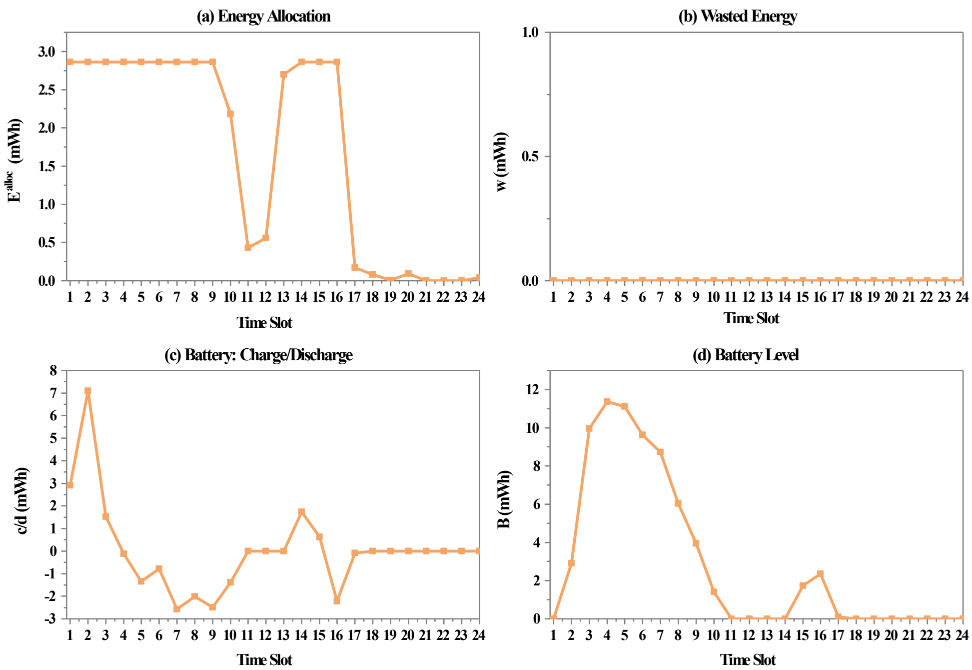

Figure 3.6: Agent 2_1: Results of utility maximisation without energy negotiation.

Figure 3.6 presents the results without energy negotiation for agent 2_1. Figure 3.6(a) shows the energy allocation (E^{alloc}) which maximises the utility of agent 2_1. Again, the allocation in this scenario is scarce and it does not supply the load of 2.86 mWh at several time slots (10-13, 17-24). This optimal allocation utilizes 98% of the energy generated by the energy source (40.56 mWh out of 41.43 mWh) when the sensor node controlled by the agent 2_1 has a fixed duty cycle of 5%. In this case, it is observed that there is no waste (see Figure 3.6(b)). This means that the energy generated by the micro-wind turbine was entirely used and the battery capacity is large enough for the dynamics generated by the allocation algorithm. Thus, the 2% of the energy that is missing from the total 41.43 mWh, is entirely associated with the loss of the battery efficiency (30% loss of e = 70%). Figures 3.6(c) and 3.6(d) show the energy flow that goes into and out of the battery and the resulting battery level over a period of 24 hours. As illustrated, the battery is charged while there exists a provision of energy from the harvester source that exceeds the load (time slots 1-3 and 14-15). The battery is discharged during the time when the energy generation is null or deficient to power the load and it can be seen that there is enough energy in the buffer (time slots 4-10, 16-17). Figure 3.6(d) of the residual energy at each time slot, matches the dynamics of the charging and discharging flows and it never exceeds the maximum battery capacity. An increment in the battery level during time slots 1-4 and 14-16 that corresponds to the charging amounts of energy is given. In the same way, the discharging amounts of time slots 4-10 and 16-17 reflect the agent's battery level during time slots 5-11 and 16-18. It is also important to note that similar to the case of agent 1_1, the charging and discharging amounts as well as the battery state are very low and respect the limit of the battery capacity, which confirms that energy is not wasted due to insufficient storage space.

Consequently, the ratio of satisfied energy consumption to the total load of the node is 0.6 (i.e. 40.56/68.64) via optimisation without cooperation. Therefore, agent 2_1 reaches a utility of 0.6 when the energy allocation algorithm is limited to its local domain.

In summary, the maximum utilities that both agents expect to achieve without cooperation are 0.8 and 0.6, for agent 1_1 and 2_1 respectively. In the next section, an energy flow is calculated using the cooperative solution NBS and compared with these results.

3.5.4 Optimisation Results With Opportunistic Energy Negotiation using NBS

The simulation setup described at the beginning of this Section 3.5 was used to obtain the optimisation results presented here.

Figures 3.7 and 3.8 illustrate the utility maximisation of agents 1_1 and 2_1, respectively. These results are obtained when agents start the process to enable the transfer of energy and use the Nash bargaining solution to compute a cooperative energy flow. The most prominent outcome in this scenario is the achievement of energy neutrality by both agents. The utility of agent 1_1 is increased from 0.8 to 1 while agent 2_1 is able to increase it from 0.6 to 1. As can be observed from Figures 3.7(a) and 3.8(a), both agents are continuously powered for the period of 24 hours by logically sharing energy resources with each other. There is a minimum amount of load that is not satisfied by agent 2_1 during time slots 4 and 5 but it corresponds only to 1.8% of its power consumption. In contrast to the results found without energy negotiation, where the energy harvested by a node of one domain is not enough to satisfy its load, the agents under this strategy improve their energy management. With these results, complex adaptive algorithms are avoided and the application performance is maintained at the same rate at all times, i.e. the duty cycle is not affected.

Thus, compared to existing power management strategies, the energy negotiation between networks permit cross-boundary transfer of energy resources. The transfer can be done by allowing a node to forward or process packets on behalf of its neighbouring counterpart. As a result, a EHWSN can manage its energy not only using resources within its own network but also among a group of networks that have a wider coverage area and resource capacity.

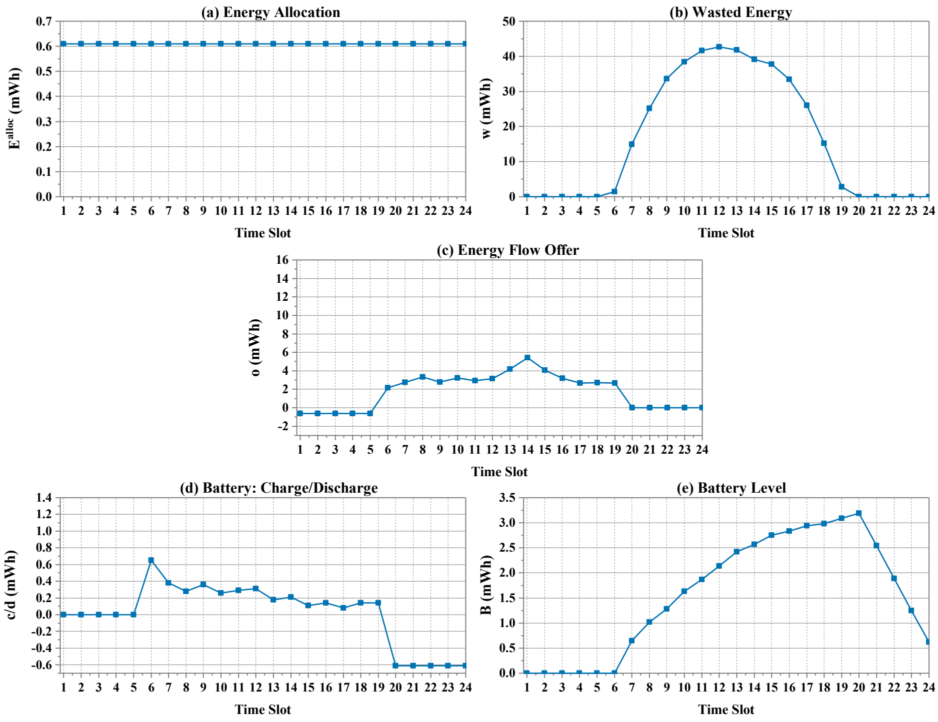

Figure 3.7: Agent 1_1: Optimal energy allocation that utilises NBS from cooperative game theory.

Figure 3.7(b) shows the energy wasted in each time slot. The optimal energy allocation for agent 1_1 which maximises its energy via cooperation increases the utilisation of energy harvested from 2.6% to 12.71% for one day of energy sharing. A total of 394.09 mWh of energy is not used in contrast to 439.31 mWh when the node simulated controlled by agent 1_1 manages its energy resources across multiple networks. With respect to this result, 87% of the energy harvested is wasted, which is less than the 97.3% without cooperation.

Figure 3.7(c) shows an energy flow between agent 1_1 and agent 2_1 with NBS for this setup. The positive amounts represent the value of energy that agent 1_1 compromises and provides to agent 2_1 through a collaborative effort, while the negative flows indicate the energy savings that agent 1_1 is able to obtain from the opponent's cooperation. This is illustrated by the fact that agent 2_1 cooperates with agent 1_1 in the early morning over time periods 1-5 by providing a service (either sensing, processing, forwarding, etc) to benefit agent 1_1 with energy savings of 3.05 mWh. This represents the incentive for agent 1_1 to cooperate with agent 2_1 over the next periods (6-19).

Figures 3.7(d) and 3.7(e) illustrate the battery life. The battery status showed in Figure 3.5(d) and the residual energy level achieved with energy transfer in Figure 3.7(e) are very similar. Such similarity is conceived by the constraint to charge only the amount of energy harvested even in the existing of a cooperative energy flow. The cooperative energy flow represents the amount of work each agent compromises to its negotiation partner. Once the agents start to cooperate, a watchdog can identify misbehaving agents and enforce the agreement reached. As can be seen from the graphics, the residual energy at each point in time matches the dynamics of the charging and discharging flows and it does not exceed the maximum battery level. Rather, it is far below the battery capacity.

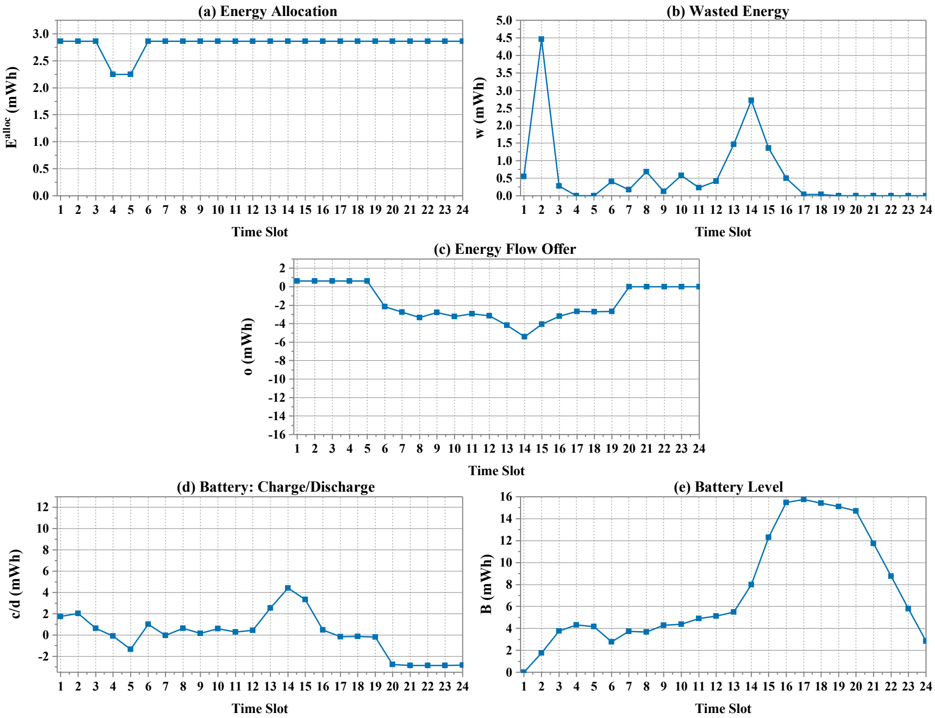

Figure 3.8: Agent 2_1: Optimal energy allocation that utilizes NBS from cooperative game theory.

Figure 3.8 shows the utility maximisation of agent 2_1, when agents logically share energy with each other through a cooperative energy allocation based on NBS. Figure 3.8(a) presents the realisation of energy-neutral operation by agent 2_1 when there is energy negotiation. A margin up to 90% load satisfied is considered. Thus, the proposed cross-network negotiation approach can enable heterogeneous co-located nodes to cooperate and optimise their network performance by selecting the optimal resources so that all involved participants gain from cooperation.

Unlike Figure 3.6(b), Figure 3.8(b) presents a waste of energy over the period of cooperation. The excess is generated by the offer of energy agreed between agents, which corresponds to a total amount of 14 mWh of energy unused. Thus, network nodes must be responsible to locally control nodes providing cooperation favours.

Figure 3.8(c) shows the energy flow from agent 2_1 to agent 1_1. As can be seen from the graphic, agent 2_1 receives the collaborative effort from agent 1_1 over periods 6-19, which is equivalent to 50.19 mWh of energy. Then, the energy transfer process between networks is fair in the way that nodes cooperate with each other when this provides benefits that justify the cooperation.

Figures 3.8(d) and 3.8(e) illustrate the dynamics of the battery. In contrast to Figure 3.6(d), the battery level of the agent in Figure 3.8(e) is never depleted. Instead, there is always a reserve available to power the node. The maximum level achieved by the battery at any time corresponds to 16 mWh, which use remains within the bounds of the battery's capacity.

In this section, a new interdomain energy allocation approach that is Pareto-efficient and fair from the perspective that both agents share an amount of energy for the realisation of cooperation was presented. Two major drawbacks of this solution are the need of a mediator and the assumption of complete information. This approach was simulated with real node's specifications and energy profiles. The experiments found that the solution based on cooperative game theory can provide a significant improvement over the utilities of each node involved. Moreover, the agreement guarantees that each network will benefit from the cooperation. The next chapter presents a solution that considers the decision-making process and not only the properties of the negotiation outcome. The cooperative approach and the heuristic model are also compared over a series of experiments.

3.6 Summary and Discussion

The main advantage of energy-harvesting WSNs compared to traditional battery-based WSNs, is that they have an unlimited power supply provided by their renewable energy sources with the capability to recharge and enhance the performance of the network as a whole. However, due to the dynamics of ambient sources such as solar light or wind that lead to low recharging rates or uneven energy distribution, this energy provision becomes virtually insufficient. Since even energy harvesting systems are not able to operate continuously, it is necessary to analyse alternative power management schemes that incorporate self-organised approaches associated with efficient operation.

Consequently, as discussed in Section 2.5, several research efforts have been made in the field of power management to model and adapt the network behaviour to the energy variation and achieve energy-neutral operation. But, these solutions are limited to a single domain, and co-existence between networks have not been considered. Moreover, these algorithms typically adjust parameters such as the duty-cycle or sampling rate. However, in addition to energy neutral operation, certain applications require the duty cycling behaviour of a node to be as stable as possible, meaning that it should have minimum variance over time [145, 146].

The cooperation problem among different WSNs has only been studied using a game-theoretical framework to model the interactions between battery-powered WSNs, while EHWSNs have been left out of this context. However, in this chapter, a novel energy allocation algorithm (Algorithm 3.3) with linear complexity that enables self-organising capabilities to nodes in a rechargeable WSN was proposed. This solution has been developed to maximise the total utility of an agent and alert it of scarce energy resources. Moreover, the energy allocation scheme can be used to start a joint strategy change by co-located nodes.

By means of the system model to support opportunistic energy negotiation described in 3.4.1, two agents can look for energy sharing agreements that optimise their energy allocation. The negotiation approach is used to address individual node preferences related to their battery information, energy consumption and energy generation profile. Using the cooperative model, an approach based on opportunistic energy negotiation where nodes from distinct networks can cooperate to optimise their use of energy was proposed. The use of negotiation in this context has two major benefits: (i) heterogeneity in terms of network resources can be solved in a decentralised manner and (ii) nodes can devise a cooperation that provides mutual benefit by communicating and compromising.

This chapter introduced a realistic scenario with real node's specifications and energy profiles that allows nodes to find an optimal solution to the cooperation problem. The presented example probes the achievement of energy neutrality by both agents, enabling energy management across network boundaries through an energy flow computed using NBS. As pointed out, the calculation of this solution has some implications in conflict with the characteristics and requirements of the domain. Thus, in the following section a strategic model of negotiation that considers the limitations of this scenario and realises intra-network power management is described.

Although the advantage of the proposed negotiation approach on opportunistic and direct cooperation have been identified, there are some limitations in its practical implementation. A WSN may be formed by hundreds of nodes and concurrent negotiation processes may share the same compromised resources. Thus, the next chapter starts by describing the establishment of OEN and how agents identify the set of available opponents.

Chapter 4

Opportunistic Energy Negotiation: A Heuristic Approach

According to the system model to support OEN introduced in Chapter 3, a heuristic approach to address the negotiation problem is presented in this chapter. Since networks are formed by multiple nodes, the first section of this chapter describes the initiation of OEN and how nodes organise to start a negotiation process. The protocol to discover negotiating agents with the desire to engage in cooperation is illustrated in Section 4.1. Along with this, network performance metrics are evaluated accordingly. A negotiation framework for automated multi-issue negotiation is presented in Section 4.2. To evaluate the effectiveness of this approach, extensive simulations based on the remaining available negotiation rounds and the agent's behaviour are conducted (Section 4.3). Results show that the model can be suitable for practical use in automatic energy re-allocation.

4.1 Establishing the Opportunistic Energy Negotiation

Before the agents face the challenge of selecting a negotiation partner, they need to discover the negotiation agents in the neighbourhood, i.e. the agents that want to cooperate and establish an OEN. This section evaluates the cost associated with the overhead of the discovery protocol in the network's performance.

4.1.1 Discovery of Neighbouring Devices

OEN adopts a publish-subscribe approach in which the agents conserve energy by sending a limited number of messages. Three types of messages are exchanged between agents: OEN_ADV, OEN_REQUEST and OEN_ACCEPT. Initially, the agents are deployed with a cross-domain link-layer protocol such as OI-MAC [147] and a standard routing protocol. OEN is implemented between the link and network layers. In this way, it takes advantage of both layers' functionalities. The agents can communicate directly with co-located agents using the capabilities of the link layer protocol. Moreover, they are still able to inform the network layer about the cooperative agreements reached with their counterparts. This approach is known as cross-layer design.

In the WSN cooperation literature, the networks increase their performance by cooperative packet forwarding. In fact, this is one of the tasks that networks can service in order to share energy. This specific type of cooperation is used to guide the decision process of being part of an OEN.

Once the agents are deployed, the first step in OEN for a negotiator is to broadcast through its immediate neighbours on all available radio frequencies, the desire to start a negotiation by sending an OEN_ADV message. From that moment, the agent becomes the main agent: the agent that will choose a negotiation partner from a set of opponents. OEN_ADV includes the list of agents in the main agent's range (nodes from its own network and from the external network domain) and a query to find the neighbours of the neighbouring agent contacted. This information is then used by an agent to decide if a cooperative packet forwarding can be established with the external neighbour. Figure 4.1 illustrates the discovery protocol in a sequence diagram and the OEN header format.

At this point, there are two possible situations per neighbouring agent reached. The main agent (call this 1_1) with another agent (call this 2_1) may have no interaction. The information provided about the nodes in range by agent 1_1 may not be ideal for agent 2_1 and it can simply ignore the main agent's request. Thus, agent 1_1, after waiting for a certain interval of time, drops the communication with agent 2_1 and stays in the initial state while the number of nodes discarded is different from the total number of its neighbours. On the other hand, agent 2_1 may accept the main agent's proposal. In this situation, agent 2_1 sends an OEN_REQUEST using the radio frequency that is associated with the main agent 1_1 to ask for participation in OEN. In this message, agent 2_1 informs agent 1_1 about the agents that are in its range.

Again, there are two possible scenarios. First, agent 1_1 may ignore agent 2_1's request. This may happen due to two reasons: agent 1_1 is already part of an OEN with another set of agents or agent 1_1 is now unreachable. Agent 2_1 then waits for a grace period before discarding the proposal of agent 1_1. The second possible scenario includes a response. Agent 1_1 may accept the agent's request and send an OEN_ACCEPT message to add agent 2_1 to the pool of opponents. This leads agent 1_1 to a selection state if the number of agents in the set of opponents is bigger than one, if not, the agent moves to a final state, the state of negotiation. In the state of negotiation, both agents can directly establish a bilateral negotiation. Conversely, in the selection state, agent 1_1 chooses a negotiation partner from the pool of agents and move to the final state of negotiation. For the purposes of this chapter, a negotiation partner is randomly

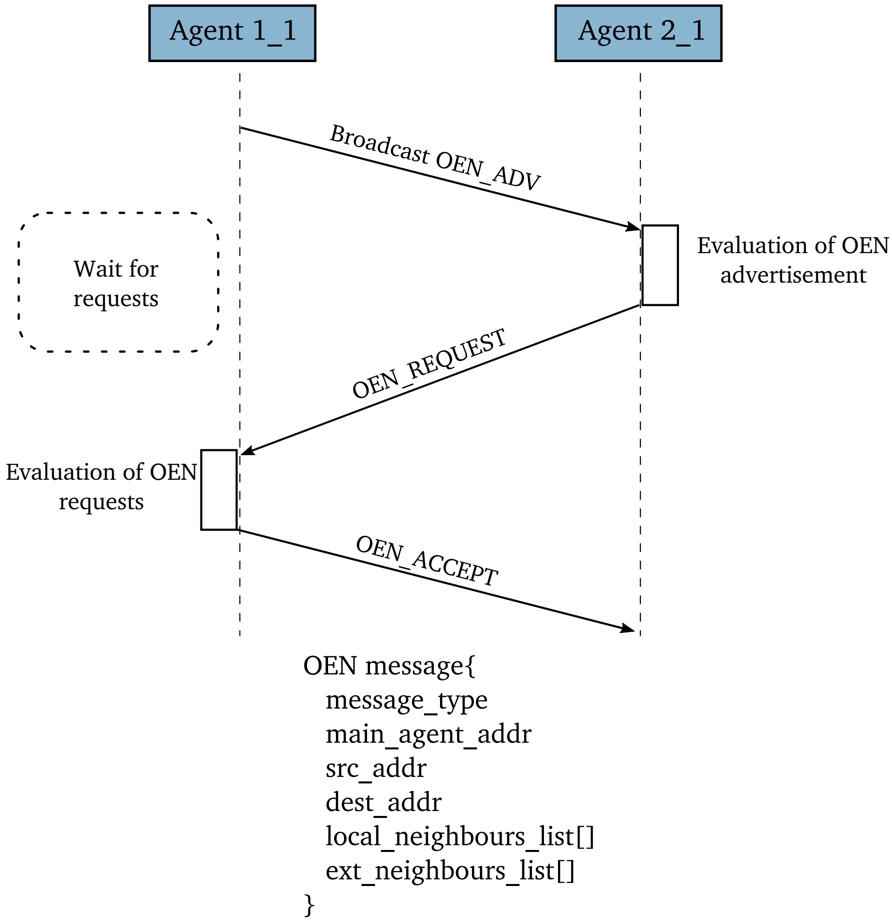

Figure 4.1: Sequence diagram of OEN establishment.

selected. However, Chapter 5 shows how appropriate action-selection policies can be introduced to improve the partner selection phase.

4.1.2 Simulation Setup in OMNeT++

The discovery protocol was implemented using OMNeT++ [148] and the INET framework [149]. OMNeT++ is a discrete event simulator primarily for building network simulations. The tool includes a graphical runtime environment with event logs, a Eclipse-based IDE to code C++ modules, and well-structured documentation. Once the components of the simulation (ned definition, computation of scalars, ini, anf and sca files) and the parameter studies are understood, OMNeT++ offers a complete simulation tool to develop network protocols or any parallel and distributed algorithm. The tool was selected for its compatibility with Linux-based systems and its easy integration with INET, an open source library that contains the implementation of wireless networks standards such as IEEE 802.15.4. IEEE 802.15.4 was designed to specify the physical layer and medium access control for low rate WSNs.

The modules and models used of INET are used to reproduce sensor nodes with the following characteristics. The 802.15.4 MAC is based on collision avoidance via CSMA/CA with ACK support. The radio interface included in the simulation is Ieee802154NarrowbandInterface. The energy storage model SimpleEpEnergyStorage simulates the residual energy capacity and initial battery level of the node. The energy storage capacity is set to 5 J. SimpleEpEnergyManagement allows describing the energy consumption model. The OEN discovery protocol is implemented as a component of the sensor node between the link and network modules following the description given in Subsection 4.1.1.

The power consumption model used in the simulation follows the one described in Subsection 3.2.2, a transceiver energy consumer model based on the radio mode, voltage, current and time. The transmission/reception states determine in these experiments the value of respective current (Tx current, Idle listen current, Rx current, Rx-Tx current). PHY and MAC layers are defined by the IEEE 802.15.4 standard, while the rest of the parameters used in the simulations are summarised in Table 4.1.

Parameter	Definition
Standard	IEEE 802.15.4
Simulation time	6 s
Tx current	17.4 mA
Idle listen current	0.02 mA
Rx current	18.8 mA
Rx-Tx current	0.02 mA
Voltage	3 V

Table 4.1: Simulation parameters for node power usage in OMNeT++.

The effects of OEN's discovery protocol are evaluated on energy consumption and latency. The simulation setup includes 5, 10, 15, 20 and 25 overlapped sensor nodes randomly deployed in an area of 100 m $\times$ 100 m. Each density represents the minimum number of deployed nodes in the defined area to have a communication between one node against 2, 3, 4, 5, 6 and 7 opponents respectively. 25 nodes in 100 m $\times$ 100 m cover one main agent against 6 or 7 opponents. Results show the average energy consumption and subscription time of 50 simulation runs for each density with random node deployments. The error bar denotes the standard deviation of the samples.

4.1.3 Results on Average Energy Consumption

Figure 4.2 shows the average energy cost of transmission of an agent against 2 to 7 opponents during the simulation period. As can be seen, the OEN discovery protocol consistently consumes more energy when the pool of opponents is increased.

The discovery protocol, however, has an insignificant impact on energy consumption (<0.01 J), and is a result of the continuous reception required for negotiation agents discovery. Once

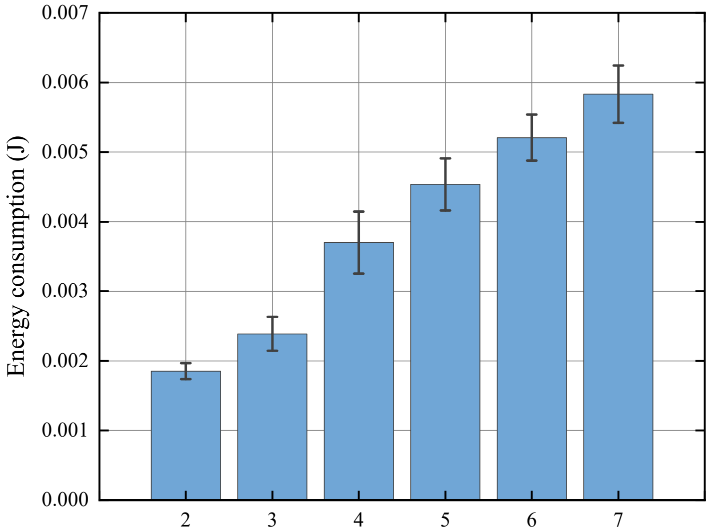

Figure 4.2: Average energy spent by one agent at the end of 6 seconds plotted against the number of opponents reached.

the agent broadcasts an advertisement message, it listens to receive the request messages of the neighbouring networks. During this process, the agents share details to decide whether they should associate with a possibility of cooperation, depending on the contribution each could give to the opponent network (by exchanging only context information). This step aligns the goals of the individual agents to find compatibility, thus ensuring that the networks can self-organise into one-to-one sets deciding how to cooperate, through a negotiation mechanism. The negotiation-based cooperation approach proposed in this thesis can facilitate the interaction and collaborative management in a wide range of applications and can lead to efficient coexistence of multiple co-located networks. For instance, the relatively minor increase in energy consumption is likely to be outweighed by the 41% increase in energy allocation that an agent may achieve when it reaches an agreement with a negotiation partner from a different network domain. The value of 41% corresponds to the amount of additional energy allocated by agent 2_1 using NBS in Subsection 3.5.4 (the agent increases its utility from 0.59 to 1).

4.1.4 Results on Average Subscription Time

The average subscription time measures the time elapsed from the broadcasting of the first OEN_ADV message to the time of the last OEN_ACCEPT message received in the network of the main agent. Figure 4.3 shows the average time spent in receiving/ transmitting mode from a neighbourhood of agents where 2 to 7 opponents subscribe to a main agent's cooperation initiative. Accordingly, every value indicates the latency introduced to a node's regular operation.

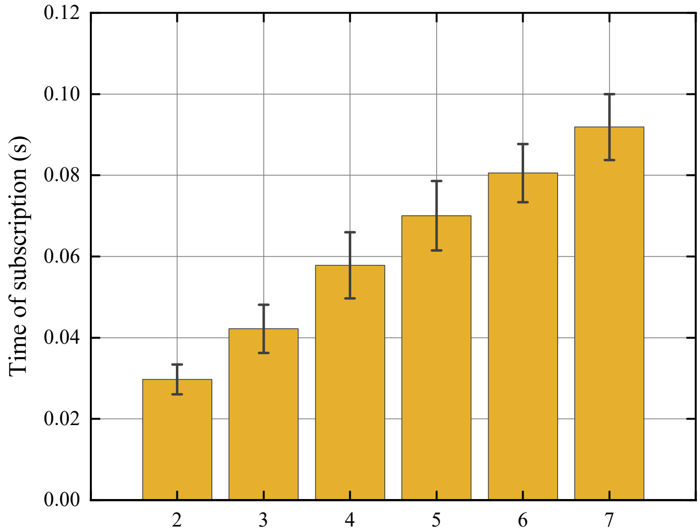

Figure 4.3: Average time spent by a neighbourhood of nodes where 3 to 8 agents form a set of agents with a cooperation incentive.

The execution time of the discovery protocol increases as the number of participants grows, however, the maximum increment corresponds to 60 ms even if the number of subscribed agents is almost four times higher (from 2 to 7 agents). The error bar length increments since the probability of collision increases in the medium where 5, 10, 15, 20 and 25 nodes share the same area.

These results describe the overhead of the discovery protocol to the methodology for establishing direct cooperation between networks. This step gives the agents the ability to detect devices with similar incentives to cooperate. The type of neighbour discovery is known as active discovery (i.e.: by broadcasting advertisement messages over different radio frequencies containing useful information about the node's properties that should be considered to interact with the advertiser). The result of this step is the identification of co-located devices or agents within 1-hop (1-hop neighbourhood) reach of the advertiser that have similar interests and are able to coordinate their tasks jointly. Such agents, however, do not know yet how their preferences will influence their need for cooperation. To cooperate with each other, the networks need to find an agreement through a negotiation process that gives them a mutual advantage but which also the agent improves its own performance. The energy negotiation model is described in the next section.

4.2 Energy Negotiation Model

In this section, bilateral, multi-issue negotiations, with an emphasis on the protocol and negotiation strategies are introduced. In bilateral negotiations, there are two agents with a desire to cooperate, but with conflicting interests regarding each given incentive for doing so. When the object in dispute is a single issue, the negotiation is referred to as single-issue negotiation. Otherwise, it is a multi-issue negotiation. A multi-issue negotiation is more common as well as challenging and complex in real-world domains. In such multi-issue negotiations, the agents should be able to negotiate agreements that are mutually beneficial for both parties. In contrast to single-issue negotiations, the complexity of the negotiation process increases as the number of issues increments, because of complex preferences over multiple attributes and the dimensionality of the solution space. For multi-issue bilateral negotiations in the domain of this work in EHWSNs cooperation, the protocol and strategies are the dominant concern.

An agent can adopt multiple strategies to create a proposal. This section describes the protocol and strategies used in this work following a generic framework for automated multi-issue negotiation [115]. The model is empirically evaluated in the next section to show a comparison with the results obtained by applying NBS. Following this, extensive simulations using multiple energy profiles and negotiation behaviours demonstrate how effective the mechanism is in comparison to NBS.

4.2.1 Notation and Assumptions

In this section, the basic concepts and notations related to the cooperation problem and the heuristic model that will be used throughout this chapter are presented. This includes the definition of networks, set of nodes, and agents of Subsection 3.2.1.

There are four fundamental parts in a negotiation model described by a heuristic approach: 1) the negotiation protocol or rules of interaction for the negotiating agents, 2) the definition of issues or objects in contention, 3) the utility function or agents' preference model, and 4) the negotiation tactics or offer generator functions that are applied during the bargaining process, which along with the utility function comprise the decision making apparatus the participants employ to act according to the negotiation protocol and reach their desired goals [74, 90].

As defined in 3.4.1, $o = (o(1), ..., o(n))$ represents the vector of issues (amounts of energy over each time slot of expected cooperation) to be negotiated in each negotiation round r, i.e. o is the energy flow offer. The agents in this domain only propose one offer in each round. Thus, $o^r_{1,1\to2,1}$ is a vector of values proposed by agent 1_1 to agent 2_1 at round r, where $o^r_{1,1\to2,1}(t)$ is the value of energy proposed from agent 1_1 to agent 2_1 for time slot t.

The issues in this domain maintain interdependencies between each other due to the use of the battery. For a time slot t, the energy flow (energy going out/into the agent) depends on how much energy an agent harvest or how much energy had been stored/withdrawn in previous time slots. The addition of such interdependencies increases the complexity of making a decision, even more, if nodes employ strategies that require them to learn about the opponent's model in order to solve the negotiation. The negotiation context (issues, deadline and initial negotiating agent) is known by both agents beforehand, and it remains unchanged during the whole encounter.

This work focuses on two-party, many-issue negotiations with a domain of limited resources (time, communication and processing among them). The number of messages exchanged between nodes is important, hence this work limits the negotiation to a short-term deadline. Little intervention by the process is required to minimise the effect of negotiation on sensor nodes' normal operations. Thus, a predefined maximum negotiation round r_{max} is used to model the deadline.

In OEN, then, the negotiation proceeds in a sequence of rounds $R = \{1, ..., r, ...r_{max}\}$ for a limited number of rounds r_{max}. In each negotiation round, an offer contains multiple issues that are negotiated simultaneously. Specifically, in this scenario, automated negotiation can complete in seconds, which makes time inappropriate to model the deadline. This is the main reason to select a discrete series of rounds for the negotiation mechanism.

The cooperation is envisaged over a finite period of time (e.g. 6 hours), which is divided into time slots of equal duration. The networks are able to pre-agree this criteria, e.g. If networks expect to cooperate for 4 hours, then they need to negotiate over an energy flow that must include 4 energy values.

Once the agents have determined the timing information (start time, duration of each slot t, and end time that the expected cooperation will last) over which they will negotiate energy resources, under the assumption they have synchronised clocks, the negotiation process between two agents 1_1 and 2_1 consists of an alternate succession of offers and counter-offers of values for the energy flow amounts. A complete description of the negotiation protocol and strategy employed in this thesis is given in the next section.

4.2.2 Negotiation Protocol

The protocol for the negotiation of energy is based on Rubinstein's alternating-offers protocol [73]. Such protocol is commonly adopted in a broad range of domains as part of their negotiation mechanism [150–152]. In this case, the protocol is used for bargaining between two neighbouring rechargeable nodes.

In a bilateral negotiation, both agents are willing to cooperate but have conflicting interests regarding their preferences (in this domain due to distinct batteries, power consumption and energy harvesting profiles). Then, agents have to negotiate and determine the most beneficial setup before cooperation. They need to agree on an energy flow that maximises their utilities.

The order of who begins the negotiation process is randomly selected during the pre-agreement phase. Usually, the buyer in other bargaining contexts is the one that starts the negotiation. However, since here the roles of the agents are the same, such order definition is proper of the domain. Then, the agent who was selected is the first to make a proposal. The agents can take actions only at certain times in the set of rounds $R = \{1, 2, 3, ..., r_{max}\}$. The agents involved have one turn per round r to respond to the current state of the negotiation. The following are the possible actions that an agent can perform:

- *offer*[o]. This action is executed from the start of the negotiation and whenever an agent rejects and offers counter proposals. An offer $o = (o(1), ..., o(n))$ is an offer of energy for each slot of time t. Its dimension depends on the expected cooperation time and length L of each time slot. An offer is valid if it respects the conventions agreed by the agents during the pre-agreement phase about time. If the expected cooperation time agreed by both agents is 3 hours starting at 06:00, and they set 15 minutes long for each time slot ($L = 15min$), then a valid energy offer would have 12 issues, $o = (o(1), o(2), o(3), o(4), o(5), o(6), o(7), o(8), o(9), o(10), o(11), o(12))$. where each issue corresponds to an amount of energy. For example, with the following $o = (5, 2.4, 6, 8, 9, 10, -4, 5, 6.7, -10, -5, -5)$ an agent is offering to contribute with 5 mWh at 06:00 while it expects to save 4 mWh at 12:00 from collaborative effort.

- *accept*[o]. When an agent i_j receives an offer o, it is able to accept the offer and reach a *provisional* agreement with the other party.

- *reject*[o]. When an agent i_j receives an offer o, it is able to reject it and opt out of the negotiation without any agreement.

- *confirm*[o]. When an agent i_j accepts an offer o, to confirm the provisional agreement, its opponent sends him a confirmation message which in turn must receive a confirm-acceptance reply to reach a final agreement. Otherwise, the negotiation fails and ends without a deal.

Figure 4.4 shows the alternating offers-based protocol for energy negotiation. One of the negotiating agents i_j starts the negotiation with an offer o to its opponent. Whenever an offer o is made, the opponent can *accept* it or *reject* it. If the offer o is accepted, then the bargaining ends and a *final agreement* is reached once the *provisional offer* is *confirmed* by the parties. If the offer is rejected, the agent with the turn can *opt out* of the negotiation and finish it without agreement, or it can propose a *counter-offer*, which again the opponent may accept or reject in the next round. The negotiation continues until a final negotiation round r_{max}. When one negotiating agent reaches a final round without a favourable response, or an agent rejects and opts out of the negotiation, or an agreement is found, the negotiation ends. In the first two cases, the negotiation fails and terminates with no deal possible.

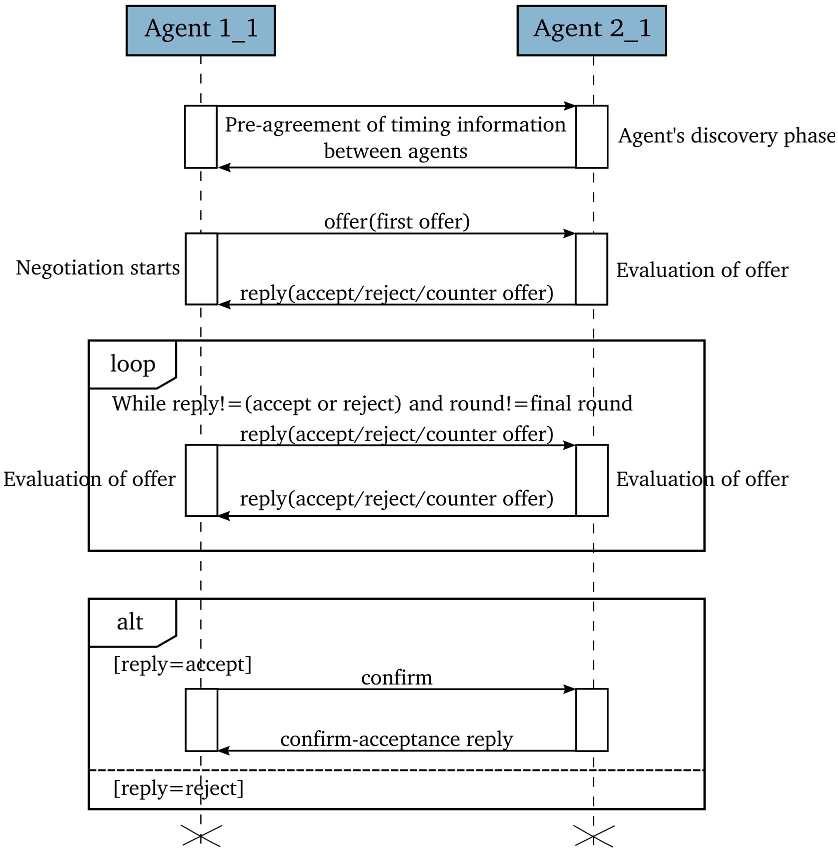

Figure 4.4: Sequence diagram of the negotiation process.

4.2.3 Negotiation Strategy in Opportunistic Energy Cooperation

As described in Section 2.4, negotiation heuristics are useful for the design of agents and the incorporation of a reasoning model. The reasoning model is based on computational approximations that produce outcomes close to Pareto-efficient solutions. A series of heuristics based on a single criterion such as time, resources and behaviour were defined by Faratin et al. [90]. Such heuristic functions, called tactics, can be combined to define a formal model of reasoning for negotiation strategy selection.

The main advantage of using heuristics in this domain is to model encounters between multi-domain nodes that are discovered opportunistically and have no information about the resources and preferences of each other. The heuristic functions allow delimitation of the search space of the solution and reduce required computation. Such heuristics, especially time-dependent concession strategies have been broadly used in several areas such as Grid environments and Cloud Computing [153–156]. The following components are part of the negotiation strategy evaluated in this domain.

4.2.3.1 The Conceding Strategy

Agents adopt time-dependent strategies [90] to determine the amount of concession required for each offer. The tactic indicates how rapidly the agent is likely to concede during the negotiation time. In the first round, the agents propose deals that give the highest utility to themselves. Afterwards, different agents may have different attitudes towards deadlines.

Rounds conduct the values of the negotiation issues, and the more rounds has passed the more pressure is induced and faster concessions are possible. The agent can adopt two behaviours: it may be impatient to reach a deal, so it concedes quickly and the offer rapidly changes to the reservation value (Conceder agent), or it may adopt a tougher strategy and maintain its initial proposal until it almost approaches the deadline (Boulware agent). Let $u^r_{i,j}$ denote the minimum utility value acceptable for agent i_j at round r. In the experiments, the following time-dependent function is employed as the concession strategy to model the target utility value (the amount of energy allocation desired for the period T) of an agent i_j at each round r of the negotiation:

$$u^r_{i,j} = min_{i,j} + (1-\alpha^r) \times (\sum_{t=1}^{n} E^c_{i,j}(t) - min_{i,j}) \tag{4.1}$$

where $min_{i,j}$ denotes the reserved value of agent i_j over T i.e., the minimal amount of energy an agent i_j can allocate for its consumption when o is null, found by Algorithm 1 in Section 3.3. The sum $\sum_{t=1}^{n} E^c_{i,j}(t)$ is the maximum amount of energy an agent can allocate to power its load over T.

Then, the target utility at each round is within the range $[min_{i,j}, \sum_{t=1}^{n} E^c_{i,j}(t)]$. A wide range of time dependent functions can be defined simply by varying the way in which α function is computed. In the experiments, α^r function is parameterised by the concession rate β, round of negotiation r and deadline r_{max} as follows:

$$\alpha^r = \begin{cases} \left(\frac{1-r}{1-r_{max}}\right)^{\beta}, & \text{if } \beta < 1 \\ \left(\frac{1-r}{1-r_{max}}\right)^{\frac{1}{2-\beta}}, & \text{if } 1 \leq \beta < 2. \end{cases} \tag{4.2}$$

The concession will depend on the strategy, which can be defined simply by varying the value of the parameter β in α^r. Following this function, the shape of the concession curve represents a human's negotiation behaviour. If $\beta < 1$, agent i_j adopts a Conceder behaviour; if $1 < \beta < 2$, the agent uses a Boulware tactic. Therefore, the function α^r as defined above is used to compute a target utility according to the round of negotiation and agent's behaviour. This information is then used to find the corresponding offer's values at each instant of time. The tactic is used to limit the search space of the solution and control the concession characteristic.

Along with the application of time-dependent tactics, the generation of the counter-offers considers the opponent's behaviour and compute the next offer based on the previous offer of the opponent. The offer generation strategy is modelled in subsequent Subsections.

The examinations are limited to these negotiation tactics, while the design of learning-based negotiation strategies would be more convenient in a dynamic environment of WSNs. However, this type of techniques is more feasible to implement in the presence of a robust and reliable intermediate entity. In the context of tactics, more complex behavioural tactics may be more appropriate, but they are less successful with short-term deadlines [90] and demand powerful devices to deal with the complexity of the negotiations [157]. In the proposing strategy below, the type of imitation tactic an agent can handle in this domain is described.

4.2.3.2 The Responding Strategy

Given an agent's utility value at round r, an agent can define its response for an opponent's offer in a way that there is no reduction in the amount of energy allocation the agent is expected to negotiate during its turn. Thus, if agent 1_1 receives an offer $o^{r}_{2,1\to1,1}$ from agent 2_1 at round $r < r'$, the interpretation of agent 1_1 defined as H at round r' for the opponent's offer is given by:

$$H^{r'}_{1,1}(o^{r}_{2,1\to1,1}) = \begin{cases} accept, & \text{if } u^{r}_{1,1}(o^{r}_{2,1\to1,1}) \geq u^{r'}_{1,1}(o^{r'}_{1,1\to2,1}) \\ reject, & \text{otherwise.} \end{cases} \tag{4.3}$$

Therefore, the agent 1_1 accepts the current offer made by agent 2_1 if the utility of this proposal is higher or equal to the amount of the utility that agent 1_1 will concede to in the next round.

If the offer is rejected, the agent in turn proposes a new agreement, which again the opponent may accept or reject in the next round. The negotiation will continue until an offer is accepted, a final negotiation round is reached, or the process is terminated by any of the participants (ending it with no deal possible). The strategy to generate offers in this domain of multiple issues is described next.

4.2.3.3 The Proposing Strategy

The agents' strategy for generating offers is implemented using the orthogonal strategy [113]. The reason to employ this strategy is its approximation of a Pareto-optimal bargaining solution over a multi-issue negotiation problem. Moreover, this offer projection strategy has a formal proof of convergence. Therefore, the heuristic guarantees the achievement of a final agreement in general automated multi-attribute negotiation, where the agents have no information about the utility function of their opponents and nonlinear utility spaces may be possible [107, 117].

The main idea behind the orthogonal strategy is to always select the point which is the closest (measured in the Euclidean distance) to its opponent's last offer on its indifference curve (i.e., the points that give the same utility for the agent). Let $o^{r-1}_{2,1\to1,1}$ be the last offer from agent 2_1 to agent 1_1 at round $r-1$. If agent 1_1 needs to generate a counter proposal that lies on the indifference curve C according to its target utility $u^r_{1,1}$ (obtained in Equation 4.1), then agent 1_1's offer at round r with the shortest distance to $o^{r-1}_{2,1\to1,1}$ can be calculated by

$$o^r_{1,1\to2,1} = \arg\min_{o\in C} \|o - o^{r-1}_{2,1\to1,1}\| \tag{4.4}$$

where $\|.\|$ denotes Euclidean distance. A nonlinear bound-constrained optimisation solver (fmincon) is used to return the offer at which the distance is minimized subject to the constraints described in Subsection 3.4.1.

A similar strategy is proposed in [112, 115]. The shortest-distance proposing mechanism is also classified under the category of alternating projection strategies. This is a feasible offer generation procedure to solve multi-issue negotiations with incomplete information. Similar to the orthogonal strategy of [113], the agent's offer at round r with the shortest distance to $o^{r-1}_{2,1\to1,1}$ can be calculated using equation 4.4. However, in this case, $o^{r-1}_{2,1\to1,1}$ represents the best offer for agent 1_1 among all the offers proposed by agent 2_1 in the past rounds, i.e. the offer made in previous rounds by agent 2_1 that yields the highest utility to agent 1_1. This strategy is also evaluated in the OEN framework.

In summary, the negotiation mechanism presented in this section allows an agent to represent its preferences and determine the desired utility level to generate a counter-offer accordingly. An agent makes use of the described negotiation model in order to fulfil the network objective of efficient energy allocation in a cooperative manner.

The outcome of the heuristic approach and the different behaviours an agent can adopt is compared with NBS. With NBS, agents declare to a trusted mediator their reservation utilities and utility functions to compute their agreement. The next section presents the experimental validation of the heuristic negotiation model by selecting some scenarios to show the agents' performance with multiple behaviours (Subsection 4.3.1). Further experiments are carried out to validate an observation in the results (Subsection 4.3.2). Finally, the comparison between OEN, the shortest-distance proposing mechanism and NBS performance is also analysed with extensive simulations (Subsection 4.3.3).

4.3 Experimental Validation

The first experiments in Subsection 4.3.1 are based on the node information and energy profiles described in Section 3.5. According to these results, a hypotheses is described and validated in

Subsection 4.3.2 with a larger dataset. Following this, extensive simulations are used to compare the performance of OEN with NBS (Subsection 4.3.3) and a similar proposing strategy.

As described earlier, the co-located nodes expect an insufficient energy level to trigger the need for OEN. This threshold anticipates an insufficient amount of energy in the battery and the absence or scarcity of energy harvested during the same period. However, the intermittent supply and different energy consumption profiles of nodes include periods where nodes can supply energy to each other. Thus, agents can share their harvested energy at some points in return for energy at other points in time. To clarify, if there is no feasible space of agreement, there is no cooperation. Given this set of possible cooperation deals, there are two agents, 1_1 and 2_1 trying to reach an agreement about the amounts of energy that can be logically shared.

Agents use the alternating-offers protocol and linear utility function described in Subsection 4.2.3.1 to compute the target value of energy allocation they want to achieve at each round of the negotiation. This target utility is used to find an offer o accordingly using the orthogonal search method described in Subsection 4.2.3.3 along with the constraints of Subsection 3.4.1.

4.3.1 Cooperative Scenarios

The first evaluation is used to compare the results obtained with NBS and reserved utilities from Chapter 3. This is tested to show the parameter model and their dynamics.
In this setup, the agents contemplate a period of cooperation that will last 24-time slots, each slot with a duration of 1 hour. At the beginning of the negotiation, the agents make offers that give the highest utility to themselves. For the analysis, a negotiation deadline of 10 rounds is set ($r_{max} = 10$). In the negotiation, an agent may adopt different negotiation behaviours with respect to different negotiation opponents: for the tough behaviour, β value is tested in 1.4, and the conceding negotiator is simulated with a β value of 0.05. The utilities for both agents are then computed for the next cases:

- Case 1: Both agents employ a Conceder tactic (β = 0.05 for agent 1_1 and agent 2_1).
- Case 2: Both agents employ a Boulware tactic (β = 1.4 for agent 1_1 and agent 2_1).
- Case 3: Agent 1_1 is tough (β = 1.4) while agent 2_1 concedes rapidly at the beginning of the negotiation (β = 0.05).
- Case 4: Agent 1_1 concedes rapidly at the beginning of the negotiation (β = 0.05) while agent 2_1 is tough (β = 1.4).

Figures 4.5 and 4.6 show the comparison of the utilities that the agents get without negotiation, by NBS and by the alternating-offers protocol and strategies described for energy sharing (OEN). For every case listed above (Case 1 - Case 4), agent 1_1 initiates with the first offer.

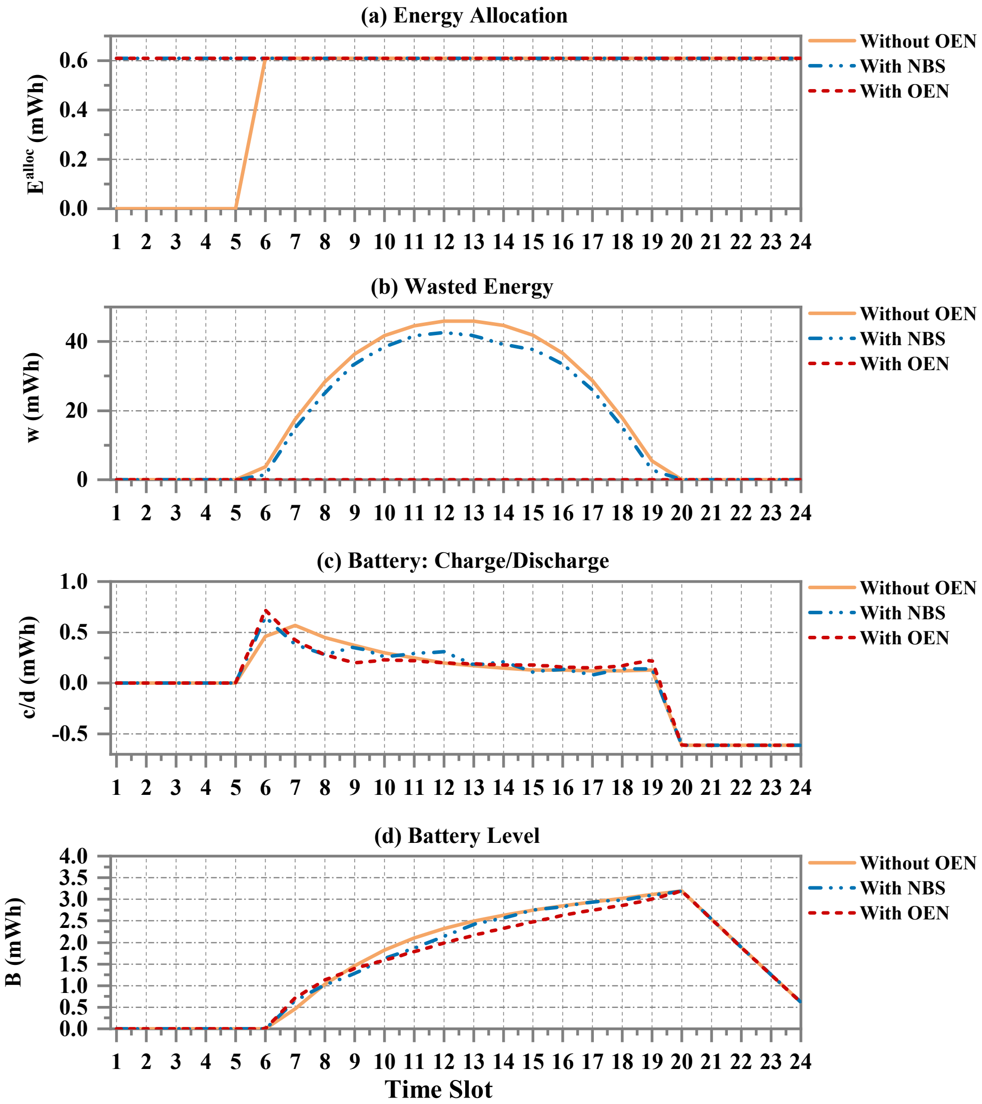

Figure 4.5: Agent 1_1: Results of utility maximisation without OEN, with NBS and with OEN.

Only two figures are shown for all the cases considered. This is due to the fact that no matter how low or high the concession shape β is varied (0.05 or 1.4), and which agent adopts any behaviour, the negotiation process ends with the same results when agent 1_1 offers the energy flow that maximises its utility and it starts the negotiation process. This scenario is an example of the cases where the energy availability between agents matches each other's need.

In this case, agent 1_1 has a distribution of energy that satisfies agent 2_1's need and 2_1 is also able to assist agent 1_1 in its lack of energy during periods 1-5. As a result, the utilisation of energy is maximised from 52.2 mWh (harvested energy used by agents 1_1 and 2_1 without cooperation) to 83.4 mWh by OEN while maintaining the application performance of both agents at the same rate at all times, i.e. their duty cycle is not affected. Then, the total energy saved via cooperation can be up to 7% for one day of the energy generated. The energy saved corresponds to the energy reallocated with OEN which would otherwise go to waste without negotiation. The reduction of energy waste is illustrated in Figures 4.5(b) Without OEN and Figure 4.6(b) with OEN.

Similarly to the results obtained by NBS in Chapter 3, the offer given by agent 1_1 to agent 2_1 in the first round of OEN assigns the maximum utility to both agents. Figures 4.5(a) and 4.6(a) illustrate the achievement of energy-neutrality for both nodes. This means that the nodes enable continuous operation when they decide to cooperate with each other. In this scenario, negotiation is successful because the interests of the nodes discovered opportunistically are not completely opposed. By cooperating, both of them optimise their power management through the extension of their networks' boundaries.

The efficiency and capacity values of the battery for agents 1_1 and 2_1 have no effect on the results. A fixed value is set for both during all the experiments in this section (70% and 708 mWh respectively). This assumption is reasonable according to the opportunistic cases dealt with in this thesis. The scenarios where nodes opportunistically enable the negotiation process to fulfil their energy allocation scheme are based on insufficient energy supply. Such a state can be either caused by low battery efficiency, ambient conditions, affected solar panel and wind turbines, or any obstacles in the environment. Besides, even if the battery efficiency of both nodes is modified in this scenario and set it to the maximum of 100%, the energy allocation of agent 1_1 remains the same while agent 2_1 slowly increases it from 40.56 mWh to 41.43 mWh without satisfying the full load of 68.64 mWh. Basically, the differences between ambient energy sources and heterogeneous energy profiles are the incentives of potential cooperation.

Figures 4.5(c), 4.5(d), 4.6(c) and 4.6(d) show the state of the battery during the day for each agent, where the battery level matches the dynamics of the charging and discharging flows and none exceeds the maximum battery capacity. The difference in the battery dynamic between NBS and the implementation of OEN depends on the negotiation's final outcome, i.e. the offer. The agreement represents the 24 energy values (in mWh) agreed for a day of cooperation and it is the energy flow from agent 1_1 to agent 2_1. The offer made using both solutions is shown in Figure 4.7.

The analysed results correspond to the negotiation outcome achieved by OEN when agent 1_1 starts the negotiation. The results vary if agent 2_1 initiates OEN. In this scenario, agent 2_1 starts the negotiation at round 1, the offer is then rejected by agent 1_1 who computes a new agreement to satisfy its target utility $u^2_{1,1}$ with the closest offer to the last offer made by agent 1_1. Such agreement is computed with the strategy described in Subsection 4.2.3.3. The negotiation ends at round 2, when agent 2_1 accepts the deal.

Table 4.2 shows the utility values of energy allocated over energy required, achieved when agents 1_1 and 2_1 reach an agreement on cooperative energy allocation. As noted, the order of alternating offers is important in this domain. Thus, the initiator of the negotiation must be defined randomly by the agents during the pre-negotiation phase.

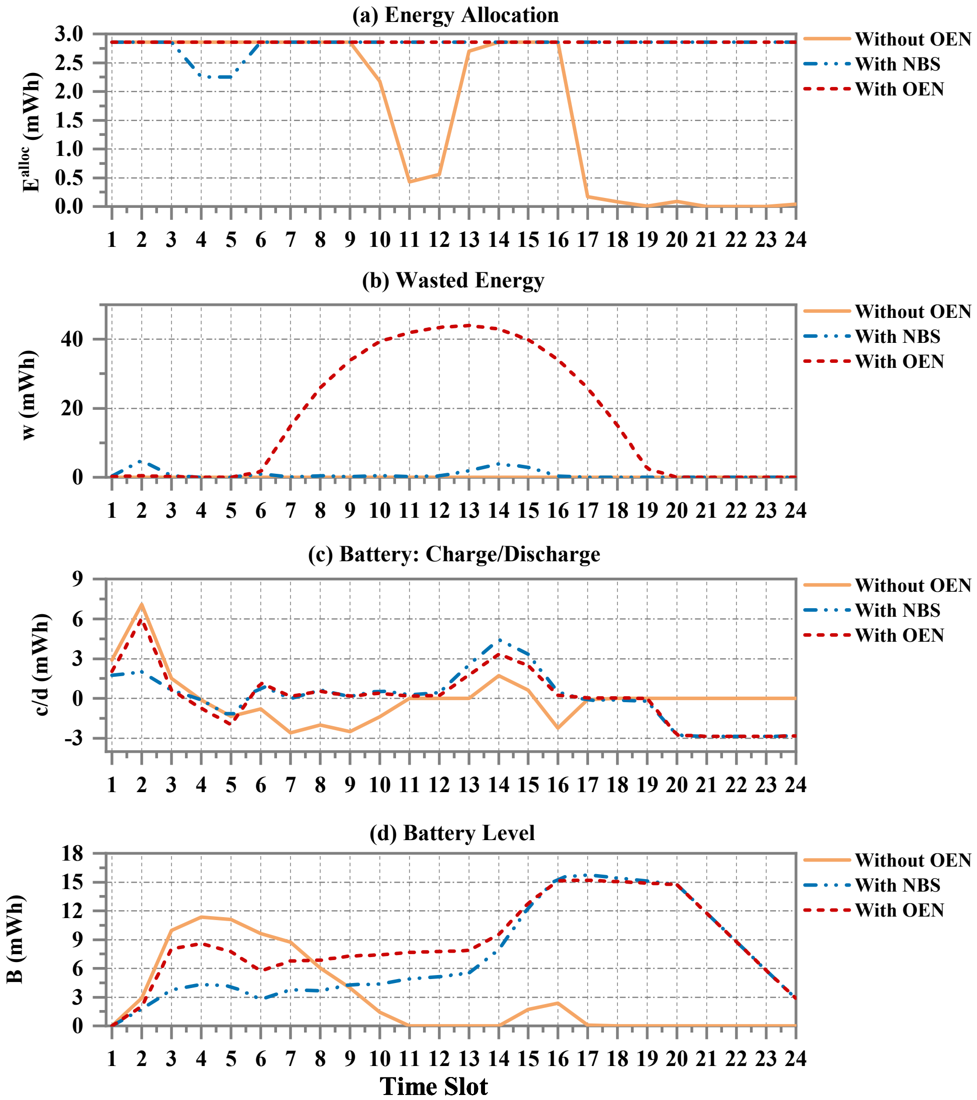

Figure 4.6: Agent 2_1: Results of utility maximisation without OEN, with NBS and with OEN.

Case	$u^2_{1,1}$	$u^2_{2,1}$
1	0.81	1
2	0.99	1
3	0.99	1
4	0.81	1

Table 4.2: Agent's utilities when agent 2_1 starts OEN.

Social welfare ($u_{1,1} + u_{2,1}$) is maximised whenever agent 1_1 adopts a tough behaviour. Without knowing agent 2_1's strategy, the best strategy of agent 1_1 is to concede less rapidly since it has a significant amount of energy to negotiate. Figures 4.8 and 4.9 show the utilities that both agents reach in case 3.

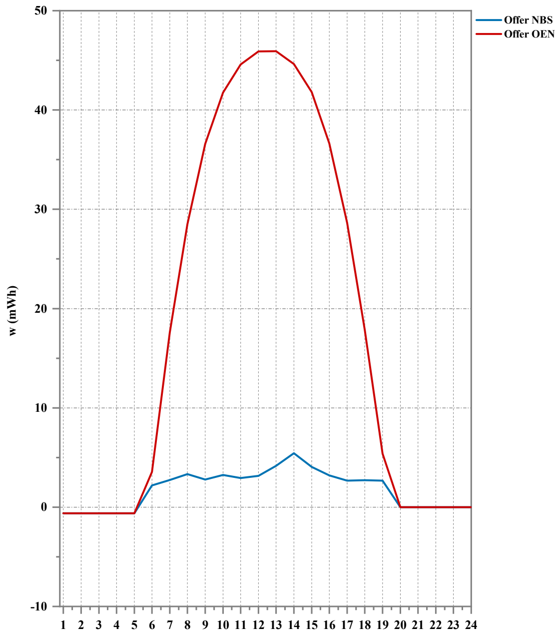

Figure 4.7: Offer made during OEN and offer found with NBS.

Both figures show the energy flow agreement reached with OEN and the effect on the model's parameters. The battery status matches the depletion and charge of the battery for the period analysed. As can be seen from the figures, the energy flow offer made by agent 1_1 exceeds the total load of agent 2_1, which lead to a waste of energy. This means that agent 2_1 will have to control the cooperative tasks, e.g. if they cooperate by relaying packets for each other, agent 2_1 will limit its traffic to agent 2_1's network and ask only for its packet forwarding capacity. To overcome this and minimise the wasted energy, nearest agents can benefit based on the percentage of remaining excess and decide how cooperation proceeds.

Now, to measure the performance of the heuristic solution when both agents have the same power consumption to satisfy, a slight change is made to match the load of agent 1_1 to agent 2_1 (both agents have a power consumption of 2.86 mW). The same energy harvesting profiles are used. This is evaluated by varying the negotiation behaviour of the agents. The concession shape β is set to 0.05 (Conceding agent) and 1.4 for tough behaviour. The deadline set for this negotiation is 10 rounds of alternating offers. The results for the same cases listed above (Case 1 to 4) are shown in Table 4.3 as the agent that starts the negotiation, who finishes it, behaviours, utilities and final round.

In these situations, the maximum achievable performance (energy neutrality) is only accomplished when the opponent is benevolent (Simulation 1 and 5) and agent 1_1 starts the negotiation process. In those cases, the behaviour of agent 1_1 is independent of the result, as long as agent 2_1 concedes, otherwise, the result is as simulations 3,4,7 or 8. Thus, in any case (Case 2 and 4) where agent 2_1 adopts a tough behaviour, it receives the lowest utilities. With such results, it can be seen that agents receive less utility when they remain reluctant to change its

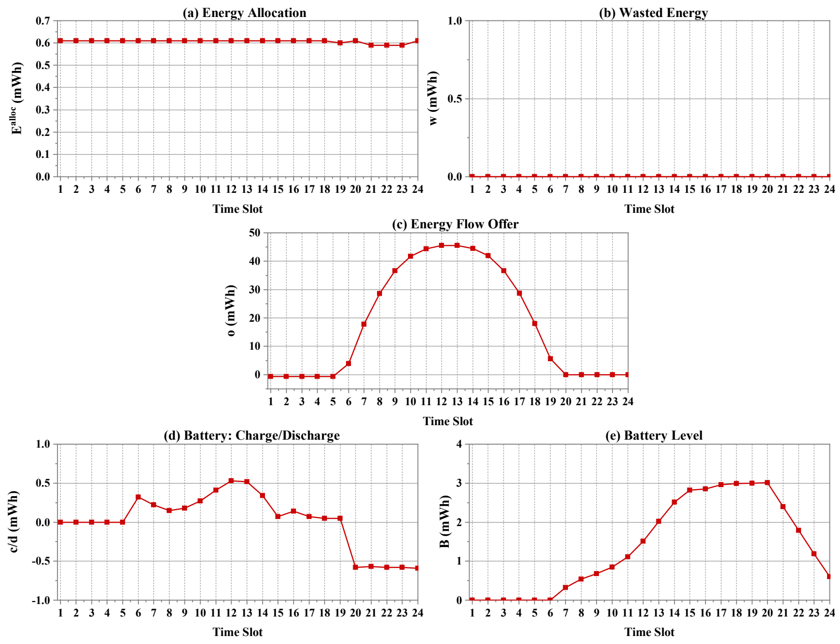

Figure 4.8: Agent 1_1: Results of utility maximisation when agent 2_1 starts OEN.

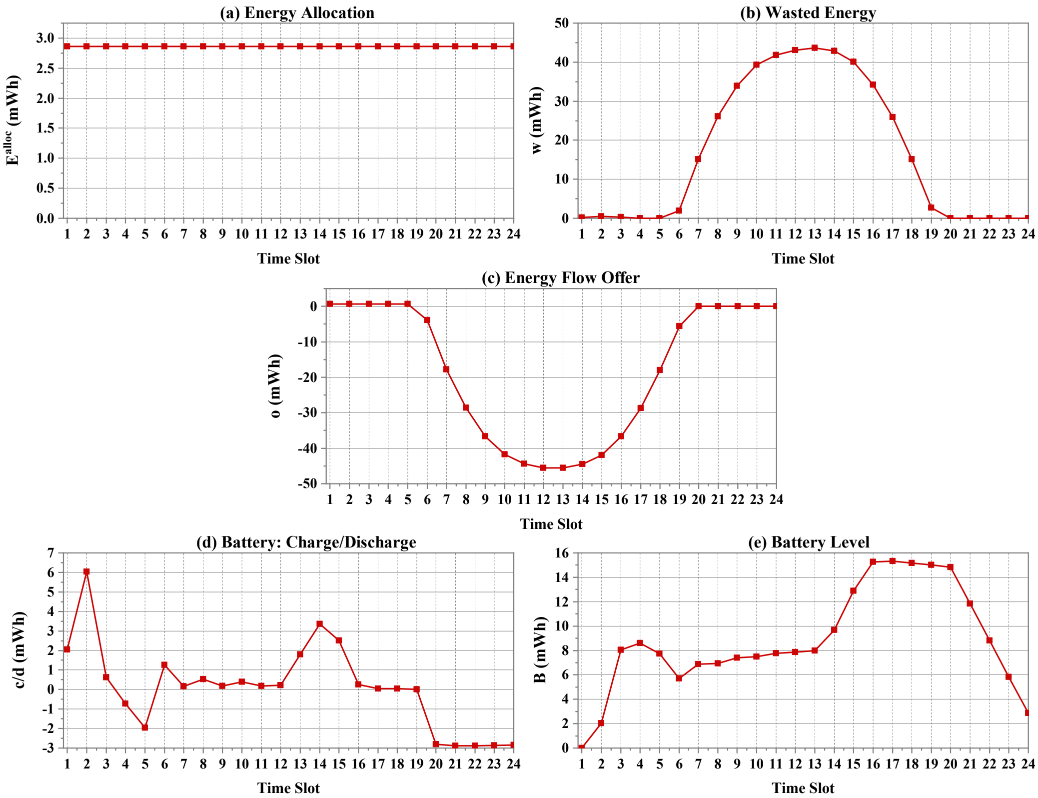

Figure 4.9: Agent 2_1: Results of utility maximisation when agent 2_1 starts OEN.

Simulation	First turn	Final turn	Case	$u_{1,1}$	$u_{2,1}$	Final round
1	1_1	2_1	1	1	0.92	1
2	2_1	2_1	1	0.81	0.94	2
3	1_1	1_1	2	0.91	0.73	8
4	2_1	1_1	2	0.92	0.66	9
5	1_1	2_1	3	1	0.92	1
6	2_1	1_1	3	0.92	0.61	5
7	1_1	1_1	4	0.91	0.73	8
8	2_1	1_1	4	0.92	0.66	9

Table 4.3: Comparison between different negotiation cases: Conceder vs Conceder, Boulware vs Boulware, Boulware vs Conceder, Conceder vs Boulware.

proposal and have less to offer (agent 1_1 has higher energy peaks than agent 2_1). The hypothesis is then, that agents should achieve better agreements when they adopt a Conceder behaviour and mimic the opponent's behaviour. The next section validates the hypothesis using a larger dataset for the experiments.

The required number of rounds is low (less than 10) in every simulation. The highest energy utilisation is given whenever agent 2_1 concedes faster at the beginning of the negotiation and agent 1_1 has the first turn. Since most of the parameters are the same for both agents (except the energy harvested), the available energy is a decisive factor in the establishment of cooperation. When sensors are energy-aware, spontaneous cooperation cannot take place and thus, negotiation is required to stimulate cooperation among directly interconnected nodes.

The presented results provide some insight on cooperation initiated by OEN. More simulations have to be conducted to benchmark the heuristic approach. The following subsection validates the hypothesis of cooperative behaviour. Next, a comparison with NBS tests a greater diversity of cooperative scenarios and the utilities reached by both solutions (NBS and OEN).

4.3.2 Hypothesis and Results

Based on the analysis of the initial results, the following observation is presented for evaluation:

Hypothesis 1. The best behaviour is to adopt a Conceder strategy when agents negotiate and have less to offer. While it is best to adopt a Boulware strategy when agents negotiate and have major peaks in their energy availability.

In order to run the experiments to validate the hypothesis, weather data of 2017 from Weather Underground about wind speed (v) and solar radiation (G_b) from PVGIS is selected for an average day of April, November and December to meet the requirement of diversity in energy generation. The corresponding measurements for solar irradiance and wind speed are shown in Figure 4.10.

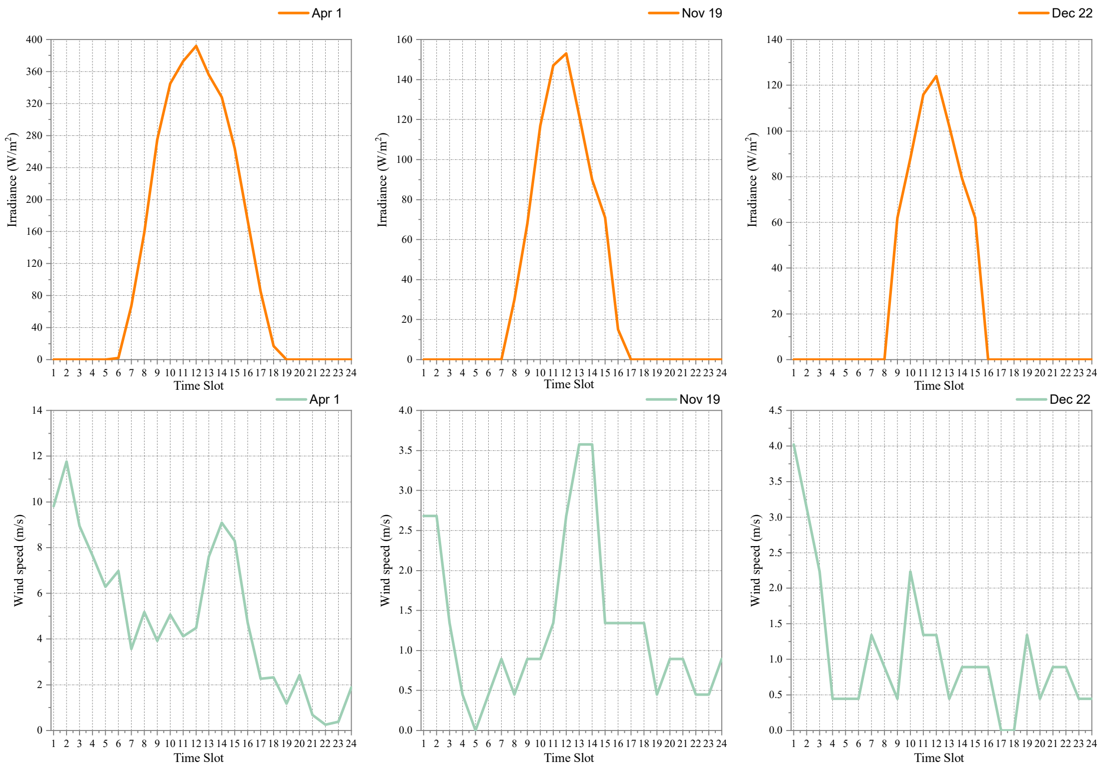

Figure 4.10: Solar irradiance and wind speed used to compute energy harvesting values.

This hypothesis is validated through 1000 simulations with 1000 records of 24-time slots each to negotiate at each run. Thus, the agents negotiate 24 energy issues for a day of cooperation at every encounter. For each simulation, the initiator is selected randomly. The deadline for every negotiation is set to 10 rounds. Each agent simulates an energy source type: Agent 1_1 is a node with a load of 5% duty cycle (3.0143 mWh) and a solar panel, while its opponent agent 2_1 has a wind turbine and consumes the same load. Thus, the energy availability is determined by the energy harvested by each. As can be seen from the figure, agent 1_1 has a higher amount of ambient energy. The parameters to compute each energy harvesting profile are described in Table 4.4. At each simulation, a weather dataset (G_b and v) is randomly selected from April, November and December to compute an energy harvesting profile using values drawn from Table 4.4. The maximum battery capacity and efficiency are set as 708 mWh and 0.7, respectively for both agents.

Agent 1_1		**Agent 2_1**	
Parameter	Value	Parameter	Value
Solar panel dimension (A)	$\sim U(4,25)$ cm^2	Swept area (A)	25 cm^2
Solar panel efficiency (f)	$\sim U(0.6,1)$	Wind turbine efficiency (f)	$\sim U(0.6,1)$
Perturbation (p)	$\sim U(0.8,1)$	Perturbation (p)	$\sim U(0.8,1)$
$E_{1,1}^{hrv} = G_b \times A \times f \times 1000 \times p$ [mW]		$E_{2,1}^{hrv} = 0.5 \times \rho \times A \times v^3 \times f \times 1000 \times p$ [mW]	

Table 4.4: Parameters to generate multiple energy harvesting profiles using irradiance (G_b) and wind speed (v) from April, November and December.

Table 4.5 shows the average utility of the deals reached when agents act with the following negotiation behaviours:

- Case 1: Both agents employ a Conceder tactic (β = 0.05 for agent 1_1 and agent 2_1).
- Case 2: Both agents employ a Boulware tactic (β = 1.4 for agent 1_1 and agent 2_1).
- Case 3: Agent 1_1 is tough (β = 1.4) while agent 2_1 concedes rapidly at the beginning of the negotiation (β = 0.05).
- Case 4: Agent 1_1 concedes rapidly at the beginning of the negotiation (β = 0.05) while agent 2_1 is tough (β = 1.4).

Case	$u_{1,1}$	$u_{2,1}$
1	0.82 ($\pm$ 0.09)	0.36 ($\pm$ 0.16)
2	0.74 ($\pm$ 0.06)	0.33 ($\pm$ 0.14)
3	0.8 ($\pm$ 0.11)	0.32 ($\pm$ 0.18)
4	0.74 ($\pm$ 0.05)	0.35 ($\pm$ 0.15)

Table 4.5: Average agent's utilities when agent 1_1 has higher peaks of energy than agent 2_1 and both adopt multiple behaviours. Standard deviations are indicated in parenthesis.

Although on average, the agent utilities show similarities, Case 1 is the most successful scenario for both agents. The similarity is due to the fact that both negotiators employ the orthogonal strategy, which mimics the behaviour of the opponent at every counteroffer. The difference, however, is observed in the percentage of agreements made at every case. Figure 4.11 illustrates the percentage of agreements achieved at every scenario.

From the figure, it is clear that the choice to apply conciliatory tactics seems to have a significant influence on the agents' resolution of conflicts. Conversely, the adoption of tough behaviour by both parties makes significantly fewer deals than Case 1. In fact, this approach leads to lower commitments than all the other cases. Having this performance measure, the final utility value achieved by the agents for every case is detailed in Figure 4.12.

For agent 1_1, which has more major peaks of energy availability than agent 2_1 the best strategy is to adopt a Conceder behaviour even if its preferences aim to a lower cooperation effort from the opponent. The second-best scenario for agent 1_1 is to be tough against a Conceder. Boulwares get high individual utilities when they manage to make deals. The average gain is up to 14.5% when agents do not give ground easily during negotiation against a Conceder negotiator.

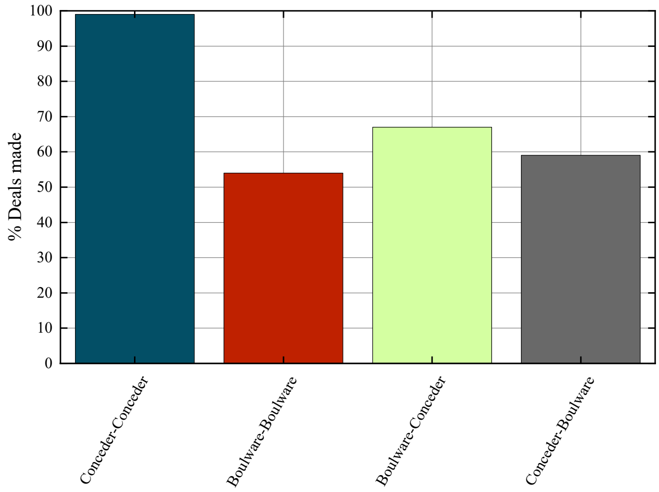

Figure 4.11: Percentage of agreements reached at every case: Conceder vs Conceder, Boulware vs Boulware, Boulware vs Conceder, Conceder vs Boulware.

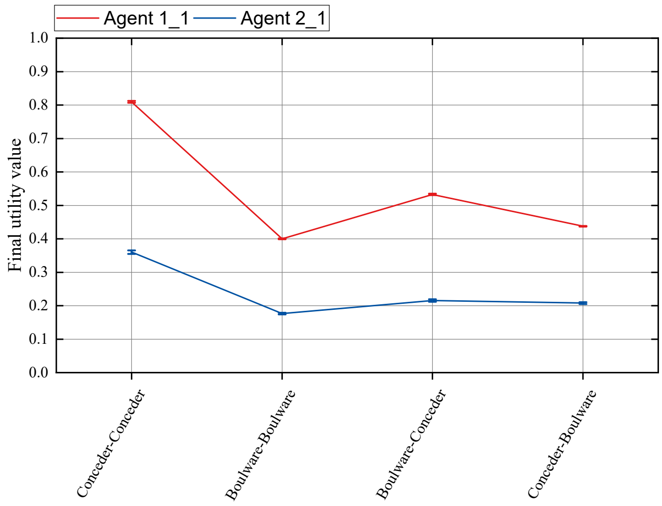

Figure 4.12: Final utility value for varying negotiation behaviours.

Thus, the results invalidate the hypothesis. According to the negotiations using real weather data, the adoption of a Conceder tactic can improve the allocation of energy for agents when they decide to cooperate regardless of the energy availability.

4.3.3 Performance Evaluation of OEN

In order to obtain the following results and compare OEN to NBS, a feasible space of energy flow agreements must exist between the agents. From this set of agreements, OEN and NBS find an energy flow agreement to satisfy the preferences of each part. Then, the optimality of the agents implementing OEN is evaluated with the average distance to NBS agreement.

Thus, the same datasets from Weather Underground and PVGIS about solar radiation and wind speed for an average day of April, November and December employed in 4.3.2 are used here to meet the requirement of feasible intersection points. The energy harvesting profiles for each agent are computed using the parameters from Table 4.4.

The performance evaluation of OEN is conducted over 1000 negotiations with agents under realistic load conditions, which duty cycles vary between 1% to 5%. Using the power consumption model described in Subsection 3.2.2 with the following parameters for agent 1_1 and agent 2_1:

Parameter	Value
Active current	20 mA
Sleep current	5 µA
Voltage	3 V
Duty cycle	1% - 5%
Battery capacity	600 mAh
Battery efficiency	0.7
Slot length	1 hr

Table 4.6: Agent's parameter values for power usage and battery storage model.

The energy consumption of each agent varies between 0.6149 mWh to 3.0143 mWh. For the OEN approach, the agent that starts the negotiation is randomly assigned between agent 1_1 and agent 2_1 at each simulation run. Each negotiation involves 24 issues in contention. The behaviour of the agents is selected randomly by choosing the concession rate β between 0.05 and 1.4. The maximum number of rounds for each negotiation is set to 10.

The efficiency of OEN is also compared with the shortest-distance proposing mechanism described in Subsection 4.2.3.3. Thus, the main goal of the simulations is to compare the performance of the agents in the following situations: without cooperation, with near-optimal energy allocation after negotiation using the OEN heuristic approach, results with the shortest-distance proposing mechanism, and finally when the agreement is computed using NBS. The error bars denote the standard error to the mean.

Figure 4.13 shows the average utilities achieved by the agents during 1000 simulations and the percentage of deals made by each negotiation mechanism.

As depicted in the figure, OEN with the orthogonal strategy performs much better than the shortest-strategy: OEN achieves 20% average increase in conflict resolution (% Deals made). From these results, the maximum utilities that each agent 1_1 and 2_1 can make are computed using NBS. These utilities are 0.87 and 0.5217 for each agent respectively, while OEN and shortest-distance exhibit 77% on average of the utilities reached by NBS (0.77 and 0.34 for agents 1_1 and 2_1 respectively). The centralised approach also increases the number of agreements, by reaching a 100% of deals made.

Considering the efficiency of OEN in the percentage of deals made, the agents can increase their utility on average up to 14.12%. While shortest-distance optimises up to 10% on average the utility of the agents that applied this solution.

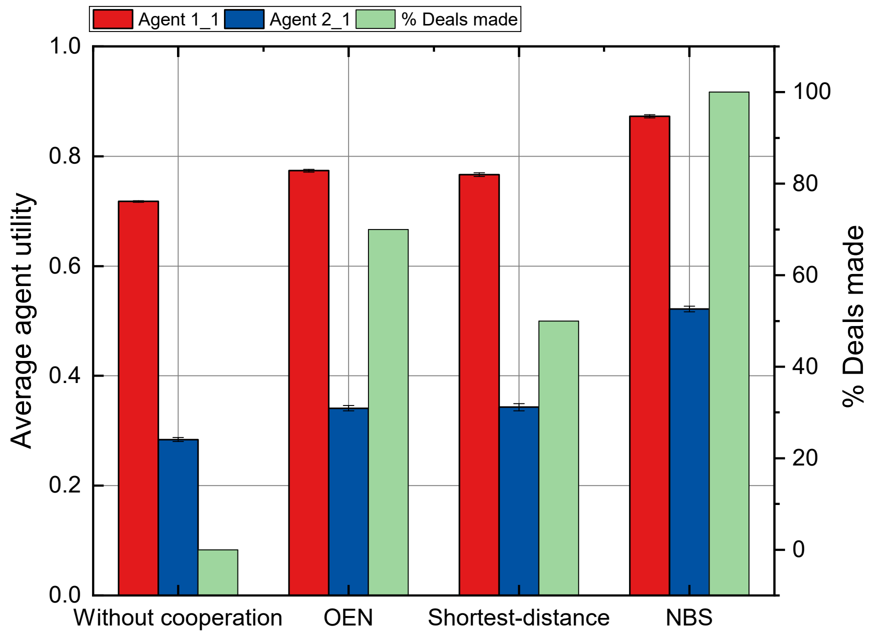

Figure 4.13: Comparison of agents' utilities without cooperation, implementing OEN, using the shortest-distance proposing mechanism and NBS.

Although the performance of the orthogonal strategy in OEN and the achieved with shortest-distance are very similar, the implementation of the last strategy demands the storage of all previous offers and identification of the best offer among this group of deals. The identification may include sorting or computation of the maximum utility associated with each offer stored in the agent's memory. Thus, the simplicity and efficiency of OEN using the orthogonal strategy are promising and its design makes it more feasible to implement cross-network cooperation in the domain of WSNs than NBS or shortest-distance.

In terms of energy neutrality conditions, OEN does not enforce the achievement of energy neutrality agreements by the agent. Instead, it relaxes this requirement to measure the utility of an agent given the power management strategy of OEN and study the potential of cross-boundary energy transfer. In order to accomplish this condition, the cooperative model of energy allocation described in 3.4.1 should include the following constraint:

$$B_{i,j}(t) \geq E^c_{i,j}(t) \tag{4.5}$$

The constraint (4.5) enforces that the residual battery level $B_{i,j}(t)$ at the beginning of each t must be bigger than the energy consumption $E^c_{i,j}(t)$ of the agent i_j at time slot t. With this, the negotiation strategy of an agent will remain in a high aspiration level every time an offer is made on a round. As proved earlier, a tough behaviour may result in less energy allocation agreements.

Thus, the proposed model with the conceding strategy without considering the constraint 4.5 is evaluated to measure the number of energy neutrality agreements reached by the solutions.

Figure 4.14 illustrates the percentage of energy neutrality agreements achieved by an agent during the simulations run in this section. As observed, although energy neutrality is not conditioned in the offers proposed in OEN, the mechanism is capable to reach this most desirable outcome. The percentage of energy neutrality deals are 20%, 19% and 34% for OEN, shortest-distance and NBS respectively. This metric is used to show that nodes are capable of satisfying continuous operation during a day using OEN.

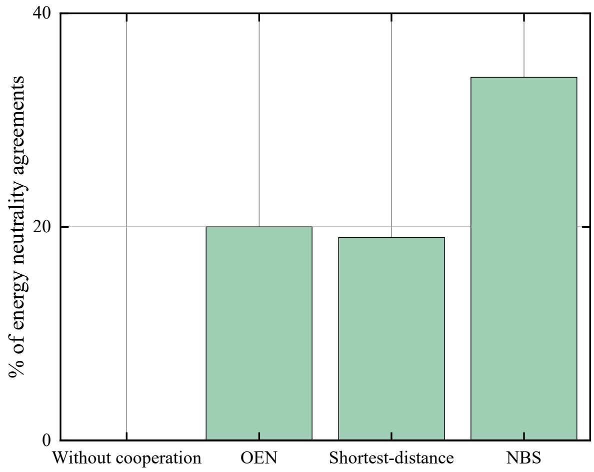

Figure 4.14: Comparison of energy neutrality agreements without cooperation, implementing OEN, using the shortest-distance proposing mechanism and NBS.

As indicated above, OEN manages to improve the utilities of an agent by up to 14% compared to its energy allocation when it does not share energy. This performance can be increased by exploiting co-located nodes that allow better energy allocation agreements to the agent. Therefore, an agent should be able to anticipate the best potential partner with which to start the negotiation process at every opportunistic encounter. Instead of choosing randomly an opponent between the co-located nodes, an agent should implement a better decision making policy for the selection of the most suitable opponent.

The next chapter presents a solution that considers relevant policies when the partner selection is modelled as a bandit problem. The performance of partner selection using different bandit policies against the best-fixed selection strategy and random selection strategy are compared over a series of experiments and scenarios.

4.4 Summary and Discussion

At the beginning of this chapter, the establishment of OEN was described. The results demonstrated the increment of energy cost and average subscription time when an agent initiates OEN with 2 to 7 opponents. Even with 7 agents reached to start an opportunistic energy negotiation, the measures of these attributes are minimum ($<$0.01 J and $<$0.1 s).

Optimising multiple co-located networks, each with a variable number of energy preferences, is a complex problem that has received little attention in the WSN research community. Using existing negotiation frameworks, this thesis proposes a heuristic model to solve the cooperation problem, with the goal being to optimise the use of harvested energy by the agents and enable cross-network power management.

The negotiation protocol and tactics are described in this chapter, followed by an analysis of the first results, which are also contrasted with the ones obtained with NBS. Then, a hypothesis is formulated to express the intuition from these results. The goal of the analysis is to compare the utility of agents in cases where they adopt two strategies: Conceder and Boulware, and the importance of energy availability in their behaviour. The results of the analysis show that the adoption of a Conceder strategy for generating offers achieves the deals with the highest utility regardless of the energy availability. Finally, the efficiency of OEN in comparison with shortest-distance and NBS is also presented.

According to the results, from the listed requirements in Section 1.3, the negotiation-based cooperation approach proposed in this chapter includes the development of a decentralised, short-term deadline multi-issue negotiation model under incomplete information. The orthogonal strategy guarantees conversion and near-optimal energy allocation solutions. Specifically, an agent can reach on average 77% of the utility achieved by NBS. The percentage of deals, however, reached by OEN is of 70% in contrast to 100% of NBS, and 50% of shortest-distance. Once considered this metric, the improvement of an agent's utility can be up to 14.12% for a day of cooperation with a co-located agent.

This difference is low. Thus, agents need to estimate the negotiation outcome or best negotiation partner to start an OEN. With these estimations, an agent is able to decide if negotiation is convenient. Then, the requirement of adaptability from Section 1.3 can be covered to provide a negotiation approach that addresses the unpredicted environment where the agents of WSNs are deployed. Using a reinforcement learning technique, the next chapter describes the policies that can be useful in the context of partner selection.

Chapter 5

Partner Selection: A Multi-Armed Bandit Based Approach

Previous results were found assuming an agent selects randomly an opponent from a co-located network. However, this process can be improved. Consulting every opponent found during the discovery of neighbouring devices about their offers for cooperation is also not efficient in this domain. This results in a waste of resources, as the exchange of proposals might be performed unnecessarily when considered the high heterogeneity of nodes in terms of energy profiles. The dynamism of the environment where: new agents may join a network, others may leave, the environment may change, is also a factor to consider in such an open domain as WSNs. Furthermore, the constant exchange of information between multiple agents leads to communication overhead. Thus, the selection of the opponent with the highest possibility for reaching a good agreement is very important in this context. To cope with such dynamism the agents must be able to adapt their behaviour according to the changing circumstances. Given this, this work relies on reinforcement learning, and more specifically on Multi-Armed Bandits, to allow networks to learn their best partner in multi-agent negotiation. Section 5.1 describes the partner selection problem in negotiation for efficient long-term energy allocation. Section 5.2 addresses the problem formally using bandits. Then, Section 5.3 introduces the proposed learning policies and their practical implementation for the partner selection problem in WSNs. Finally, Section 5.4 provides the experimental setup and empirical evaluation of the developed model using MAB learning.

5.1 Partner Selection Problem for Long-term Energy Allocation

Suppose a neighbourhood of agent i_j is defined as $\Omega_{i,j}$, such that $\Omega_{i,j} \subseteq I_i$, and that the agent i_j knows about all other agents in its 1-hop neighbourhood. Thus, a neighbourhood is a subset of agents in N_i that control sensors with overlapping radio range. A link or edge between two agents (i_u, l_v) represented by $e_i(i\_u, l\_v)$, can communicate offers and traffic directly from agent i_u to an agent in a different network l_v, where $1 \leq i, l \leq m$ and $u \in I_i, v \in I_l$.

Each agent i_j can maintain a map of energy conditions across its neighbourhood $\Omega_{i,j}$ employing broadcast messages used by the routing protocol. Consensus via local communication with their neighbours is only required when more than one agent discover a lack of energy at the same time in the same neighbourhood. If that is the case, agents must decide the order in which negotiations take place. In that regard, the following assumptions about the neighbourhoods, the pre-negotiation communication phase and the transmission properties are made:

(a) Neighbourhoods do not necessarily have the same number of members, and each agent i_j belongs to only one neighbourhood $\Omega_{i,j}$.

(b) OEN is proactive: each agent i_j periodically broadcasts HELLO messages that contain its energy status and the list of its current neighbours along with their status. In this way, each agent i_j can maintain a map of energy conditions across its neighbourhood $\Omega_{i,j}$. To eliminate OEN overhead, HELLO information is introduced onto the broadcast updates required by the used routing protocol.

(c) If multiple agents discover a lack of energy at the same time in the same neighbourhood, this work assumes that agents are assigned a priority level and rotate with time, but how such assignment will be performed is out of the scope of this work.

(d) No packet loss occurs during cross-network communication, which is relevant for the delivery performance of offers during the negotiation process. This is a valid assumption since no loss is observed under the introduction of ODI architectures [147].

(e) The cost of negotiation is negligible as compared to the energy aimed by a node to win after negotiation (e.g. collected data is valuable to communicate). This assumption is reasonable in negotiations with pre-established short deadlines [158]. Moreover, [57] shows that the energy cost to maintain ODI functionality is also insignificant.

Given the agent's utility function for period T of cooperation, defined in equation (3.3):

$$u_{i,j} = \sum_{t=1}^{n} E_{i,j}^{alloc}(t) \tag{5.1}$$

where $E_{i,j}^{alloc}(t)$ is equal to:

$$E_{i,j}^{alloc}(t) = E_{i,j}^{hrv}(t) - c(t) + d(t) - w_{i,j}(t) - o(t) \qquad (c_{10})$$

Since the global objective is to maximise the total energy allocation in the WSN over a period of cooperation T, the network utility function is maximised when the sum of all agent's functions for the energy allocation in the network is maximised. Then, the utility function for the whole network N_i is:

$$u_i = \sum_{j \in I_i} u_{i,j} \tag{5.2}$$

where I_i as defined before denotes the set of agents in network N_i. Thus, the global objective is to maximise the total energy allocation in the WSN over T. This may imply that the communication between all agents in the network is required. However, this is not the case for OEN, which considers the suboptimal approach where interactions between agents are performed only with the agents in the same neighbourhood Ω. Each agent i_j can maintain a map of energy conditions across its neighbourhood $\Omega_{i,j}$ employing broadcast messages used by the routing protocol and avoiding the possible communication overhead. Given this, the objective of network N_i is to maximise the total energy allocation over the negotiator agents on each neighbourhood $\Omega_{i,j}$.

Ultimately, the goal of each agent at every opportunistic encounter is to decide and choose a partner among the alternative nodes it has discovered in its immediate neighbour networks so as to maximise its energy allocation in the long term. Thus, such an opponent must be selected as the most prospective negotiation partner with whom the expectation of successful negotiation and the achievement of the best agreement are the highest. In order to realise this approach, the agent seeking a partner first needs to learn the performance of all the neighbouring agents expected to cooperate. Then, the decision of the negotiation partner involves a trade-off: the negotiation with an opponent provides feedback about its effectiveness (**exploration**), but the collection of that feedback ignores the immediate benefit of selecting a partner that is already known to be effective (**exploitation**).

The selection method of an appropriate partner must be able to learn the dynamism of the environment and adversarial setting introduced by the negotiation behaviour of the opponents. To solve the partner selection problem, a MAB model for each agent within the network is proposed. In probability theory, MAB learning provides a theoretical framework for sequential learning and decision-making to address the trade-offs between exploration and exploitation under uncertainty. Unlike traditional partner selection methods, which require historical data to calculate the outcomes of negotiation and predict the possibility of successful negotiation, this work uses MAB to estimate the profitability of each agent and develop an online (or adaptive) scheme able to tolerate dynamically changing environments and adversarial conditions without prior knowledge.

5.2 Multi-Armed Bandits for Partner Selection in WSNs

In this section, the K-armed bandit problem is defined formally. The following shows how the bandit problem can model the partner selection in negotiation for an efficient long-term energy allocation in WSNs. In doing so, existing policies are discussed and later their comparison is reported.

5.2.1 The Multi-Armed Bandit Problem

The multi-armed bandit problem originally proposed by Robbins [159] refers to the gambler's dilemma. Correspondingly, the goal of a gambler is to maximise the total rewards earned through a sequence of lever pulls over a row of slot machines. Specifically, a set of K machines (arms) is available to the decision maker. At each trial, a gambler must choose which of these arms to play. To keep the terminology of MAS consistent, from here onwards the term gambler is replaced by agent, and the lever pulling action of the gambler is specified as an action of that particular agent. Without any prior knowledge on the machines' profitability, the agent can still collect partial information while it observes the reward of each chosen arm. Such information can be used to estimate the revenue of the machines. It thus becomes a dilemma, between *exploiting* the machine that has the highest expected reward or *exploring* the set of different machines to gain more information and learn about their reward density. The fundamental challenge in bandit problems is to define the pulling strategies (also referred to as policies) for decision making in situations under uncertainty to trade-off between exploration and exploitation. A MAB learning model is particularly useful to model agents that learn a hidden reward distribution while maximising their gains.

Formally, let $Tr = \{1, 2, ..., \mathrm{Tr}\}$ be a set of sequential trials and $a(tr)$ the action of an agent at trial tr, which raises the reward $r_a(tr)$. An agent's objective is to maximise the sum of its observed rewards as follows:

$$\max \sum_{tr=1}^{Tr} r_a(tr) \tag{5.3}$$

As such, it is clear that the agent has to choose a policy (i.e. a sequence of actions) that maximise the total rewards earned through a sequence of trials.

The performance of the policy applied by an agent at a given trial is measured in terms of *regret*, defined as the expected loss of applying the policy with respect to the maximal expected reward by a policy assumed to be optimal. Given the stochastic nature on the reward processes, this notion of expected regret is often considered. However, in this domain, a different concept of regret is incorporated, suitable for the adversarial MAB problem of our environment. This notion of regret, known as *weak regret*, is described in the next subsection.

The perception of optimality and bandit policies vary according to the environment. The following subsection presents a description of the practical application of MAB and the existing policies applied to the specific problem of partner selection.

5.2.2 Multi-Armed Bandits Formulation for Partner Selection in OEN

In this domain, the environment is adversarial. Unlike classical stochastic MAB problems whereby the rewards are independently drawn following a fixed but unknown distribution, for the adversarial MAB problem, no statistical assumption about the generation of rewards is made. Instead, the rewards are chosen by an adversary. The applicability of the solution also depends on the type of adversary. If the adversary chooses the reward ahead of the actual selection process, it is known as an oblivious adversary. Whereas, if the reward is simultaneously chosen by the opponent with the agent's choice, an adversary is called a non-oblivious adversary. The second category describes our case.

An adversarial MAB formulation is a natural fit for modelling the research problem of this thesis, where an agent and opponent interact to solve their conflicts and the opponent is adaptive. More precisely, in this context, the outcome of a negotiation between one agent and its opponent forms the reward value that the MAB model gets by selecting a partner for opportunistic energy negotiation. In OEN, an agent has an incentive to negotiate but may adopt different behaviours based on its preferences and observations at any time. The preferences of the networks vary according to their energy availability which is influenced by the amount of energy harvested during each time slot. The networks can then adopt a responsive attitude towards their environment using conceding strategies during the negotiation.

Since agents have to negotiate with incomplete knowledge of the opponents and have no control of the environmental factors affecting the outcomes of the negotiation, it is very difficult to estimate a distribution for the rewards. This thesis focuses on environmental changes such as varying energy availability, which influences the different patterns of the agent's negotiation behaviour. Furthermore, the topology of the networks is also intrinsically dynamic as sensors may fail, move, or enter in sleep/active state. An agent can also reject a negotiation encounter or be added opportunistically at any time. Thus, this variant of the MAB problems is considered here, where the stochastic assumption about the processes of rewards is removed and their realisation rely on the agents involved, their status, preferences and negotiation behaviours.

Repeated bilateral negotiation encounters are considered over a finite number of trials Tr where three or four WSNs overlap within a geographical area. In each trial tr, there are two or three agents that belong to different networks in the immediate neighbourhood of the main agent. The main agent needs to select one opponent between these two or three agents, as the most preferred negotiation partner to reach energy cooperation agreements that maximise its energy allocation. The action is easy to identify then, for each agent i_j in a wireless sensor network N_i that needs to start an OEN, an action of agent i_j at trial tr denoted as $a_{i,j}(tr)$, corresponds to the election of a negotiation partner (e.g. $a_{1,1}(1) = \{negotiation\_partner : 2\_1\}$) among a set of K opponents. The action is constant over time since this work only contemplates bilateral negotiations ("one-to-one") as a decentralised decision-making process to not require a mediator. The negotiation also includes short-term deadlines to avoid transmission overhead.

Given this, let $r_{i,j}(tr)$ be the linear reward function of agent i_j for each trial tr, defined as the amount of energy allocation reached on agreement at the OEN encounter tr (Equation (5.1) with a selected partner $a_{i,j}(tr)$ from a set of K opponents, the objective of network N_i to maximise the total energy allocation over a number of trials Tr, can be formulated as follows:

$$\max \sum_{tr=1}^{Tr} \sum_{\Omega \in I_i} \sum_{j \in \Omega} r_{i,j}(tr) \tag{5.4}$$

Therefore, the network objective is to maximise the sum of reward functions of all agents on each neighbourhood Ω, from the OEN encounter 1 to the trial Tr. Thus, each agent i_j's chosen action (i.e. the chosen negotiation partner) will determine the value of the global network objective.

Once the action an agent can perform is defined, and the reward function associated with each action is clear, the partner selection problem in OEN of each agent i_j can be reduced to a MAB problem. The agent's goal then is to efficiently maximise the expected total rewards against the adaptive environment and the adversarial opponent or equivalently, to minimise the cumulated loss over time, i.e. the energy that an agent doesn't get when it misses the chance to cooperate with the best partner.

In the setting where agents have no prior knowledge about the preferences of their opponents, and the outcomes are affected by unexpected environmental factors, the achievement of low-regret bounds (i.e. high performance) is not possible with any deterministic policy (especially for the non-oblivious adversary case). Alternatively, state-of-the-art policies in an adversarial setting are assessed, which aim to minimise a regret with respect to the best-fixed strategy in hindsight, i.e., by having access to the history of negotiation's outcome against every opponent. This weak regret is common in similar situations in which it is impossible to learn the optimal (adaptive) strategy [160, 161], mostly because the payoffs are adversarially decided by the opponent. Thus, although the optimal strategy cannot be learned, the best-fixed strategy in hindsight becomes feasible to analyse from the history of previous negotiations. Consequently, the cumulative expected regret over Tr represented by R_{Tr} with respect to the optimal fixed strategy is:

$$R_{Tr} = max \sum_{tr=1}^{Tr} r_{i,j}(tr) - \mathbb{E}\left[\sum_{tr=1}^{Tr} r_{i,j}(tr)\right] \tag{5.5}$$

Where the first term describes the cumulative reward by the best-fixed strategy over trials Tr and the second part corresponds to the total expected reward achieved by the policy applied in the system.

The next section describes three well-known policies for this problem. Following this, the experiment scenarios are defined and the performance results for these algorithms are analysed.

5.3 Mixed Policies

In order to handle the partner selection problem, this thesis makes use of adversarial bandit algorithms. The contribution of this chapter consists of adopting these policies to the domain of WSNs and compare them in three practical settings to study their performance.

The policies ε-Greedy, EXP3, and FPL-UE are selected as the bandit strategies for the experiments. These three bandit algorithms explicitly make use of an exploration parameter, they are widely used in the MAB literature and have proven to obtain sub-linear upper regret bounds with an appropriate choice of the exploration factor.

5.3.1 ε-Greedy Action Selection Strategy

A well-known and low-complexity heuristic policy for the bandit problem is the ε-greedy action selection strategy. The ε-greedy strategy is sketched in Algorithm 2. The policy selects at each trial tr an action with uniform random probability for a fraction ε of the trials (exploration), and choose the best arm (exploitation) with a probability $1-\varepsilon$ (Steps 4 and 6 respectively). The specification of the exploration factor ε is made based on the experiment, i.e. there is no standard value that fit-for-all scenarios.

Algorithm 2: Algorithm ε-greedy for each agent i_j

Input : $\varepsilon \in [0,1]$, opponents $1,\dots,K$;
Output: Negotiation partner $a_{i,j}(tr)$

1 Initialisation: $\hat{r}_k = 1, pulls_k = 0, rewards_k = 0$ for $k = 1,2,\dots,K$;
2 **for** $tr \leftarrow 1$ **to** Tr **do**
3   **if** $\sim U(0,1) \leq \varepsilon$ **then**
4         $a_{i,j}(tr) \sim U\{1,K\}$;
5   **else**
6         $a_{i,j}(tr) = \underset{k\in\{1,\dots,K\}}{\arg\max}\{\hat{r}_k\}$;
7   **end**
8   Receive reward $r_{i,j}(tr)$ as $E^{alloc}_{i,j}(t)$ for all t in negotiation against selected partner $a_{i,j}(tr)$;
9     $pulls_k = pulls_k + 1$ where $k = a_{i,j}(tr)$;
10     $rewards_k = rewards_k + r_{i,j}(tr)$ where $k = a_{i,j}(tr)$;
11     $\hat{r}_k = \frac{rewards_k}{pulls_k}$ where $k = a_{i,j}(tr)$;
12 **end**

Given this, it can be seen that the estimated reward ($\hat{r}_k$) of a selected action is updated using its cumulative reward ($rewards_k$) and the number of times the action k has been executed ($pulls_k$). ε-Greedy adds some randomness when deciding between negotiation partners: instead of relying always on the best partner, it randomly explores other opponents with a probability ε.

5.3.2 Follow the Perturbed Leader with Uniform Exploration (FPL-UE) Algorithm

The policy considered here is based on the online prediction scheme Following-the-Perturbed-Leader (FPL) [162], which has efficient treatment of problems with a linear cost function by following the perturbed leader. The original algorithm only works for oblivious adversaries and focuses on choosing the action of minimal cost by observing the loss incurred of each selected action. The goal of this work, instead, is to efficiently maximise the total rewards an agent can successively achieve against an adaptive and adversarial opponent. To address this problem, a novel strategy for repeated interactions called Follow the Perturbed Leader with Uniform Exploration (FPL-UE) [160] is employed. In this approach, the learning algorithm proposed by Neu and Bartok [163] is extended by introducing uniform random exploration for the reward maximisation scenario (Algorithm 4). Similar to ε-Greedy, the selection of a probability (ε in ε-Greedy, λ in FPL-UE) determines the exploration rate of the pulling strategy. That is, the agent will uniformly randomly choose a negotiation partner with a probability λ (Step 8) and select the partner that reaches the maximum estimated reward perturbed by the noise factor z_k (Step 5) every $1-\lambda$ of the cases. The efficiency of the mixed strategy FPL-UE on finding the best partner from a set of opponents within the repeated opportunistic encounters between networks is measured.

FPL-UE makes use of Geometric Resampling (GR) (Algorithm 3) in order to compute the estimated reward for the chosen action at every trial (Algorithm 4, Steps 11-12). The application of GR in our setting is shown in Algorithm 3. Basically, GR measures the reoccurrence where simulated a, denoted as ã, may appear. Thus, K_val_k represents the reciprocal of the probability of action k (p_k^{-1}), i.e. K_val provides a 1-in-M scale for probabilities, where M is a finite value that bounds the number of samples. For example, the reciprocal of 0.01 is 100, so an event with probability 0.01 has a 1 in 100 chance of happening.

Algorithm 3: Algorithm GR

Input : $M \in \mathbb{Z}^+$, $a_{i,j}(tr)$;
Output: $K\_val_k \in \mathbb{Z}^+$

for $i \leftarrow 1$ **to** M **do**
  Repeat steps 3 $\sim$ 9 in Algorithm FPL-UE once to sample ã;
  if $i < M$ **and** $\tilde{a} = a_{i,j}(tr)$ **then**
    $K\_val_k = i$;
  else
    $K\_val_k = M$;
  end
  if $K\_val_k > 0$ **then**
    break;
  end
end

Algorithm 4: Algorithm FPL-UE for each agent i_j

Input : $\lambda \in [0,1]$, $\eta \in \mathbb{R}^+$, $M \in \mathbb{Z}^+$, opponents $1,...,K$;
Output: Negotiation partner $a_{i,j}(tr)$

Initialisation: $\hat{r}_k = 0$ for $k = 1,2,...,K$;
for $tr \leftarrow 1$ **to** Tr **do**
    Set flag $\in \{0,1\}$ such that $flag = 0$ with prob. λ;
    if $flag$ **then**
        $a_{i,j}(tr) = \underset{k \in \{1,...,K\}}{\arg\max}\ (\hat{r}_k + z_k)$;
        where $z_k \sim exp(\eta)$ independently for $k = 1,2,...,K$;
    else
        $a_{i,j}(tr) \sim U\{1,K\}$;
    end
    Receive reward $r_{i,j}(tr)$ as $E^{alloc}_{i,j}(t)$ for all t in negotiation against selected partner $a_{i,j}(tr)$;
    Run GR($M, a_{i,j}(tr)$) to estimate p_k^{-1} as K_val_k;
    $\hat{r}_k = \hat{r}_k + K\_val_k \cdot r_{i,j}(tr)$ where $k = a_{i,j}(tr)$;
end

5.3.3 Exponential-weight Algorithm for Exploration and Exploitation (EXP3)

Unlike FPL-UE, EXP3 employs the value of the probabilities for each action more explicitly. The partner selection strategy using EXP3 is described in Algorithm 5. At each trial, EXP3 chooses a partner $a_{i,j}(tr)$ according to the distribution p (Step 4) learned from the iterations. EXP3 as FPL and ε-Greedy is a mixed strategy that introduces uniform randomisation into the action selection process. Once the action has been determined, the received reward is used to update the weight value of the chosen action (Steps 5-7), which affects proportionally to the probability of each action in the next trial (Step 3) (i.e. the higher the current estimate is, the higher the probability an agent chooses that action). Thus, at each trial, EXP3 updates the value of the distribution p, and it defines the action with higher probability and vice versa. Although EXP3 is classified under the category of MAB algorithms with partial information, that is, only the reward of the selected action can be observed, the update of its weight affects proportionally the weights of each respective arm. According to this, EXP3 selects at each trial the best-estimated action and provides an updated probability as the learning process continues, i.e. it guarantees that an agent can efficiently adapt to different environmental situations. The exploration rate, as in FPL-UE and ε-Greedy is given parametrically and affects the efficiency of the algorithm as well (γ in EXP3). It is important then, to analyse first the definition of this parameter before any comparison is executed.

5.4 Experimental Validation

This section describes the goal of the experiments, the description of the scenarios and their implementation, followed by the evaluation of the policies.

Algorithm 5: Algorithm EXP3 for each agent i_j

Input : $\gamma \in [0,1]$, opponents $1,...,K$;
Output: Negotiation partner $a_{i,j}(tr)$

Initialisation: $w_k = 1$ for $k = 1,2,...,K$;
for $tr \leftarrow 1$ **to** Tr **do**
  Set $p_k = (1-\gamma) \cdot \frac{w_k}{\sum_{k=1}^{K} w_k} + \frac{\gamma}{K}$ for $k = 1,2,...,K$;
  Draw $a_{i,j}(tr)$ randomly according to the probabilities $p_1, p_2, ..., p_K$;
  Receive reward $r_{i,j}(tr)$ as $E_{i,j}^{alloc}(t)$ for all t in negotiation against selected partner $a_{i,j}(tr)$;
  $\hat{r}_k = \frac{r_{i,j}(tr)}{p_k}$ where $k = a_{i,j}(tr)$;
  $w_k = w_k \cdot \exp\left(\frac{\gamma \cdot \hat{r}_k}{K}\right)$ where $k = a_{i,j}(tr)$;
end

5.4.1 Goal of the Experiments

The goal of the experiments are:

- Apply MAB learning to the setting of partner selection between multiple sensor networks for an efficient energy allocation in the long term.
- Compare three state-of-the-art policies for the adversarial MAB problem presented here, using as a baseline the best-fixed strategy in hindsight and the uniform random selection of a partner in each OEN encounter.
- Evaluate through extensive simulations the performance and validate the theoretical properties of the online prediction policies in a practical case study under different circumstances.

5.4.2 Experiment Scenarios

These simulations assume four authorities that deploy their sensor network in the same geographic area, in such a way that there may be between three to four distinct agents within overlapping radio coverage, i.e. for each agent, there is a pool of K parties formed by 2 or 3 opponent agents from which an agent can choose one partner to initiate a bilateral negotiation. As already mentioned, in the context of partner selection for OEN, agents in the pool may be viewed as arms. An agent must decide between these 2 or 3 agents which arm is expected to provide the best payoff. This setup is suitable for the experiments conducted; however, the pool of arms can be formed with any number of nodes, greater than two (depending on the memory limitations) to evaluate the MAB algorithms.

As previously described, the networks of this study periodically report readings to the sink. These networks are typically deployed for long-term operation, and their design constraints are application-dependent, also based on the monitored environment. This implies that if there is a pool of agents (arms) from which to select a negotiation partner, each arm will have unique characteristics that will determine its reward. This reward will depend on the negotiation outcome, which is directly affected by the negotiation strategy used by each party and their mutual zone of agreement. A mutual agreement relies on the energy availability of the arms and their ability to meet the current aspirational demand of the other agent.

Regarding the dynamic nature of these networks, besides taking into account their varying status (due to node failures, time-delays, active/sleep modes) that define their network topology, this work considers changes in their attitude towards negotiation (Conceder/Boulware tactics) and environmental conditions that modify the energy availability of the agents involved.

Thus, the following possible situations where the MAB model is applicable are examined. These scenarios define the changes that characterise the dynamic and heterogeneous domain of WSNs. This thesis focuses on environmental changes such as varying energy availability, which influences the different patterns of the agent's negotiation behaviour, and also instances of network topology variation.

Cooperative scenario. All the agents are Conceder negotiators. In proposed approaches to enable cooperation in multi-domain sensor networks [12, 13, 43], sensor nodes are assumed to be spontaneously cooperative on routing tasks in order to fulfil their application requirement. The utility function in these studies is characterised by the effective gain of minimising node energy consumption. The battery-powered networks represented in these works find that the equilibrium state with the highest payoff (where the lifetime of the sensors is the highest) consists of cooperative strategies. Similarly, this first scenario assumes the Conceder strategy for the generation of offers, as a cooperative effort. Thus, the cooperative behaviour of an agent is represented here as concessions quickly performed at the beginning of the negotiation.

Multiple behaviours. The opponents adapt their negotiation behaviour according to their energy availability, which is determined by the weather conditions. In these experiments, if the agent requires more energy than the amount it can provide to its opponent, it adopts a tough behaviour, otherwise, it employs a Conceder strategy. The functions in Subsection 4.2.3.1 are used to model the concessions. In multi-authority WSNs, a resource-constrained node may be reluctant to forward packets received from other network domain, or to do any other task on behalf of an external network to save its own resources. When an agent is aware of its power level [9, 14, 164, 165], it adapts its strategy to avoid being exploited by selfish decisions.

Dynamic topology. The networks change their topology. This thesis seeks for an efficient learning method that finds a trade-off between exploring and exploiting the available options of opponents by jointly considering the dynamically changing environment and varying network topology. The changing environmental characteristics are depicted by the ambient energy sources and their wide temporal variation that may also control the agents' behaviour (as in the previous case), while the varying network topology is taking into account as well. Some of the existing works on multi-domain cooperation either assume a static network topology [16] or consider dynamic node operation and topology independently [11, 42, 43]. In practice, topology changes are more frequent in a sensor network and can be attributed either to node mobility, failure or status. The networks that this thesis is studying are not mobile networks. However, a dynamic topology is considered in terms of node failures, the new addition of nodes, as well as nodes in different states such as active and sleep.

In each scenario, the energy availability of an agent is determined by its energy harvested. Each agent in the pool of arms may have an associated reward, which corresponds to the outcome of the negotiation. To show the dynamism of the domain the simulation time period is divided into intervals called epochs, and each epoch lasts a number of trials Tr. Each trial involves an OEN interaction between two agents. The characteristics of each party are constant along with a fixed number of trials, or epoch. At the end of an epoch, the features amongst arms change to set a different optimal partner (which is unknown by the agent during the selection). In this work, a preferred opponent is one that has more energy availability and a Conceder behaviour. The feature of energy availability changes per epoch in all situations, and the behaviour varies in the situations of multiple behaviours and dynamic topology. All runs in all scenarios involve 5000 trials. Four cases are considered: long epochs or static environment (Tr = 5000 iterations), moderately dynamic epochs (Tr = 1000), dynamic epochs (Tr = 500) and extremely dynamic epochs (Tr = 200). The length of epochs is obviously not known by the agents. The goal of an agent is to maximise its total reward over these trials, by finding the partner with the highest expected payoff. This determination is accomplished by observing the reward to know the efficiency of the chosen opponent, and thus, learn which opponents are the most efficient ones.

5.4.3 Design of Experiments

Now the conditions used to alter the environment and describe the values set for the energy availability and behaviour of the agents in every situation are formulated. The simulation of topology changes is also described below.

5.4.3.1 Energy Harvesting Profiles

In all three scenarios, the weather data collected from PVIGIS and Weather Underground from the year 2017 is used to generate the energy harvesting values for each agent. For the main agent, energy harvesting values from solar radiance were computed, while the set of opponents are simulated with energy values from wind speed. Similar to Subsection 4.3.3, from Table 4.4 the values of solar panel dimension, its efficiency and wind turbine efficiency are selected over the given ranges. There is no perturbation value set for these records, (p=1). Every energy value corresponds to the same day, same period. The records are used for 5000 trials, where negotiations of 6 issues each are considered, i.e. At every trial, the agents exchange offers to reach an agreement of cooperation for 6 hours.

5.4.3.2 Energy Consumption Profiles

The duty cycle of each agent is set uniformly random between 1% to 5%, which defines its load using the power consumption model from Subsection (3.2.2). The following parameters are used to compute the energy consumption profile of each agent, voltage = 3 V, sleep current = 5 μ A, active current = 20 mA. Each time slot is equal to 1 hr. Thus, the load for each agent is randomly selected from the following set $\{0.61, 1.21, 1.81, 2.41, 3\}$ [mWh].

As previously mentioned, a fixed value for the battery maximum capacity and its efficiency is set for all the simulated agents (708 mWh and 70% respectively).

5.4.3.3 Negotiation Parameters

Each negotiation includes a deadline of 5 rounds. For every interaction, the first to generate an offer is randomly selected. The number of negotiation issues in every trial is 6. The agents use the heuristic framework described in 4.2 to generate offers, evaluate them and make concessions.

5.4.3.4 Simulation of Environmental Changes

In every scenario, the energy harvested by the agents in the pool is modified in order to simulate environmental changes [1] that affect the performance of the energy source and affect the energy availability of every agent in the pool. These conditions determine a setting where one opponent is the best choice in every possible negotiation. Any setting with different conditions also shows the same broad patterns in the result of the simulations. The information about the characteristics of the opponents (how quickly they change per epoch and how they differ) is unknown to the agents. If there are three agents in the pool of arms, three different environmental conditions are simulated:

- Condition 1. First opponent is the best option.
  - First opponent: $E^{hrv}_{i,j}$ is not affected ($p = 1$).
  - Second opponent: $E^{hrv}_{i,j}$ is reduced to 40% ($p = 0.4$).
  - Third opponent: $E^{hrv}_{i,j}$ is reduced to 10% ($p = 0.1$).
- Condition 2. Third opponent is the best option.
  - First opponent: $E^{hrv}_{i,j}$ is reduced to 10% ($p = 0.1$).
  - Second opponent: $E^{hrv}_{i,j}$ is reduced to 40% ($p = 0.4$).
  - Third opponent: $E^{hrv}_{i,j}$ is not affected ($p = 1$).

[1] Energy harvested can be affected by multiple causes as obstruction of power source, weather conditions, solar panel and wind turbine efficiency. [8, 144, 166]

- Condition 3. Second opponent is the best option.
  - First opponent: $E_{i,j}^{hrv}$ is reduced to 40% ($p = 0.4$).
  - Second opponent: $E_{i,j}^{hrv}$ is not affected ($p = 1$).
  - Third opponent: $E_{i,j}^{hrv}$ is reduced to 10% ($p = 0.1$).

The experiments consider pools with two or three opponents. The agents that have two options in the set of arms will have only two situations where the first or second opponent is the best option, respectively.

The **cooperative scenario** follows the conditions described above, and the strategic behaviour of each agent in the set of opponents is not affected no matter how low its energy availability is. All the agents concede more rapidly at the beginning of the negotiation with a 0.05 concession shape value.

For the **multiple behaviours scenario**, the tactic of an agent is modelled using three concession shape values (β): 0.05 for a Conceder agent, while 1.4 and 1.9 model a Boulware strategy. In this case, the same conditions described above accompanied by the varying tactics are followed: the best option has a Conceder behaviour 0.05, while the rest of the set will have 1.4 and 1.9 respectively.

The third **dynamic topology** scenario exhibits the conditions of the multiple behaviours scenario plus the assumptions made on the networks' topology. Every 20 trials, this simulation assigns a probability of 0.4 to allow the absence of any opponent chosen uniformly random. This represents the dynamic behaviour of the network topology, where the absence can be seen as an agent's rejection of being part of an OEN, an agent's failure, or an agent in the sleep state. As a result, every 20 opportunistic encounters, any node from the pool of opponents may be unavailable.

Finally, in order to capture the dynamic nature of the environment, the four cases described in subsection 5.4.2 are simulated: static characteristics over time (1 epoch), moderately dynamic changes (5 epochs), dynamic (10 epochs) and extremely dynamic case (25 epochs). When the environment changes its epoch, it uniformly randomly chooses one of the three conditions specified above (Condition 1 - Condition 3). If there are two agents in the pool, then only two conditions are swapped.

5.4.4 Comparison of the MAB Algorithms

5.4.4.1 Selection of Exploration Factor

The algorithms used in this study condition their performance on the election of an exploration rate. In order to make a comparison between them, it is important to carefully choose the exploration factor, ε in ε-Greedy, λ in FPL-UE and γ in EXP3. Thus, this work first evaluates the policies over 25 epochs with 200 trials each, where one agent is assigned randomly between 2 and 3 opponents. The negotiation is excluded from these tests, instead, the three conditions rewards are sampled from a Bernoulli distribution with probability $1/K$: 1 represents a succesfull negotiation and maximum payoff, while 0 indicates no agreement between the parties. Then, in every trial an agent will have 2 or 3 opponents, where at least one of them will give the agent a reward. This condition is maintained over the epoch. The results shown in Figure 5.1 average the payoff over 100 simulations for each exploration factor, between 0 and 1 in 0.1 steps. The figure illustrates the effect of the exploration rate at which an agent can operate using every algorithm. The standard deviation is also shown.

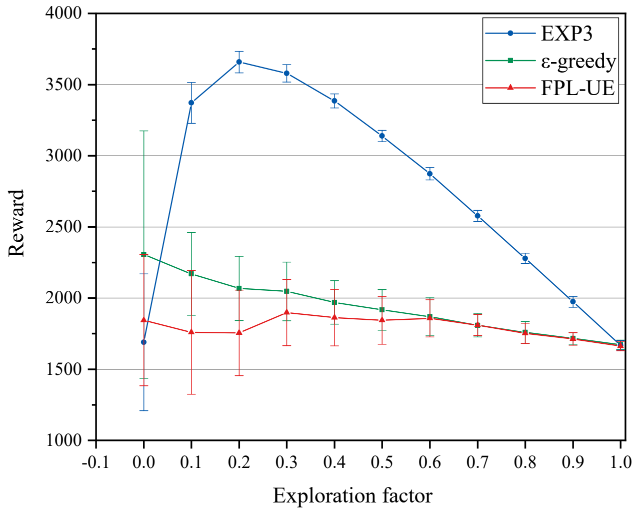

Figure 5.1: Reward for each exploration factor in EXP3, ε-Greedy and FPL-UE.

As can be seen from the figure, the average reward achieved by FPL-UE during 25 epochs slightly changes through the different exploration factors, except for the case where its rate is of 0.3. It can be observed that in the ε-Greedy case as the exploration factor increases, the performance of the algorithm gets more scarce. Whereas the average cumulative reward with EXP3 is significantly larger than the reported with the other two policies. Particularly, when the exploration rate is 0.2. In conclusion, the exploration rates are defined as $\varepsilon = 0.1$ in ε-Greedy, $\lambda = 0.3$ in FPL-UE and $\gamma = 0.2$ in EXP3 for the rest of simulations, since these factors provide in average the highest cumulative reward among 100 runs.

5.4.4.2 Performance Comparison

For the results presented here, 10 simulations are run and each simulation consists of a network where 5 agents need to select a partner among a set of opponents reached by opportunistic and direct interconnection (each opponent belongs to a different network). Every simulation run in all scenarios involves 5000 OEN encounters. All simulation results correspond to the arithmetic mean of these 10 simulation runs for every main agent (5 agents, 10 runs: 50 simulations per trial), with differences in the agents' energy profiles (as described in 5.4.3.2 and 5.4.3.1), number of opponents (between 2 and 3), and environmental conditions (Condition 1 - Condition 3). Each environmental condition varies over time according to the static (1 epoch, 5000 trials), moderately dynamic (5 epochs, 1000 trials each), dynamic (10 epochs, 500 trials each) and extremely dynamic case (25 epochs, 200 trials each). The experiments use as a baseline the best-fixed strategy and random selection of opponent to measure the performance of the MAB algorithms.

Cooperative Scenario

Figure 5.2 illustrates how the agents perform when they use every policy described in this work in a static, moderately dynamic, dynamic, and extremely dynamic environment when the topology is static and all agents behave in a cooperative way. Figure 5.2(a) shows ε-Greedy as the best algorithm to select a negotiation partner. During 5000 opportunistic encounters in constant competitiveness between the opponents, the most appropriate policy is a simple greedy approach. It enforces only 10% of randomness in its strategy to explore sub-optimal options, but also to consider environmental changes. Although the exploration factor is higher in FPL-UE with respect to EXP3, the second best choice in this configuration is the first algorithm. An explanation for these results is that in FPL-UE the best partner is the one that generates the maximum estimated reward over time, which in this case, is fixed for the entire epoch. While EXP3 uses the exploration factor to maintain a list of weights for each of the opponents, to further support its mixed strategy on deciding which action to take next. However, EXP3 may benefit of its methodology if the environment changes over time, since it will use these weights to adapt to such conditions. In any case, the algorithms learn to play actions that enhance the overall performance of the agents and need to be applied in the specific scenario to know which strategy is the best against the corresponding problem. For instance, in the context of adversarial online learning in defender-attacker encounters, FPL-UE has proved to achieve efficient results against the best fixed strategy on hindsight [160]. The hypothesis, in fact, was that negotiating agents using the FPL-UE policy can tackle better the adaptive behaviour of the opponents. Although FPL-UE has certainly improved the performance of the agents, in general, EXP3 technique has shown to efficiently deal with the partner selection problem in highly dynamic settings. The non-learning approach or random selection of the agents consistently shows a poor performance. The random selection is the worst of the benchmarks but it is implicitly suggested in the existing literature. In this case, the random selection strategy in the experiments achieves up to 63% of the energy that can be allocated when the selection of a partner is performed intelligently. From the scenario covered in Figure 5.2(a) ε-Greedy is efficient with an average efficiency of 93%, followed by FPL-UE with 87% and EXP3 with 84%, respectively. The performance of the action-selection strategies compared to that of the best fixed strategy is degraded when changes appear in the environment. Their efficiency is affected even more when environmental transitions take place more frequently (see Figure 5.2(b-d)). Temporal changes in the reward distribution structure are an intrinsic characteristic of this domain. These changes at every decision epoch vary the expectation of the rewards and motivate the agents to dismiss information gathered about the opponents, which in turn encourages exploration. However, the less time the agents have to adapt to these variations, the less they are able to characterise the reward distributions. In this regard, EXP3 implements the best approach. In particular, EXP3 is almost consistent with 84%, 84%, 83% and 82% of efficiency among respective environments: static, moderately dynamic, dynamic and extremely dynamic. The sensitivity to disturbance is more notorious in ε-Greedy, which efficiency decreases up to 16% when changes occur. Similarly, FPL-UE on average loses its ability to opportunistically select a negotiation partner up to 15%. Thus, EXP3 is better in all three conditions. In conclusion, in this first scenario, the ε-Greedy policy is enough when temporal changes can be avoided over time in a partner selection problem. EXP3, however, outperforms the other two policies for a broad range of temporal uncertainties in the environment.

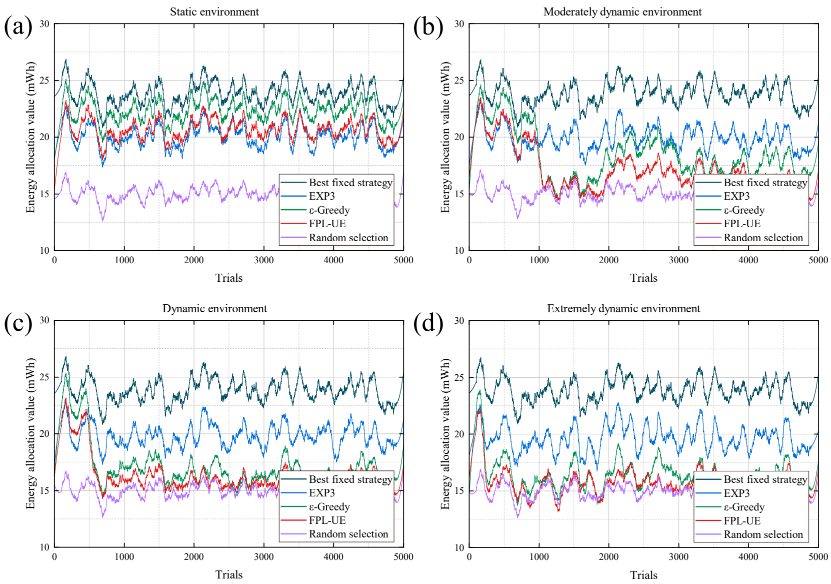

Figure 5.2: Cooperative scenario. Energy allocation in a 5-agent network with static network topology and Conceder agents ($\beta = 0.05$), in static, moderately dynamic, dynamic, and extremely dynamic environments.

Multiple Behaviours Scenario

The second scenario where agents are simulated with multiple negotiation behaviours is evaluated in Figure 5.3. The negotiation behaviours, in this case, determine the target energy allocation value an agent desires in each round of the negotiation encounter. Similarly to Figure 5.2, the results are shown under four degrees of environmental dynamism with respect to the energy availability that directly affects the negotiation behaviour of the agents: in static, moderately dynamic, dynamic, and extremely dynamic environments. As can be observed in the figure, the energy allocation on average has changed in comparison to the energy allocation achieved when all the agents are Conceder. Following the results from the random selection strategy, there is a slight reduction. Specifically, the results obtained when all agents behave "cooperatively" report

11% more than the amount of energy allocated in this scenario (52%). In any case, even if the agents offer concessions rapidly at the beginning of the encounters, the selection of the most appropriate partner by intelligently choosing the opponent, makes a difference in this model.

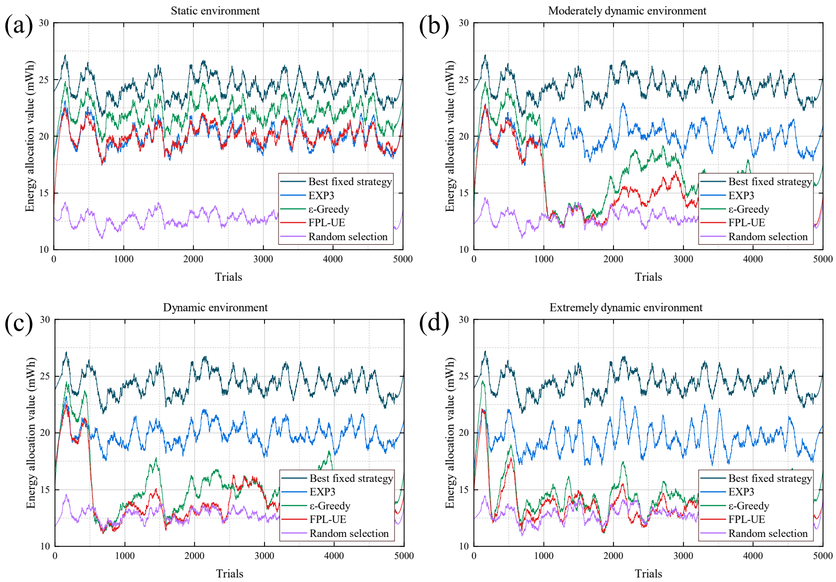

Figure 5.3: Multiple behaviours scenario. Energy allocation in a 5-agent network with static network topology and agents with β between 0.05, 1.4 and 1.9 in static, moderately dynamic, dynamic, and extremely dynamic environments.

Now, from Figure 5.3(a) is noted that the performance of every approach is decreased, compared to that of the case of cooperative networks. This is, however, due to the fact that the average energy allocation amount that the agent's best strategy achieves in this scenario has increased. The main reason behind this difference is that there are fewer concessions among the agents and the desired utility levels are higher. In the cases where there are agents with a Conceder behaviour against an opponent with a Boulware behaviour, if the first one has enough energy to power itself and share or requires a minimum amount of cooperation (less than the opponent), the agent playing a Boulware tactic gets a better agreement. This meets the following statement, when Boulwares make deals, they receive a higher individual utility [90]. The second reason for this variation in the policies' performance is the dynamism introduced by the multiple negotiation's behaviours. In this scenario, the set of opponents offer different amounts of energy values and the diversity of potential agreements is increased between the agents. The environment is then more dynamic from an agent's perspective since the agent's behaviours change according to the amount of energy they harvest. Consequently, the variability of the opponent's negotiation tactics directly impacts the learning curve of the agents. The adaptation to these variations is however best approached by EXP3, as shown in the figures. Specifically, in the static environment, EXP3 technique achieves up to 82% on average, of the total energy that can be allocated with the best fixed strategy. That in comparison with the first scenario is only 2% less of its original capacity. Such level remains stable, as seen in Figure 5.3(b-d), where the efficiency of EXP3 is of 82%, 82%, and 80% for the moderately dynamic, dynamic and extremely dynamic environment, respectively. The other two policies reduce their performance as more dynamicity is considered in the agent's behaviours. ε-Greedy reduces its performance, on average, up to 9% in every environment, while FPL-UE policy reports a decrement up to 10% of the amount obtained in the cooperative scenario. This indicates that EXP3 is less sensitive to the negotiation strategy changes than the rest of the policies. Most important, the EXP3 estimation method is not affected by the introduction of negotiation in the system. Overall, the learning approaches achieve better results compared to the random selection of the negotiation partner over time.

Dynamic Topologies Scenario

The results for the last scenario are depicted in Figure 5.4. The changes in the networks' topology are taken into account while the environmental changes on weather conditions are also studied. In this regard, the energy availability determined by the ambient energy sources affects the negotiation strategy of the agents. Thus, the agents have to deal with the challenges of environmental changes and the varying operational status of the opponents. The figure shows how the performance of the policies is again decreased by the introduction of the agent's movements (because of failure, rejection to be part of OEN, activity commute between active/sleep status). For instance, the energy amount allocated by the agent using EXP3 is reduced up to 9% in comparison to the amount allocated in the first scenario and decreased up to 7% of the average energy amount allocated when there are no topology changes but multiple negotiation behaviours. The same occurs with the rest of the policies on a different level. For both FPL-UE policy and ε-Greedy, the efficiency is reduced up to 12% from the results obtained in the cooperative scenario to these of the last scenario, and up to 3% from the values reported in the multiple behaviours scenario without topology changes to these values of multiple behaviours with topology changes. Although these policies are affected in less proportion than EXP3, this algorithm presents the best results as the environmental changes become more frequent (see Figure 5.4(b-d), respectively). In fact, an agent deciding a partner using the EXP3 algorithm achieves 74% efficiency while the use of ε-Greedy supply 60% efficiency and FPL-UE learning approach obtains 55% in the most challenging case i.e. in extremely dynamic environments with environmental and topology changes.

Despite the fact that the agents' performance using EXP3 is affected by the topology changes, this policy achieves the best results with respect to the best fixed strategy and random selection. EXP3 approach is consistent through the performance evaluation in each scenario and its reward estimation method proved to handle more realistic domains of complex and dynamic environments. This is supported by the results depicted in the variety of scenarios studied here. Moreover, EXP3 is not sensitive to the negotiation strategies incorporated in the decision process. Thus, the adaptive learning feature provided by EXP3 is the most suitable solution for the problem of partner selection. Furthermore, the EXP3 policy can be applied in a broader range of negotiation agents interactions where computationally-lightweight solutions are required. The results of this research are quite useful for designing agents in open environments that need to

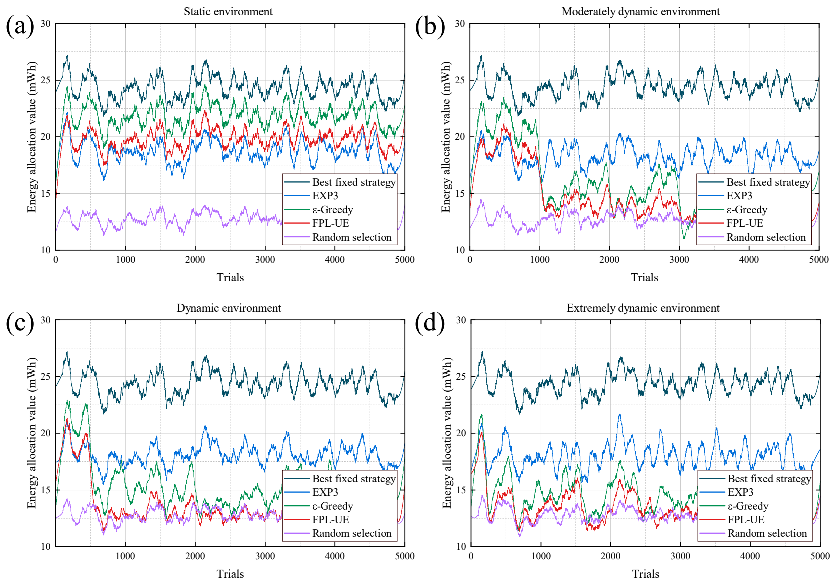

Figure 5.4: Dynamic topology scenario. Energy allocation in a 5-agent network with dynamic network topology and β between 0.05, 1.4 and 1.9 in static, moderately dynamic, dynamic, and extremely dynamic environments.

cope with the uncertainty of the adversarial setting and network conditions. In this case, the MAB learning model presented in this work allows an agent to select the most prospective partner from a set of opponents and reach efficient energy allocation agreements in the long term.

5.5 Summary and Discussion

The topic of negotiation has been widely studied to solve cooperation problems in complex systems. In the domain of WSNs, the limited resources along with the dynamism, distribution and heterogeneity of the networks present key considerations for the design of an effective automated negotiation technique. Subject to the requirements, the problem lies in the fact that multiple nodes report the same desire to cooperate. In this regard, the potentially resource-consuming negotiation with a large number of agents, especially in open dynamically changing environments, can be impractical. Therefore, the OEN methodology includes the partner selection step to propose a new model tested under different scenarios. The challenge is to enable an agent to adaptively adjust its partner selection depending on the characteristics of such scenarios with the possibility of maximising its energy allocation in the long term.

Three scenarios are designed for the experiments, which vary three characteristics: the negotiation behaviour of the agents, the energy availability of the ambient energy sources and the network topology. Accordingly, the following situations are simulated. A cooperative scenario, where the environment varies the energy availability and the agents adopt a conceder behaviour during all the negotiation interactions. A multiple behaviours scenario, which covers both environmental conditions variations and different concession strategies. Finally, a dynamic topology scenario, with the conditions of the multiple behaviours scenario and the movement of agents.

Consequently, for the partner selection problem, a MAB based approach is proposed to address the decision of choosing a partner under uncertainty of the opponents and the dynamism of the environment. The setup of negotiation fits naturally with an adversarial MAB problem, where the reward obtained by an agent in every opportunistic interaction not only depends on the actions taken by the agent but also on the adversaries behaviour. Against this background, state-of-the-art MAB policies are applied in the negotiation context of non-oblivious adversaries: ε-greedy, EXP3, and FPL-UE. The algorithms are adapted as the strategies for efficiently select a partner and repeatedly re-learn the current best partner for cooperation.

The efficiency of the MAB policies is then compared against the random selection of a partner and the best strategy in hindsight. The results show that even in a cooperative scenario, where agents offer concessions rapidly at the beginning of the encounters, the agents improve their benefit by choosing a partner strategically instead of select it randomly. Every simulated setup establishes stationary events in the duration of epochs. The length of such epochs defines the four cases considered: long epochs or static environment (5000 iterations), moderately dynamic epochs (1000), dynamic epochs (500) and extremely dynamic epochs (200). In every scenario, the ε-greedy algorithm reaches the best performance for the static environment. Therefore, the application of a simple heuristic as selecting a 10% of exploration and exploit the best opponent the rest of the time is enough to handle the partner selection problem in a stable scenario.

The problem becomes even more challenging with setups where agents employ tougher negotiation strategies and the presence of them is unstable. In any case, the bandit strategies achieve improved energy allocation agreements by adjusting to dynamic environments. In this direction, the EXP3 policy produces better results at a large number of unexpected events as the environment becomes more dynamic. Moreover, the EXP3 policy consistently shows the best performance despite wide variations. Thus, EXP3 policy increases the adaptability of the negotiation-based cooperation.

Furthermore, the partner selection step may support the decision to start a negotiation based on the estimations of the profitability of the opponents. The partner selection policy can also reduce the complexity of addressing reasoning to negotiate. For instance, instead of selecting a partner based on the reward associated with it, the negotiation may involve strategies to model the negotiation behaviour of its opponent. In particular, some strategies include regression techniques to estimate the concessions of the opponents and predict possible agreements. Since this thesis is interested in resource-constrained WSNs, it concentrates on low complexity solutions that don't require learning mechanisms that predict the opponent's future offers. The prediction techniques require a sufficient number of the opponent's offers to apply the learning approach and start the estimation of the counterpart's information (such as its deadline or reservation values) in order to obtain better deals. In fact, the complexity of the utility space increases with the interdependent issues and the number of time slots involved in the energy cooperation domain. Thus, the MAB partner selection strategy can serve as an alternative to maximise an agent's utility in the long-term.

Chapter 6

Conclusions and Future Work

In this chapter, the conclusions and future work are presented. The first section consolidates the contributions of this research. The following section highlights some research opportunities towards the realisation of a negotiation-based cooperation between WSNs.

6.1 Conclusions

Based on the literature review on Chapter 2, this work investigates a feasible approach to enable cooperation between co-located WSNs that meet opportunistically. Such a problem had not been addressed before. Although networks can have multiple cooperation incentives, this work focuses on the optimisation of energy use. Negotiation-based networking to enable cooperation across heterogeneous co-located networks has been previously studied using a solution that incorporates a central and powerful monitoring application. Game-theoretic models have also been proposed to analyse the cooperation problem. However, due to the characteristics of WSNs, the proposed methods are not practically feasible. The availability of complete information, the presence of a mediator with high computational power and the conception of full rationality are not suitable in this domain.

To address this challenge, a new methodology to enable negotiation-based cooperation between co-located WSNs was proposed in Chapter 3. The steps of the methodology were developed in every chapter to accomplish the specific goal of this work: opportunistic energy negotiation, called OEN. As mentioned above, networks can have different or multiple optimisation goals. In that case, the proposed methodology can be custom-tailored towards a specific objective.

With the aim to optimise a network's power management using the suggested approach, the first step for an agent is to identify its own efficiency. In the domain of OEN, it corresponds to the energy allocation scheme that a node can employ to power its load. Thus, the first contribution of this thesis is the optimal energy allocation algorithm described in Section 3.3. This power management technique is tested during every simulation presented in this work. Such algorithm enables self-organised agents that can anticipate insufficient energy allocation and the opportunity to start an OEN.

The establishment of OEN is described in Chapter 4. Using a discrete event simulator as OMNeT++, the discovery protocol to reach the negotiation agents in a 1-hop neighbourhood is implemented using a publish-subscribe protocol. The results shown on maximum energy cost and latency introduced by the protocol, assuming up to 7 opponents subscribed to the OEN process, have a negligible impact in the agent's performance (<0.01 J and <0.1 s). Thus, the obtained results demonstrate that a node can engage in OEN with a minimum cost even in the emergence of seven co-located and distinct WSNs.

In Chapter 4, a heuristic approach which employs existing negotiation methods is proposed to model the cooperation problem between networks. The negotiation framework is tested through extensive simulations using real energy profiles: load with regular duty cycles, and energy harvesting from weather data. The heuristic approach is used to reallocate the renewable energy of the agents' sources and enhance the power management of the networks extending their boundaries. The heuristic model is validated using time-dependent concession strategies. Results found that independently of the agent's energy availability, a conciliatory behaviour is the best strategy in this domain.

The negotiation strategies of OEN were compared against NBS and the shortest-distance proposing mechanism. The results have shown how an agent using OEN can reach on average 77% of the utility reached by NBS. OEN mechanisms lead to the maximisation of an agent's energy allocation in a 14%. The percentage of deals reached was of 70% for the OEN's approach, against 50% and 100% for shortest-distance and NBS respectively. OEN also proved to reach 59% of the energy neutrality agreements reached by the optimal solution. Although this last result is promising, the agreements reached with OEN were later improved with the inclusion of a partner selection policy.

Chapter 5 presented a partner selection model based on bandits. State-of-the-art policies on adversarial bandits were compared to find the most appropriate for OEN's domain and its dynamism. The multi-armed bandit approach enables an agent to reach efficient energy allocation in the long term. In every of the studied scenarios, the bandit strategies achieved improved energy allocation agreements compared with the random selection mechanism. Using these estimations, an agent is also able to decide if envisaged negotiation is in fact helpful. The results improved up to 39% against a random selection strategy in static conditions and up to 30%, 29% and 27% in moderately dynamic, dynamic and extremely dynamic environments respectively. Such improvement represents the maximisation of an agent's energy allocation up to 53% in static environments and up to 41% in an extremely dynamic condition.

The EXP3 policy achieved better results than FPL-UE and ε-Greedy at a large number of unexpected events as the environment tested became more dynamic. EXP3 can reach an efficiency of 74% against the best selection strategy in the most challenging scenario studied in this thesis. Thus, the introduction of reinforcement learning techniques feasible in the domain of WSNs, can bring more efficient energy agreements than the random selection strategy implicit in previous studies.

Some negotiation techniques were reviewed and the heuristic approach was selected as the most appropriate for this domain. The list of specifications for the domain of WSNs (Section 1.3) was covered in all aspects, but further investigations can be derived to improve this first attempt of negotiation-based cooperation model for the WSN domain.

The vision of WSNs cooperation demands mechanisms to enable an agent with self-configuration, self-optimisation and self-management capabilities. The inclusion of negotiation in a highly heterogeneous environment helps to resolve conflicts and permit optimisation of network performance by taking into account a wider geographical coverage and resource capacity. Therefore, the design and evaluation of a negotiation process between WSNs become an important research topic towards autonomic environments in IoT.

6.2 Future Work

The research developed in this thesis has successfully addressed the cooperation problem between co-located and independent WSNs. However, there are still interesting research opportunities to realise the full potential of cooperative networks:

- **Management of security.** For the safe flow of offers and authenticity of the agreements, new approaches are needed for addressing security in bilateral negotiation between sensor nodes. Additionally, these privacy controls can protect the data exchanged autonomously by the agents when cooperation proceeds and create a trust relationship between networks. Security, in general, is a major concern in WSNs. The inclusion of security schemes has a significant impact on energy consumption and memory usage [167]. Although there are security controls for the E-commerce context [168], its nature is totally different, which make these approaches incompatible with this domain. Thus, ensuring holistic security in automated negotiation between agents in OEN represents a key research challenge for its realisation.

- **New cooperative routing protocols.** Once the cooperation is agreed between networks, the cooperation may proceed. As described in Chapter 2, energy sharing can take place through the acceptance of energy-hungry services as data processing or packet forwarding. In case agents implement cooperative packet forwarding, cross-network routing needs to be possible. In the same way as cross-boundary data transmission at MAC level is already feasible with ODI. Having different networks, there might be different network-layer protocols and multiple addressing schemes. Thus, there should be support for address translation, address mapping between the networks involved, or common identification of nodes.

Additionally, new negotiation approaches can be developed. A scenario where multiple nodes share the same area may consist of multi-lateral negotiations modelled as auctions. One-to-many models in which multiple agents negotiate with a single agent can be proposed for the multi-issue problem. In a single auction, agents place bids (buyers) on the energy resource offered by an agent experiencing a surplus of energy and acting as a seller. As noted, in this type of negotiation the roles are different and there is no way to perform communication of offers and counteroffers. Moreover, auctions may require a mediator. However, new negotiation approaches can be implemented based on single or double auction to study their impact and feasibility.

Appendix A

Selected Publications

This appendix includes the following paper published as a result of this research:

- Ortega, Andre P., Geoff V. Merrett, and Sarvapali D. Ramchurn. "Automated Negotiation for Opportunistic Energy Trading Between Neighbouring Wireless Sensor Networks." *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. IEEE, 2018.

Automated Negotiation for Opportunistic Energy Trading Between Neighbouring Wireless Sensor Networks

Andre P. Ortega[1,2], Geoff V. Merrett[1], Sarvapali D. Ramchurn[1]
[1]Electronics and Computer Science, University of Southampton, UK
[2]Escuela Superior Politécnica del Litoral, Ecuador
apoa1g15@soton.ac.uk, gvm@ecs.soton.ac.uk, sdr1@soton.ac.uk

***Abstract*—As the Internet of Things grows, the number of wireless sensor networks deployed in close proximity will continue to increase. By nature, these networks are limited by the battery supply that determines their lifetime and system utility. To counter such a shortcoming, energy harvesting technologies have become increasingly investigated to provide a perpetual energy source; however, new problems arise as a result of their wide spatio-temporal variation. In this paper, we propose opportunistic energy trading, which enables otherwise independent networks to be sustained by sharing resources. Our goal is to provide a novel cooperation model based on negotiation to solve coordination conflicts between energy harvesting wireless sensor networks. Results show that networks are able to satisfy their loads when they agree to cooperate.**

I. Introduction

Internet of Things (IoT) deployments in industries, cities, healthcare and home automation are spread all over the world. A core technology required for IoT are wireless sensor networks (WSNs), which gather information from the environment, analyse it, make decisions and act accordingly. In many of these applications, sensor nodes are battery-powered and limited in energy supply. Thus, one proposed solution is to extend performance optimisation to the inter-network approach by enabling cooperation among networks that co-exist in a physical location [1]–[4].

Energy harvesting technologies have also gained widespread attention to enhance node lifetime. Moreover, ways to capture green energy from regenerative sources for self-sustainable operation is a key driver in today's low-power devices for smart applications. However, energy harvesting wireless sensor networks (EHWSNs) are conditioned to spatio-temporal variations of energy availability. The main objective of EHWSNs, because of their unlimited power supply, is the optimisation of their energy use to operate continuously. This mode of operation is called energy-neutral operation: a harvesting node achieves it if the energy supply during a harvesting period is sufficient to replace the amount consumed during the same time [5].

Adaptive algorithms have been developed to address the spatio-temporal variation of ambient energy sources and scale a node's performance appropriately, in order to deliver energy-neutrality. These algorithms typically adjust parameters such as the duty-cycle or sampling rate [6], [7]. Other energy-neutral algorithms exploit the spatial variation and distribute load according to energy reserves [8]. However, these algorithms are limited by the bounds of one network domain; i.e. if one node is expecting insufficient energy and the rest has a scarce energy input, no solution exists.

The cooperation problem among distinct WSNs has been studied in a game theoretic setting [1], [2]. These works model the behaviour of a network as a game to analyse the existence of strategies, looking for equilibrium among rational players that negotiate with each other to maximise their own benefit. They focus on the conditions under which cooperation is the best strategy in multi-domain WSNs, and make an exhaustive search on the available space to find a solution for each network's authority (i.e. those that form a Nash equilibrium with the highest possible lifetimes). For a WSN, this would necessitate nodes making a significant effort to calculate and store not only all their possible actions at each decision point, but also the ones corresponding to the other nodes. This is not feasible for devices with limited memory and power. One approach to deal with this complexity is to simplify

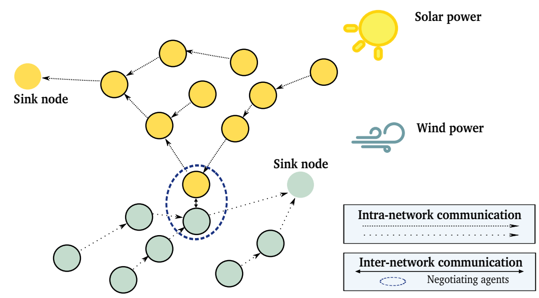

Fig. 1: Direct interconnection between co-located EHWSNs.

the settings in which nodes interact with each other and use heuristic methods. Before cooperation can be established, networks should be able to interact and find a mutually-acceptable agreement in favour of maximising their utilities. In the domain of EHWSNs, they must find an energy flow that deals with the spatio-temporal profile of their energy sources and satisfies as much as possible their energy consumption profile from collaborative effort. A multi-agent approach is a natural fit to this setting as individual sensor nodes need to autonomously negotiate and form an agreement as to how to share their resources [9], [10].

Against this background, we motivate the use of a solution based on automated negotiation and propose a novel approach to model cooperation between nodes with a direct interconnection architecture, i.e. without an intermediary (Figure 1). The contributions of the work reported in this paper are:

- An alternating offers protocol for the nodes to exchange offers to trade energy-hungry services.
- An optimisation algorithm based on Linear Programming (LP) to optimise the allocation of energy to maximise individual actors' preferences.
- An analysis of time-dependent negotiation strategies in EHWSNs for energy re-allocation of distinct energy harvesting sources.
- Results showing how negotiation can be delimited by a short-term deadline and end in social-welfare maximising deals.

II. System Model and Problem Formulation

We assume an initial simplified setting where two distinct EHWSNs with a different type of energy harvesting source (e.g. solar and wind) share the same location and direct interconnection is possible between each pair of nodes with overlapping radio range [11]. Our main motivation in investigating negotiation applied in this domain is to observe the effects of cross-boundary energy transfer for sensor's power management and this setting is suitable for that purpose.

While it is convenient to envisage opportunistic energy trading as physically transferring energy across a network boundary, energy is actually logically transferred by accepting energy-consuming tasks as data processing or packet forwarding [1]–[3]. For example, each network involved may control the agreed energy flow by asking for/providing routing favours.

A. Model Assumptions

Each network N_k, $k \in \{1, 2\}$ is formed by a set of unique sensor nodes and a sink. Each node is controlled by an agent, which is denoted as $\alpha_{k,i}$, $i \in \mathbb{N}$. The agent has complete knowledge of all the relevant node's information, such as its energy profile variables, battery capacity and residual energy.

We assume that the period at which energy is harvested by a node is T (e.g. 24 hours for solar energy) and it is divided into discrete time slots $T = (1, ..., n)$ of equal duration L.

B. Energy Consumption Model

Each node controlled by $\alpha_{k,i}$ consists of an energy harvester unit, a rechargeable battery and several loads: a radio, CPU and sensors. We use the energy model introduced in [12] and define $\boldsymbol{E^c_{k,i}} = (E^c_{k,i}(1), ..., E^c_{k,i}(n)) : \boldsymbol{E^c_{k,i}} \in \mathbb{R}^+$ as the energy consumed by $\alpha_{k,i}$ over n time slots. At any given slot t, we can calculate the total energy $E^c_{k,i}$ that an agent $\alpha_{k,i}$ consumes as:

$$E^c_{k,i}(t) = V \cdot \left[D \cdot I^{active} + (1 - D) \cdot I^{sleep} \right] \cdot L \quad (1)$$

Then, the energy is dependent on the duty cycle D, supplied voltage V, active mode current I^{active} and sleep mode current I^{sleep}. D is chosen by the application, while I^{active}, I^{sleep} and V can be known in advance using datasheet information.

C. Energy Management Model

Our model is built on the models proposed by [5] and [13]. We assume that all nodes can harvest energy and store it in their battery for future use. Without loss of generality, we assume that the replenishment of energy occurs at the beginning of each time slot t.

The expected energy input during each slot t can be forecast from historical information with a high level of accuracy. Energy can then be allocated to each slot t. We use $E^c_{k,i}(t)$ and $E^{hrv}_{k,i}(t)$ to denote the energy profile variables for each time slot. The amount of energy that can be generated by the harvesting unit in n time slots is defined as $\boldsymbol{E^{hrv}_{k,i}} = (E^{hrv}_{k,i}(1), ..., E^{hrv}_{k,i}(n)) : \boldsymbol{E^{hrv}_{k,i}} \in \mathbb{R}^+$. For example, if the harvesting period starts at 00:00 and L is 1 hour, then $E^{hrv}_{k,i}(1)$ is the expectation for the energy harvested during slot 1 (from 00:00 to 01:00), $E^{hrv}_{k,i}(2)$ is the expectation of energy during slot 2 (01:00 to 02:00), etc.

$B_{k,i}(t)$ is used to represent the residual battery energy at the beginning of slot t in agent $\alpha_{k,i}$. Then $\boldsymbol{B_{k,i}} = (B_{k,i}(1), ..., B_{k,i}(n)) : \boldsymbol{B_{k,i}} \in \mathbb{R}^+$ denotes the battery level in n time slots. The battery is characterised by a limited capacity $B^{max}_{k,i}$ and charging efficiency η. The battery enables an agent to save and use energy throughout a day, which helps the agent to compute an energy allocation, $\boldsymbol{E^{alloc}_{k,i}} = (E^{alloc}_{k,i}(1), ..., E^{alloc}_{k,i}(n)) : \boldsymbol{E^{alloc}_{k,i}} \in \mathbb{R}^+$, to assign the harvested energy $E^{hrv}_{k,i}$ to the energy consumed $E^c_{k,i}$ by the load of the node.

When $E^{hrv}_{k,i}(t)$ is lower than $E^c_{k,i}(t)$, some of the energy used by the sensor node is discharged from the battery. We use $\boldsymbol{d} = (d(1), ..., d(n)) : \boldsymbol{d} \in \mathbb{R}^+$ to represent this amount. When $E^{hrv}_{k,i}(t)$ is higher than $E^c_{k,i}(t)$, all the energy used in the node is provided by the energy source and the battery is charged with the excess, as required. We use $\boldsymbol{c} = (c(1), ..., c(n)) : \boldsymbol{c} \in \mathbb{R}^+$ to denote this amount in n time slots. Any excess energy received at times when the battery is full is discarded by the node. The energy that the agent is unable to use or store is waste, denoted by $\boldsymbol{w_{k,i}} = (w_{k,i}(1), ..., w_{k,i}(n)) : \boldsymbol{w_{k,i}} \in \mathbb{R}^+$. Then we can calculate the energy used from the battery in any slot t as:

$$B_{k,i}(t) - B_{k,i}(t + 1) = d(t) - \eta \cdot c(t) \quad (2)$$

In our domain, an opportunistic energy trade is triggered when a node's energy level has dropped below a threshold. Then, the initial battery status $B_{k,i}(1)$ is equal to $\eta \cdot b$ where b is the energy level at $t = 1$. At each time t, $\alpha_{k,i}$ also considers the amount of energy to receive/give from the negotiation, which is defined by $\boldsymbol{o} = (o(1), ..., o(n)) : \boldsymbol{o} \in \mathbb{R}^+$. o represents the offer of energy at each time slot, i.e. The issues of this negotiation domain. We call these offers *energy flow offers*. A valid energy flow offer must include the energy values for the predetermined time of cooperation, e.g. If networks expect to cooperate for 24 hours, then the energy flow must include 24 values. The direction of the energy flow is denoted by a positive or negative sign. If positive, the amount is an offer of energy from the agent to its opponent, otherwise, it represents the energy to be received from the opponent. For example, if two agents are willing to cooperate with each other for a period of 2 hours and L is set to 30 minutes, then an offer of energy from agent $\alpha_{k,i}$ to the other party can be $o = [-1.88, -0.7, 18, -4]$; where -1.88 mWh, -0.7 mWh and -4 mWh represent the energy savings of $\alpha_{k,i}$ from the opponent's cooperation (e.g. by packet routing) at time slots 1, 2 and 4 respectively, while $\alpha_{k,i}$ compromises to provide 18 mWh through collaborative effort to its opponent at time slot 3.

D. Utility Function

The objective function of this model is described as the total energy consumption that is satisfied (i.e. energy allocation $E^{alloc}_{k,i}$) at period T. Then the utility of an agent represented by u is defined as follows:

$$Objective \quad \max u = \sum_{t=1}^{n} E^{alloc}_{k,i}(t) \tag{3}$$

Subjected to the following constraints:

$$E^{alloc}_{k,i}(t) = E^{hrv}_{k,i}(t) - c(t) + d(t) + o(t) - w(t) \qquad (c_1)$$

$$E^{alloc}_{k,i}(t) \leq E^{c}_{k,i}(t) \qquad (c_2)$$

$$B_{k,i}(t) - B_{k,i}(t+1) = d(t) - \eta \cdot c(t) \qquad (c_3)$$

$$B_{k,i}(1) = \eta \cdot b \qquad (c_4)$$

$$0 \leq c(t) \leq B^{max}_{k,i} \qquad (c_5)$$

$$E^{c}_{k,i}(t) - E^{hrv}_{k,i}(t) \leq d(t) \leq B^{max}_{k,i} \qquad (c_6)$$

$$0 \leq B_{k,i}(t) \leq B^{max}_{k,i} \qquad (c_7)$$

$$0 \leq w(t) \leq E^{hrv}_{k,i}(t) \qquad (c_8)$$

Equations (c1) and (c2) represent the energy balancing condition. The allocated energy to a node defined by the harvested energy, battery flow, the energy offer and waste is equal or smaller than the node's load at time slot t. Equations (c3)-(c7) define the battery status and flows constraints regarding its capacity. Equation (c8) is used to guarantee that the energy waste is an excess of the energy harvested.

The solution to the optimisation problem yields the amount of energy that must be allocated to a sensor node in every t and the evolution of residual energy in its battery over period T. Following the model described, an agent can compute the optimal energy flow that benefits both agents, but this requires complete information and high computation capabilities since the set of all possible agreements is exponential in the number of time slots. Cooperative approaches must ideally result in Pareto-efficient outcomes, which means that one agent cannot be better off without making the other agent worse off. In section IV, we present a cooperative solution that satisfies this property of efficiency known as the Nash Bargaining Solution (NBS) [14] to find an agreed energy flow between agents, but first, we describe the heuristic model used for the bargaining process of this domain.

III. Heuristic Approach for Opportunistic Energy Trading

There are four fundamental parts in a negotiation model described by a heuristic approach: 1) the negotiation protocol or rules of interaction for the agents, 2) the definition of issues or objects in contention (see II-C), 3) the utility function or agents' preference model (see II-D), and 4) the tactics or offers' generator functions that are applied during the bargaining process, which along with the utility function comprise the decision making apparatus the participants employ to act according to the negotiation protocol and reach their desired goals [15], [16]. The protocol and tactic employed are defined below.

A. Multi-issue Bilateral Negotiation Protocol

We adopt Rubinstein's alternating-offers protocol [17] for the negotiation of energy among neighbouring EHWSNs. In a bilateral negotiation, both agents desire to cooperate but have conflicting interests regarding their preferences (in this domain due to distinct batteries, power consumption and energy harvesting profiles).

According to the protocol, all the agents involved have one turn per round to respond to the current state of the negotiation. One of the negotiating agents starts with an offer to its opponent. Whenever an offer is made, the opponent can accept or reject the offer. If the offer is accepted, then the bargaining ends and an agreement is reached. If the offer is rejected, the agent in turn proposes an agreement, which again the opponent may accept or reject in the next round. We continue the negotiation until a final negotiation round. When one negotiating agent reaches a final round without a favorable response or an agreement is found, the negotiation ends. In the first case, the negotiation fails and terminates with no deal possible.

In our domain, we must consider the number of messages exchanged between nodes and limit the negotiation to a short-term deadline. Thus, a predefined maximum negotiation round is set. Specifically, in our scenario, automated negotiation can complete in seconds, which makes time inappropriate to model the deadline. In each negotiation round, an offer contains multiple issues that are negotiated simultaneously. We assume that the knowledge of the negotiation domain (issues, deadline, initial negotiating agent) is known by both agents beforehand, and is not changed during the whole negotiation process. As defined in II-C, o represents the offer of energy. Thus, $o^r_{1,1\to2,1}$ is a vector of values proposed by agent $\alpha_{1,1}$ to agent $\alpha_{2,1}$ at round r, where $o^r_{1,1\to2,1}(t)$ is the value of energy proposed from $\alpha_{1,1}$ to $\alpha_{2,1}$ for time t. Each issue $o(t)$ has an acceptable range of values represented as the interval $[min_{k,i}o(t), max_{k,i}o(t)]$.

B. Negotiation tactic

In the negotiation context, heuristics are useful for the generation of initial offers, evaluation of proposals and decision of counter offers, based on computational approximations that produce good close to Pareto-efficient outcomes.

The main advantage of using heuristics in this domain is to model encounters between networks that are discovered opportunistically and have no information about the resources and preferences of each other.

Faratin *et al.* [15] studied strategic negotiation between autonomous computational agents and develop a formal model of reasoning to address the coordination problem. They defined a number of heuristic functions, which receive the name of *tactics* and use a single criterion (time, resources, behaviour, etc.) to generate new values for each issue in the negotiation set. The following family of tactics for counter-offer generation were applied in this domain.

1) Time-dependent tactics (TDT): The time elapsed in the negotiation is what conducts the values of the negotiation issues. It is the same for rounds, the more rounds has passed the more pressure is induced and faster concessions are possible. Then the value of $o(t)$ proposed by agent $\alpha_{1,1}$ to agent $\alpha_{2,1}$ at round r is giving by the following equation:

If $\alpha_{1,1}$'s utility decreases with issue $o(t)$:

$$o^r_{1,1\to 2,1}(t) = min_{1,1}o(t) + \gamma^r_{o(t)}(max_{1,1}o(t) - min_{1,1}o(t)) \tag{4}$$

If a's utility increases with issue $o(t)$:

$$o^r_{1,1\to 2,1}(t) = min_{1,1}o(t) + (1 - \gamma^r_{o(t)})(max_{1,1}o(t) - min_{1,1}o(t)) \tag{5}$$

We define $\gamma^r_{o(t)}$ as a polynomial function parameterised by $\beta \in \mathbb{R}^+_n$ as follows:

$$\gamma^r_{o(t)} = k_{1,1}o(t) + (1 - k_{1,1}o(t))(\min(r, rmax_{1,1})/rmax_{1,1})^{1/\beta} \tag{6}$$

The constant k at $t = 1$ represents the initial bargaining value of $o(t)$ while $rmax$ is the deadline. $\beta_{>0}$ defines the convexity degree of the curve. When $\beta > 1$, the agent is benevolent and characterised by a conceder behaviour (such tactic is called *Conceder*) and the offer rapidly changes to the reservation value. At $0 < \beta < 1$, the agent is tough and maintains its initial offer until it almost approaches the deadline (such tactic is known as *Boulware*). We limit our examinations to these negotiation tactics, while behavioural heuristics would be more appropriate in a dynamic environment as EHWSNs. But, these are less successful in short-term deadlines [15].

IV. Numerical Analysis

We assume agents observe that their residual energy level has dropped below the threshold set, they are appropriately synchronised and plan to cooperate for the next 24 hours, which start at 00:00 and end at 23:00 local time with L=1 hr, i.e. agents negotiate an energy flow of 24 values. Then T corresponds to the same period and time slots $T = (1, ..., 24)$. Numerical results are shown to demonstrate the performance of the agents with trading over the individualistic approach. All the results are obtained using MATLAB.

A. Simulation Setup

In this section, we study the problem of cooperation in a simplified scenario with a pair of nodes (each from a different network); negotiating agents $\alpha_{1,1}$ and $\alpha_{2,1}$ of N_1 and N_2, respectively.

Agent $\alpha_{1,1}$ is simulated as controlling a Memsic eKo mote, containing a 3.3 cm $\times$ 6.35 cm photovoltaic cell (assumed to be 10% efficient) to recharge a 600 mAh battery. We consider $\eta = 0.7$, which is typical of NiMH batteries. Agent $\alpha_{2,1}$ is simulated as controlling a Memsic MICAz node, with a micro-wind turbine to recharge a 600 mAh battery. The energy model in II-B is used to evaluate the energy consumption of both agents, using parameters obtained from empirical measurements and datasheets [18]. We consider a realistic scenario where an eKo node operates at 1% duty cycle, and the average power consumption is 0.615 mW. For the MICAz mote, an average load of 2.86 mW at 5% duty cycle of operation is expected. Thus, agent $\alpha_{1,1}$ and $\alpha_{2,1}$ demand 0.615 mWh and 2.86 mWh of energy in each time slot, respectively.

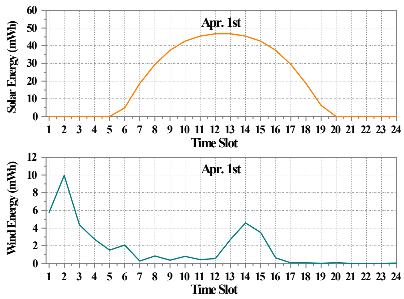

Fig. 2: Solar (agent $\alpha_{1,1}$) and wind energy (agent $\alpha_{2,1}$) harvested throughout a day.

The meteorological information used to compute the energy generation corresponds to the area of Southampton, UK (50.8997°N, -1.3955°W, Elevation 32 m) [19], [20].

The values of solar irradiance from April 2017 are used to estimate the hourly power output of a photovoltaic system for a day, which is proportional to the solar radiation, the panel dimension, and its efficiency. The estimated hourly energy output is shown in Figure 2. The energy exhibits a temporal variation that favours time slots 6-19 which correspond to times 05:00-18:00. The total energy generated is 452 mWh.

We adopt daily data from April 2017 to estimate the hourly average wind speed for a day. The power from the wind source can be calculated from its speed as in [21] considering a swept area of 5 cm $\times$ 5 cm for the wind turbine. From April data, we chose April 1st. The diversity between generation times in solar and wind creates an opportunity for energy trading. Then we scale this data to get the hourly power output of a highly efficient micro-turbine (Figure 2). The total energy generated in a day is 41.4 mWh.

B. Results

With the nodes' information and energy profiles described above, agents can compute their utilities (without trading) using the LP model described in Section II-D when the offer o is null. We compare the utility achieved by the agents without

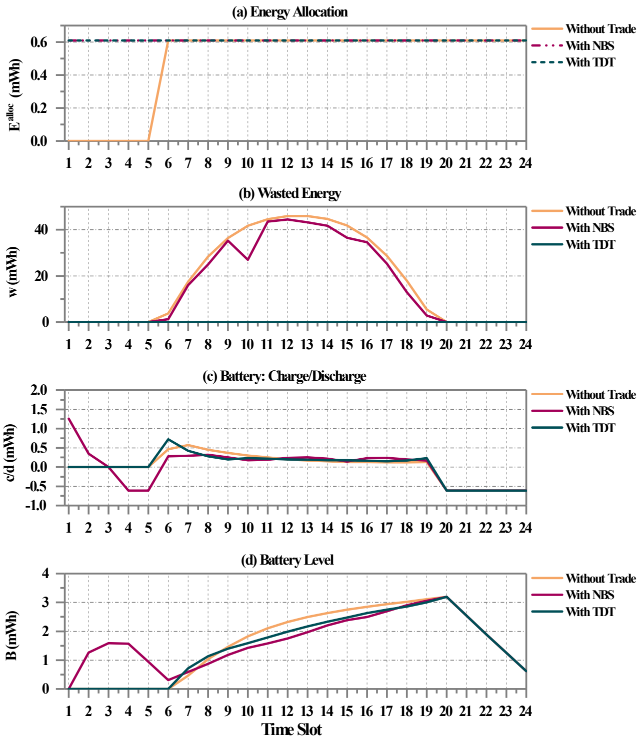

Fig. 3: Agent $\alpha_{1,1}$: Results of utility maximisation without trade, with NBS and with TDT.

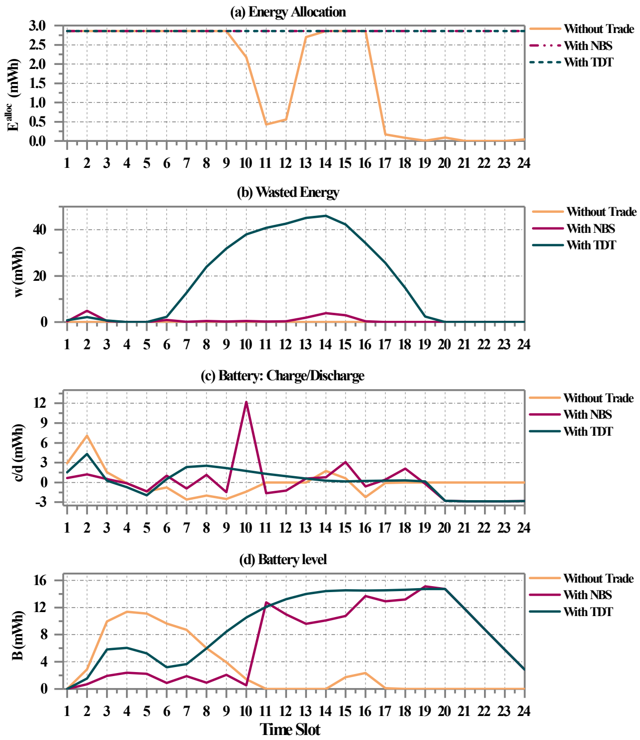

Fig. 4: Agent $\alpha_{2,1}$: Results of utility maximisation without trade, with NBS and with TDT.

trade, by NBS and by the bilateral negotiation protocol using TDT. Figures 3 and 4 for agents $\alpha_{1,1}$ and $\alpha_{2,1}$ respectively, show how agents can increase their utilities via cooperation and reduce the waste of energy excess. As presented in Figures 3.(a) and 4.(a), the energy allocation without trade is insufficient at time slots 1-5 for agent $\alpha_{1,1}$ and 10-13, 17-24 for $\alpha_{2,1}$, while it is equal to the load when there is energy trading. Thus, u for $\alpha_{1,1}$ increases from 0.79 to 1 and 0.59 to 1 for agent $\alpha_{2,1}$ when they reach an agreement to cooperate. The achievement of energy-neutrality in both scenarios depends in this case on the amount of unused energy from both agents and their matching requirements.

The results shown of TDT are obtained for a negotiation deadline set to 10 rounds. At the beginning of the negotiation, the agents make the offers that give the highest utility to themselves. No matter how low or high we vary the concession shape β (0.5 or 1.8) for any agent, the negotiation process with TDT ends with these results. If agent $\alpha_{1,1}$ starts, the process ends in the first round, otherwise it ends in the second round after $\alpha_{2,1}$ agrees with the counter offer of agent $\alpha_{1,1}$. Agent $\alpha_{1,1}$ has a large excess of energy to offer that satisfies agent $\alpha_{2,1}$ requirements (Figure 3.(b) Without Trade) and $\alpha_{2,1}$ is also able to assist $\alpha_{2,1}$ in its lack of energy during periods 1-5. In result, the utilisation of energy is maximised from 52.2 mWh to 83.3 mWh by negotiation while maintaining the application performance at the same rate at all times, i.e. the duty cycle is not affected. Then, the total energy saved via cooperation can be up to 7.08% for one day of the energy generated. The energy saved corresponds to the energy reallocated in cooperation which would otherwise go to waste without trade. The reduction of energy waste is illustrated in Figures 4.(b) Without Trade and Figure 4.(b) with TDT. Figures 3.(c), 3.(d), 4.(c) and 4.(d) show the state of the battery during the day for each agent, where the battery level matches the dynamics of the charging and discharging flows and none exceeds the maximum battery capacity. The difference in the battery dynamic between NBS and TDT depends on the negotiation's final outcome. For NBS, the solution corresponds to the offer $o_{1,1\rightarrow 2,1} = [-1.88, -0.96, -0.62, 0, 0, 2.75, 1.78, 3.57, 1.23, 14.77, 1.07, 1.37, 2.7, 2.88, 5.33, 1.94, 3.25, 4.91, 2.64, 0, 0, 0, 0, 0]$ while TDT finishes in $o_{1,1\rightarrow 2,1} = [-0.61, -0.61, -0.61, -0.61, -0.61, 3.54, 17.61, 28.49, 36.59, 41.75, 44.59, 45.91, 45.92, 44.63, 41.8, 36.63, 28.63, 17.86, 5.39, 0, 0, 0, 0, 0]$, which represent the 24 energy values (in mWh) agreed for a day of cooperation.

To evaluate the proposed cooperation model and compare the different agent behaviours, we make a slight change and match the load of agent $\alpha_{1,1}$ to agent $\alpha_{2,1}$. The results are shown in Table I as the agent that starts the negotiation, who finishes it, behaviours, utilities and final round. The following cases are considered:

- Case 1: Both agents employ a Conceder tactic.
- Case 2: Both agents employ a Boulware tactic.

- Case 3: $\alpha_{1,1}$ is tough while $\alpha_{2,1}$ concedes.
- Case 4: $\alpha_{1,1}$ concedes while $\alpha_{2,1}$ is tough.

TABLE I: Comparison between different negotiation cases

First turn	Final turn	Case	u $\alpha_{1,1}$	u $\alpha_{2,1}$	Final round
$\alpha_{1,1}$	$\alpha_{2,1}$	1	0.93	1	2
$\alpha_{2,1}$	$\alpha_{2,1}$	1	1	0.94	2
$\alpha_{1,1}$	$\alpha_{2,1}$	2	0.97	0.96	7
$\alpha_{2,1}$	$\alpha_{2,1}$	2	0.97	0.96	8
$\alpha_{1,1}$	$\alpha_{2,1}$	3	1	0.93	2
$\alpha_{2,1}$	$\alpha_{2,1}$	3	1	0.94	2
$\alpha_{1,1}$	$\alpha_{1,1}$	4	0.91	1	3
$\alpha_{2,1}$	$\alpha_{1,1}$	4	0.91	1	3

In those situations, energy-neutrality is only accomplished by one agent at a time and only when the opponent is benevolent. We can see that it is not possible to satisfy any energy consumption profile if both agents adopt a tough negotiation strategy. Similar to our simulation result before, the required number of rounds is low. The highest energy utilisation is given whenever agent $\alpha_{2,1}$ concedes faster at the beginning of the negotiation and while it has the first turn. In both cases, agent $\alpha_{1,1}$ reaches energy-neutrality. The second result of the table matches the utility levels reached by the optimal solution if a central and trustable authority is available to collect the information about the agents and calculate NBS. Since most of the parameters are the same for both agents (except the energy harvested), the available energy is a decisive factor in the establishment of cooperation. When sensors are energy-aware, spontaneous cooperation cannot take place and thus, a negotiation is required.

The presented results provide some insight on cooperation initiated by a negotiation, but more simulations have to be conducted to evaluate the model. For example, in our scenarios, additional costs for energy re-allocation, e.g. due to offers exchange, are not yet considered. Such issues, as well as further investigation in the effect of the network's dynamism on the negotiation model, is required.

V. Conclusions and Future Work

In this paper, we have proposed a model of EHWSNs and also applied a negotiation heuristic based on a TDT to optimally allocate the harvested energy by the nodes involved at each time slot. The main advantage of this cooperation initiated by a negotiation is that it allows to establish an opportunistic interaction between networks that cannot be conceived at design time about the resources of their neighbours, leading to a more integrated system of EHWSNs, while at the same time the use of the energy harvested is maximised.

The vision of WSNs cooperation brings many implications (from protocol diversity to security concerns), where several steps must be taken in order to ensure an effective interaction. One of the main challenges in extending power management to an area wider than the boundaries of one domain is the heterogeneity in terms of resources, which is the problem addressed in this work. An essential factor to establish cooperation is to know the costs and benefits that will incur to the parties. Here, a negotiation approach has been evaluated as a mechanism for networks to communicate and compromise to reach mutually beneficial results. In the future, we expect to extend the model to consider multiple nodes and the uncertainty generated by the energy availability and unexpected weather conditions.

Acknowledgment

This research is supported by Secretaría de Educación Superior, Ciencia, Tecnología e Innovación (SENESCYT) and Escuela Superior Politécnica del Litoral (ESPOL) under Ph.D. studies 2016, and the UK EPSRC under grant EP/P010164/1. Data supporting the results presented in this paper is available at DOI: https:/doi.org/10.5258/SOTON/D0621.

References

[1] Levente Buttyán et al. Spontaneous cooperation in multi-domain sensor networks. In *European Workshop on Security in Ad-hoc and Sensor Networks*, pages 42–53, 2005.
[2] Márk Félegyházi et al. Cooperative packet forwarding in multi-domain sensor networks. In *Pervasive Computing and Communications Workshops*, pages 345–349, 2005.
[3] Teng Jiang et al. Opportunistic energy trading between co-located energy-harvesting wireless sensor networks. In *Proc. of the 1st International Workshop on Energy Neutral Sensing Systems*, page 11, 2013.
[4] Eli De Poorter et al. A negotiation-based networking methodology to enable cooperation across heterogeneous co-located networks. *Ad Hoc Networks*, vol. 10(6):901–917, 2012.
[5] Aman Kansal et al. Power management in energy harvesting sensor networks. *ACM Transactions on Embedded Computing Systems*, vol. 6(4):32, 2007.
[6] Christian Renner et al. Adaptive energy-harvest profiling to enhance depletion-safe operation and efficient task scheduling. *Sustainable Computing: Informatics and Systems*, vol. 2(1):43–56, 2012.
[7] Jason Hsu et al. Adaptive duty cycling for energy harvesting systems. In *Proc. of the 2006 International Symposium on Low power electronics and design*, pages 180–185.
[8] Zhi Ang Eu et al. Opportunistic routing in wireless sensor networks powered by ambient energy harvesting. *Computer Networks*, vol. 54(17):2943–2966, 2010.
[9] Victor Lesser et al. *Distributed sensor networks: A multiagent perspective*, volume vol. 9. 2012.
[10] Meritxell Vinyals et al. A survey on sensor networks from a multiagent perspective. *The Computer Journal*, vol. 54(3):455–470, 2011.
[11] Teng Jiang et al. Opportunistic direct interconnection between co-located wireless sensor networks. In *Int. Conf. on Computer Communications and Networks*, pages 1–5, 2013.
[12] Xiaofan Jiang et al. Perpetual environmentally powered sensor networks. In *Proc. of the 4th international symposium on Information processing in sensor networks*, page 65, 2005.
[13] Alam Muddasser et al. A negotiation protocol for multiple interdependent issues negotiation over energy exchange. In *Proc. of the AI for an Intelligent Planet*, page 1, 2011.
[14] John F. Nash Jr. The bargaining problem. *Econometrica: Journal of the Econometric Society*, pages 155–162, 1950.
[15] Peyman Faratin et al. Negotiation decision functions for autonomous agents. *Robotics and Autonomous Systems*, vol. 24(3-4):159–182, 1998.
[16] Nicholas R. Jennings et al. Automated negotiation: prospects, methods and challenges. *Group Decision and Negotiation*, vol. 10(2):199–215, 2001.
[17] Ariel Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica: Journal of the Econometric Society*, pages 97–109, 1982.
[18] MEMSIC Corporation, Inc. http://www.memsic.com/. Accessed on March 01, 2018.
[19] Southampton Weather. http://www.southamptonweather.co.uk/wxabout.php. Accessed on March 01, 2018.
[20] PVGIS. http://re.jrc.ec.europa.eu/pvgis/apps4/pvest.php. Accessed on March 01, 2018.
[21] Shad Roundy et al. Power sources for wireless sensor networks. In *European workshop on wireless sensor networks*, pages 1–17, 2004.

References

[1] Andre P Ortega, Geoff V Merrett, and Sarvapali D Ramchurn. Automated negotiation for opportunistic energy trading between neighbouring wireless sensor networks. In *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pages 1–6. IEEE, 2018.

[2] Gabriel Martins Dias, Simon Oechsner, and Boris Bellalta. A centralized mechanism to make predictions based on data from multiple wsns. In *International Workshop on Multiple Access Communications*, pages 19–32. Springer, 2015.

[3] Huma Zia, Nick R Harris, Geoff V Merrett, Mark Rivers, and Neil Coles. The impact of agricultural activities on water quality: A case for collaborative catchment-scale management using integrated wireless sensor networks. *Computers and electronics in agriculture*, 96:126–138, 2013.

[4] Krongboon Singhanat, Teng Jiang, Geoff V Merrett, and Nick R Harris. Empirical evaluation of oi-mac: Direct interconnection between wireless sensor networks for collaborative monitoring. In *Sensors Applications Symposium (SAS), 2015 IEEE*, pages 1–5. IEEE, 2015.

[5] Simon Oechsner, Boris Bellalta, Desislava Dimitrova, and Tobias Hoßfeld. Visions and challenges for sensor network collaboration in the cloud. In *IMIS*, pages 16–22, 2014.

[6] Eli De Poorter, Benoît Latré, Ingrid Moerman, and Piet Demeester. Symbiotic networks: Towards a new level of cooperation between wireless networks. *Wireless Personal Communications*, 45(4):479–495, 2008.

[7] Tim De Pauw, Bruno Volckaert, Anna Hristoskova, Veerle Ongenae, and Filip De Turck. Symbiotic service composition in distributed sensor networks. *International Journal of Distributed Sensor Networks*, 9(12):684563, 2013.

[8] Teng Jiang, Geoff V Merrett, and Nick R Harris. Opportunistic energy trading between co-located energy-harvesting wireless sensor networks. In *Proceedings of the 1st International Workshop on Energy Neutral Sensing Systems*, page 11. ACM, 2013.

[9] MJ Shamani, Hossein Gharaee, Sahba Sadri, and Fereidoon Rezaei. Adaptive energy aware cooperation strategy in heterogeneous multi-domain sensor networks. *Procedia Computer Science*, 19:1047–1052, 2013.

[10] Eli De Poorter, Pieter Becue, Milos Rovcanin, Ingrid Moerman, and Piet Demeester. A negotiation-based networking methodology to enable cooperation across heterogeneous co-located networks. *Ad Hoc Networks*, 10(6):901–917, 2012.

[11] Kemal Bicakci and Bulent Tavli. Prolonging network lifetime with multi-domain cooperation strategies in wireless sensor networks. *Ad Hoc Networks*, 8(6):582–596, 2010.

[12] Levente Buttyán, Tamás Holczer, and Péter Schaffer. Spontaneous cooperation in multi-domain sensor networks. In *European Workshop on Security in Ad-hoc and Sensor Networks*, pages 42–53. Springer, 2005.

[13] Márk Félegyházi, J-P Hubaux, and Levente Buttyán. Cooperative packet forwarding in multi-domain sensor networks. In *Pervasive Computing and Communications Workshops, 2005. PerCom 2005 Workshops. Third IEEE International Conference on*, pages 345–349. IEEE, 2005.

[14] Kazuhiko Kinoshita, Natsuki Inoue, Yosuke Tanigawa, Hideki Tode, and Takashi Watanabe. Fair routing for overlapped cooperative heterogeneous wireless sensor networks. *IEEE Sensors Journal*, 16(10):3981–3988, 2016.

[15] Garth V Crosby and Niki Pissinou. Evolution of cooperation in multi-class wireless sensor networks. In *32nd IEEE Conference on Local Computer Networks (LCN 2007)*, pages 489–495. IEEE, 2007.

[16] Pedro OS Vaz de Melo, Felipe D Cunha, and Antonio AF Loureiro. A distributed protocol for cooperation among different wireless sensor networks. In *2013 IEEE International Conference on Communications (ICC)*, pages 6035–6039. IEEE, 2013.

[17] Florent Garcin, Mohammad Hossein Manshaei, and Jean-Pierre Hubaux. Cooperation in underwater sensor networks. In *Game Theory for Networks, 2009. GameNets' 09. International Conference on*, pages 540–548. IEEE, 2009.

[18] Aman Kansal, Jason Hsu, Sadaf Zahedi, and Mani B Srivastava. Power management in energy harvesting sensor networks. *ACM Transactions on Embedded Computing Systems (TECS)*, 6(4):32, 2007.

[19] Dong Kun Noh, Lili Wang, Yong Yang, Hieu Khac Le, and Tarek Abdelzaher. Minimum variance energy allocation for a solar-powered sensor system. In *International Conference on Distributed Computing in Sensor Systems*, pages 44–57. Springer, 2009.

[20] Dong Kun Noh and Kyungtae Kang. A practical flow control scheme considering optimal energy allocation in solar-powered wsns. In *2009 Proceedings of 18th International Conference on Computer Communications and Networks*, pages 1–6. IEEE, 2009.

[21] Chin Keong Ho and Rui Zhang. Optimal energy allocation for wireless communications with energy harvesting constraints. *IEEE Transactions on Signal Processing*, 60(9):4808–4818, 2012.

[22] Shaobo Mao, Man Hon Cheung, and Vincent WS Wong. An optimal energy allocation algorithm for energy harvesting wireless sensor networks. In *2012 IEEE International Conference on Communications (ICC)*, pages 265–270. IEEE, 2012.

[23] Shengbo Chen, Prasun Sinha, Ness B Shroff, and Changhee Joo. A simple asymptotically optimal joint energy allocation and routing scheme in rechargeable sensor networks. *IEEE/ACM Transactions on Networking*, 22(4):1325–1336, 2013.

[24] Mohammad Hossein Anisi, Gaddafi Abdul-Salaam, Mohd Yamani Idna Idris, Ainuddin Wahid Abdul Wahab, and Ismail Ahmedy. Energy harvesting and battery power based routing in wireless sensor networks. *Wireless Networks*, 23(1):249–266, 2017.

[25] Aliyu Aliyu Babayo, Mohammad Hossein Anisi, and Ihsan Ali. A review on energy management schemes in energy harvesting wireless sensor networks. *Renewable and Sustainable Energy Reviews*, 76:1176–1184, 2017.

[26] Mohamed Kashef and Anthony Ephremides. Optimal packet scheduling for energy harvesting sources on time varying wireless channels. *Journal of Communications and Networks*, 14(2):121–129, 2012.

[27] Shengbo Chen, Prasun Sinha, Ness B Shroff, and Changhee Joo. Finite-horizon energy allocation and routing scheme in rechargeable sensor networks. In *INFOCOM, 2011 Proceedings IEEE*, pages 2273–2281. IEEE, 2011.

[28] Marios Gatzianas, Leonidas Georgiadis, and Leandros Tassiulas. Control of wireless networks with rechargeable batteries. *IEEE Trans. Wireless Commun*, 9(2):581–593, 2010.

[29] Lianyou Jing, Chengbing He, Jianguo Huang, and Zhi Ding. Energy management and power allocation for underwater acoustic sensor network. *IEEE Sensors Journal*, 17(19):6451–6462, 2017.

[30] Shaobo Mao, Man Hon Cheung, and Vincent WS Wong. Joint energy allocation for sensing and transmission in rechargeable wireless sensor networks. *IEEE Transactions on Vehicular Technology*, 63(6):2862–2875, 2014.

[31] Longbo Huang and Michael J Neely. Utility optimal scheduling in energy-harvesting networks. *IEEE/ACM Transactions on Networking (TON)*, 21(4):1117–1130, 2013.

[32] Jason Hsu, Sadaf Zahedi, Aman Kansal, Mani Srivastava, and Vijay Raghunathan. Adaptive duty cycling for energy harvesting systems. In *Proceedings of the 2006 international symposium on Low power electronics and design*, pages 180–185. ACM, 2006.

[33] Wai Hong Ronald Chan, Pengfei Zhang, Ido Nevat, Sai Ganesh Nagarajan, Alvin C Valera, Hwee-Xian Tan, and Natarajan Gautam. Adaptive duty cycling in sensor networks with energy harvesting using continuous-time markov chain and fluid models. *IEEE Journal on Selected Areas in Communications*, 33(12):2687–2700, 2015.

[34] Yongmin Zhang, Shibo He, and Jiming Chen. Data gathering optimization by dynamic sensing and routing in rechargeable sensor networks. *IEEE/ACM Transactions on Networking*, 24(3):1632–1646, 2015.

[35] Christian Renner and Volker Turau. Adaptive energy-harvest profiling to enhance depletion-safe operation and efficient task scheduling. *Sustainable Computing: Informatics and Systems*, 2(1):43–56, 2012.

[36] Bo Zhang, Robert Simon, and Hakan Aydin. Maximum utility rate allocation for energy harvesting wireless sensor networks. In *Proceedings of the 14th ACM international conference on Modeling, analysis and simulation of wireless and mobile systems*, pages 7–16. ACM, 2011.

[37] Cesare Alippi, Giuseppe Anastasi, Mario Di Francesco, and Manuel Roveri. Energy management in wireless sensor networks with energy-hungry sensors. *IEEE Instrumentation & Measurement Magazine*, 12(2):16–23, 2009.

[38] Vana Jelicic. Power management in wireless sensor networks with high-consuming sensors. *Qualifying Doctoral Examination, University of Zagreb*, 2011.

[39] Zhi Ang Eu and Hwee-Pink Tan. Adaptive opportunistic routing protocol for energy harvesting wireless sensor networks. In *2012 IEEE international conference on communications (ICC)*, pages 318–322. IEEE, 2012.

[40] Zhi Ang Eu, Hwee-Pink Tan, and Winston KG Seah. Opportunistic routing in wireless sensor networks powered by ambient energy harvesting. *Computer Networks*, 54(17):2943–2966, 2010.

[41] Pedro OS Vaz de Melo, Felipe D da Cunha, Jussara M Almeida, Antonio AF Loureiro, and Raquel AF Mini. The problem of cooperation among different wireless sensor networks. In *Proceedings of the 11th international symposium on Modeling, analysis and simulation of wireless and mobile systems*, pages 86–91. ACM, 2008.

[42] Kemal Bicakci, Ibrahim Ethem Bagci, Bulent Tavli, and Zeydin Pala. Neighbor sensor networks: Increasing lifetime and eliminating partitioning through cooperation. *Computer Standards & Interfaces*, 35(4):396–402, 2013.

[43] Natsuki Inoue, Kazuhiko Kinoshita, Takashi Watanabe, Koso Murakami, Yosuke Tanigawa, and Hideki Tode. A cooperative routing method with shared nodes for overlapping wireless sensor networks. In *2014 International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 1106–1111. IEEE, 2014.

[44] Sara Berri, Vineeth Varma, Samson Lasaulce, Mohammed Said Radjef, and Jamal Daafouz. Studying node cooperation in reputation based packet forwarding within mobile ad hoc networks. In *International Symposium on Ubiquitous Networking*, pages 3–13. Springer, 2017.

[45] Zahra Mohamad Nezhad and Siavash Khorsandi. Cooperation enforcement based on dynamic pricing in multi-domain sensor network. In *Consumer Communications and Networking Conference (CCNC), 2011 IEEE*, pages 1055–1060. IEEE, 2011.

[46] Milos Rovcanin, Eli De Poorter, Daniel van den Akker, Ingrid Moerman, Piet Demeester, and Chris Blondia. Experimental validation of a reinforcement learning based approach for a service-wise optimisation of heterogeneous wireless sensor networks. *Wireless Networks*, 21(3):931–948, 2015.

[47] Steve Munroe and Michael Luck. Motivation-based selection of negotiation opponents. In *International Workshop on Engineering Societies in the Agents World*, pages 119–138. Springer, 2004.

[48] Shaheen S Fatima, Michael Wooldridge, and Nicholas R Jennings. The influence of information on negotiation equilibrium. In *International Workshop on Agent-Mediated Electronic Commerce*, pages 180–193. Springer, 2002.

[49] Serban Radu, Eugenia Kalisz, and Adina Magda Florea. A model of automated negotiation based on agents profiles. *Scalable Computing: Practice and Experience*, 14(1):47–56, 2013.

[50] Jakub Brzostowski and Ryszard Kowalczyk. On possibilistic case-based reasoning for selecting partners for multi-attribute agent negotiation. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, pages 273–279. ACM, 2005.

[51] Victor Lesser, Charles L Ortiz Jr, and Milind Tambe. *Distributed sensor networks: A multiagent perspective*, volume 9. Springer Science & Business Media, 2012.

[52] Paritosh Padhy, Rajdeep K Dash, Kirk Martinez, and Nicholas R Jennings. A utility-based sensing and communication model for a glacial sensor network. In *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, pages 1353–1360, 2006.

[53] Meritxell Vinyals, Juan A Rodriguez-Aguilar, and Jesus Cerquides. A survey on sensor networks from a multiagent perspective. *The Computer Journal*, 54(3):455–470, 2010.

[54] Teng Jiang. Opportunistic direct interconnection and cooperation between co-located wireless sensor networks. In *Ph.D. dissertation*. University of Southampton, 2015.

[55] Takayuki Ito, Hiromitsu Hattori, and Mark Klein. Multi-issue negotiation protocol for agents: Exploring nonlinear utility spaces. In *IJCAI*, volume 7, pages 1347–1352, 2007.

[56] Teng Jiang, Geoff V Merrett, and Nick R Harris. Opportunistic direct interconnection between co-located wireless sensor networks. In *2013 22nd International Conference on Computer Communication and Networks (ICCCN)*, pages 1–5. IEEE, 2013.

[57] Krongboon Singhanat, Nick R Harris, and Geoff V Merrett. Experimental validation of opportunistic direct interconnection between different wireless sensor networks. In *2016 IEEE Sensors Applications Symposium (SAS)*, pages 1–6. IEEE, 2016.

[58] Milos Rovcanin, Eli De Poorter, Ingrid Moerman, and Piet Demeester. An lspi based reinforcement learning approach to enable network cooperation in cognitive wireless sensor network. In *Advanced Information Networking and Applications Workshops (WAINA), 2013 27th International Conference on*, pages 82–89. IEEE, 2013.

[59] Min Chen, Sergio Gonzalez, Athanasios Vasilakos, Huasong Cao, and Victor CM Leung. Body area networks: A survey. *Mobile networks and applications*, 16(2):171–193, 2011.

[60] Giancarlo Fortino, Stefano Galzarano, Raffaele Gravina, and Wenfeng Li. A framework for collaborative computing and multi-sensor data fusion in body sensor networks. *Information Fusion*, 22:50–70, 2015.

[61] Jun Wu, Mianxiong Dong, Kaoru Ota, Muhammad Tariq, and Longhua Guo. Cross-domain fine-grained data usage control service for industrial wireless sensor networks. *IEEE Access*, 3:2939–2949, 2015.

[62] Carlo Fischione, Karl Henrik Johansson, Fabio Graziosi, and Fortunato Santucci. Distributed cooperative processing and control over wireless sensor networks. In *Proceedings of the 2006 international conference on Wireless communications and mobile computing*, pages 1311–1316. ACM, 2006.

[63] Chengpei Tang, Sanesy Kumcr Shokla, George Modhawar, and Qiang Wang. An effective collaborative mobile weighted clustering schemes for energy balancing in wireless sensor networks. *Sensors*, 16(2):261, 2016.

[64] Virginia Pilloni, Pirabakaran Navaratnam, Serdar Vural, Luigi Atzori, and Rahim Tafazolli. Cooperative task assignment for distributed deployment of applications in wsns. In *Communications (ICC), 2013 IEEE International Conference on*, pages 2229–2234. IEEE, 2013.

[65] Chongmyung Park, Youngtae Jo, and Inbum Jung. Cooperative processing model for wireless sensor networks. *International Journal of Distributed Sensor Networks*, 9(9):317214, 2013.

[66] Vidyasagar Potdar, Atif Sharif, and Elizabeth Chang. Wireless sensor networks: A survey. In *Advanced Information Networking and Applications Workshops, 2009. WAINA'09. International Conference on*, pages 636–641. IEEE, 2009.

[67] J Nagata, Y Tanigawa, K Kinoshita, H Tode, and K Murakami. A routing method for cooperative forwarding in multiple wireless sensor networks. In *Proceedings of International Conference on Networking and Services*, pages 43–46, 2012.

[68] Bora Karaoglu, Ilker Demirkol, and Wendi Heinzelman. Exploring the benefits of symbiotic routing. In *Computer Communications and Networks (ICCCN), 2011 Proceedings of 20th International Conference on*, pages 1–6. IEEE, 2011.

[69] Drew Fudenberg and Jean Tirole. Game theory mit press. *Cambridge, MA*, page 86, 1991.

[70] R Duncan Luce, Howard Raiffa, and T Teichmann. Games and decisions. *Physics Today*, 11:33, 1958.

[71] Yuanjie Li and Xiaojun Wu. Cooperative packet-forwarding strategies in mobile ad hoc networks with unreliable channels: An evolutionary game approach. *International Journal of Distributed Sensor Networks*, 15(9):1550147719875651, 2019.

[72] John F Nash et al. Equilibrium points in n-person games. *Proceedings of the national academy of sciences*, 36(1):48–49, 1950.

[73] Ariel Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica: Journal of the Econometric Society*, pages 97–109, 1982.

[74] Nicholas R Jennings, Peyman Faratin, Alessio R Lomuscio, Simon Parsons, Michael J Wooldridge, and Carles Sierra. Automated negotiation: prospects, methods and challenges. *Group Decision and Negotiation*, 10(2):199–215, 2001.

[75] Alessio R Lomuscio, Michael Wooldridge, and Nicholas R Jennings. A classification scheme for negotiation in electronic commerce. *Group Decision and Negotiation*, 12(1):31–56, 2003.

[76] Shaheen Fatima, Sarit Kraus, and Michael Wooldridge. *Principles of automated negotiation*. Cambridge University Press, 2014.

[77] Alvin E Roth. Lecture notes in economics and mathematical systems. 1979.

[78] John F Nash Jr. The bargaining problem. *Econometrica: Journal of the Econometric Society*, pages 155–162, 1950.

[79] Ken Binmore. Fun and games, a text on game theory. 1992.

[80] Andreu Mas-Colell, Michael Dennis Whinston, Jerry R Green, et al. *Microeconomic theory*, volume 1. Oxford university press New York, 1995.

[81] Ehud Kalai and Meir Smorodinsky. Other solutions to nash's bargaining problem. *Econometrica: Journal of the Econometric Society*, pages 513–518, 1975.

[82] Kenneth J Arrow, Amartya Sen, and Kotaro Suzumura. *Handbook of social choice and welfare*, volume 2. Elsevier, 2010.

[83] Roger B Myerson. Utilitarianism, egalitarianism, and the timing effect in social choice problems. *Econometrica: Journal of the Econometric Society*, pages 883–897, 1981.

[84] JS Rosenchein and Gilad Zlotkin. Rules of encounter, 1994.

[85] Shaheen Fatima, Sarit Kraus, and Michael Wooldridge. The negotiation game. *IEEE Intelligent Systems*, 29(5):57–61, 2014.

[86] Ariel Rubinstein. A bargaining model with incomplete information about time preferences. *Econometrica: Journal of the Econometric Society*, pages 1151–1172, 1985.

[87] Guoming Lai, Cuihong Li, Katia Sycara, and Joseph Giampapa. Literature review on multi-attribute negotiations. *Robotics Inst., Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep. CMU-RI-TR-04-66*, 2004.

[88] Peyman Faratin. *Automated service negotiation between autonomous computational agents*. PhD thesis, Queen Mary University of London, 2000.

[89] Carles Sierra, Peyman Faratin, and Nick R Jennings. A service-oriented negotiation model between autonomous agents. In *Collaboration between human and artificial societies*, pages 201–219. Springer, 1999.

[90] Peyman Faratin, Carles Sierra, and Nick R Jennings. Negotiation decision functions for autonomous agents. *Robotics and Autonomous Systems*, 24(3-4):159–182, 1998.

[91] Dean G Pruitt. *Negotiation behavior*. Academic Press, 2013.

[92] Howard Raiffa. *The art and science of negotiation*. Harvard University Press, 1982.

[93] Peyman Faratin, Carles Sierra, and Nicholas R Jennings. Using similarity criteria to make issue trade-offs in automated negotiations. *artificial Intelligence*, 142(2):205–237, 2002.

[94] Ralph L Keeney and Howard Raiffa. *Decisions with multiple objectives: preferences and value trade-offs*. Cambridge university press, 1993.

[95] Linlin Wu, Saurabh Kumar Garg, Rajkumar Buyya, Chao Chen, and Steve Versteeg. Automated sla negotiation framework for cloud computing. In *2013 13th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing*, pages 235–244. IEEE, 2013.

[96] Xudong Luo, Ho-fung Leung, and Jimmy Ho-Man Lee. A multi-agent framework for meeting scheduling using fuzzy constraints. In *Proceedings Fourth International Conference on MultiAgent Systems*, pages 409–410. IEEE, 2000.

[97] Xudong Luo, Nicholas R Jennings, Nigel Shadbolt, Ho-fung Leung, and Jimmy Homan Lee. A fuzzy constraint based model for bilateral, multi-issue negotiations in semicompetitive environments. *Artificial Intelligence*, 148(1-2):53–102, 2003.

[98] Catholijn M Jonker, Valentin Robu, and Jan Treur. An agent architecture for multiattribute negotiation using incomplete preference information. *Autonomous Agents and Multi-Agent Systems*, 15(2):221–252, 2007.

[99] Gerald Tesauro. Efficient search techniques for multi-attribute bilateral negotiation strategies. In *Proceedings. Third International Symposium on Electronic Commerce,*, pages 30–36. IEEE, 2002.

[100] Cuihong Li and Gerald Tesauro. A strategic decision model for multi-attribute bilateral negotiation with alternating. In *Proceedings of the 4th ACM Conference on Electronic Commerce*, pages 208–209. ACM, 2003.

[101] Cuihong Li, Joseph Giampapa, and Katia Sycara-Cyranski. A review of research literature on bilateral negotiations. 2003.

[102] Bo An and Victor Lesser. Yushu: a heuristic-based agent for automated negotiating competition. In *New Trends in Agent-Based Complex Automated Negotiations*, pages 145–149. Springer, 2012.

[103] Liviu Dan Şerban, Gheorghe Cosmin Silaghi, and Cristian Marius Litan. Agentfsega: time constrained reasoning model for bilateral multi-issue negotiations. In *New Trends in Agent-Based Complex Automated Negotiations*, pages 159–165. Springer, 2012.

[104] Jianye Hao and Ho-Fung Leung. Abines: An adaptive bilateral negotiating strategy over multiple items. In *Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology-Volume 02*, pages 95–102. IEEE Computer Society, 2012.

[105] Shogo Kawaguchi, Katsuhide Fujita, and Takayuki Ito. Agentk: Compromising strategy based on estimated maximum utility for automated negotiating agents. In *New Trends in Agent-Based Complex Automated Negotiations*, pages 137–144. Springer, 2012.

[106] Ronghuo Zheng, Nilanjan Chakraborty, and Sycara K Dai T. Automated multiagent negotiation on multiple issues with private information, 2013.

[107] Ronghuo Zheng, Nilanjan Chakraborty, Tinglong Dai, and Katia Sycara. Multiagent negotiation on multiple issues with incomplete information. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, pages 1279–1280. International Foundation for Autonomous Agents and Multiagent Systems, 2013.

[108] Ronghuo Zheng, Tinglong Dai, Katia Sycara, and Nilanjan Chakraborty. Automated multilateral negotiation on multiple issues with private information. *INFORMS Journal on Computing*, 28(4):612–628, 2016.

[109] Aodah Diamah, Michael Wagner, and Menkes van den Briel. A comparative study on vector similarity methods for offer generation in multi-attribute negotiation. In *Australasian Joint Conference on Artificial Intelligence*, pages 149–156. Springer, 2015.

[110] Ronghuo Zheng, Ying Xu, Nilanjan Chakraborty, and Katia Sycara. Multiagent coordination for demand management with energy generation and storage. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pages

1587–1588. International Foundation for Autonomous Agents and Multiagent Systems, 2014.

[111] Claudia Di Napoli, Dario Di Nocera, and Silvia Rossi. Computing pareto optimal agreements in multi-issue negotiation for service composition. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 1779–1780. International Foundation for Autonomous Agents and Multiagent Systems, 2015.

[112] Lei Niu, Fenghui Ren, and Minjie Zhang. Feasible negotiation procedures for multiple interdependent negotiations. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pages 641–649. International Foundation for Autonomous Agents and Multiagent Systems, 2018.

[113] DJA Somefun, Enrico H Gerding, Sander Bohte, and Johannes A La Poutré. Automated negotiation and bundling of information goods. In *International Workshop on Agent-Mediated Electronic Commerce*, pages 1–17. Springer, 2003.

[114] DJA Somefun, Enrico H Gerding, and Johannes A La Poutré. Efficient methods for automated multi-issue negotiation: Negotiating over a two-part tariff. *International Journal of Intelligent Systems*, 21(1):99–119, 2006.

[115] Guoming Lai and Katia Sycara. A generic framework for automated multi-attribute negotiation. *Group Decision and Negotiation*, 18(2):169, 2009.

[116] Mengxiao Wu, Mathijs de Weerdt, and Han La Poutré. Efficient methods for multi-agent multi-issue negotiation: Allocating resources. In *International Conference on Principles and Practice of Multi-Agent Systems*, pages 97–112. Springer, 2009.

[117] Ronghuo Zheng, Nilanjan Chakraborty, Tinglong Dai, Katia Sycara, and Michael Lewis. Automated bilateral multiple-issue negotiation with no information about opponent. In *2013 46th Hawaii International Conference on System Sciences*, pages 520–527. IEEE, 2013.

[118] Hongseok Yoo, Moonjoo Shim, and Dongkyun Kim. Dynamic duty-cycle scheduling schemes for energy-harvesting wireless sensor networks. *IEEE communications letters*, 16(2):202–204, 2011.

[119] Deyu Zhang, Zhigang Chen, Ju Ren, Ning Zhang, Mohamad Khattar Awad, Haibo Zhou, and Xuemin Sherman Shen. Energy-harvesting-aided spectrum sensing and data transmission in heterogeneous cognitive radio sensor network. *IEEE Transactions on Vehicular Technology*, 66(1):831–843, 2016.

[120] Muddasser Alam, Enrico H Gerding, Alex Rogers, and Sarvapali D Ramchurn. A scalable interdependent multi-issue negotiation protocol for energy exchange. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.

[121] Muddasser Alam, Alex Rogers, and Sarvapali D Ramchurn. A negotiation protocol for multiple interdependent issues negotiation over energy exchange. In *Proceedings of the AI for an Intelligent Planet*, page 1. ACM, 2011.

[122] Ting Zhu, Ziguo Zhong, Yu Gu, Tian He, and Zhi-Li Zhang. Leakage-aware energy synchronization for wireless sensor networks. In *Proceedings of the 7th international conference on Mobile systems, applications, and services*, pages 319–332. ACM, 2009.

[123] Milos Rovcanin, Eli De Poorter, Ingrid Moerman, and Piet Demeester. A reinforcement learning based solution for cognitive network cooperation between co-located, heterogeneous wireless sensor networks. *Ad Hoc Networks*, 17:98–113, 2014.

[124] Xuedong Liang, Ilangko Balasingham, and Sang-Seon Byun. A reinforcement learning based routing protocol with qos support for biomedical sensor networks. In *2008 First International Symposium on Applied Sciences on Biomedical and Communication Technologies*, pages 1–5. IEEE, 2008.

[125] Shaoqiang Dong, Prathima Agrawal, and Krishna Sivalingam. Reinforcement learning based geographic routing protocol for uwb wireless sensor network. In *IEEE GLOBECOM 2007-IEEE Global Telecommunications Conference*, pages 652–656. IEEE, 2007.

[126] Mohammad Abu Alsheikh, Shaowei Lin, Dusit Niyato, and Hwee-Pink Tan. Machine learning in wireless sensor networks: Algorithms, strategies, and applications. *IEEE Communications Surveys & Tutorials*, 16(4):1996–2018, 2014.

[127] Reza GhasemAghaei, Md Abdur Rahman, Wail Gueaieb, and Abdulmotaleb El Saddik. Ant colony-based reinforcement learning algorithm for routing in wireless sensor networks. In *2007 IEEE Instrumentation & Measurement Technology Conference IMTC 2007*, pages 1–6. IEEE, 2007.

[128] Ping Wang and Ting Wang. Adaptive routing for sensor networks using reinforcement learning. In *The Sixth IEEE International Conference on Computer and Information Technology (CIT'06)*, pages 219–219. IEEE, 2006.

[129] Rocio Arroyo-Valles, Rocio Alaiz-Rodriguez, Alicia Guerrero-Curieses, and Jesús Cid-Sueiro. Q-probabilistic routing in wireless sensor networks. In *2007 3rd International Conference on Intelligent Sensors, Sensor Networks and Information*, pages 1–6. IEEE, 2007.

[130] Long Tran-Thanh, Alex Rogers, and Nicholas R Jennings. Long-term information collection with energy harvesting wireless sensors: a multi-armed bandit based approach. *Autonomous Agents and Multi-Agent Systems*, 25(2):352–394, 2012.

[131] Jiang Zhu, Yonghui Song, Dingde Jiang, and Houbing Song. Multi-armed bandit channel access scheme with cognitive radio technology in wireless sensor networks for the internet of things. *IEEE access*, 4:4609–4617, 2016.

[132] Glacsweb project. Glacsweb. https://glacsweb.org/technology/probe/. Last accessed on July 01, 2017.

[133] Zhao Cheng, Mark Perillo, and Wendi B Heinzelman. General network lifetime and cost models for evaluating sensor network deployment strategies. *IEEE Transactions on mobile computing*, 7(4):484–497, 2008.

[134] Xiaofan Jiang, Joseph Polastre, and David Culler. Perpetual environmentally powered sensor networks. In *Proceedings of the 4th international symposium on Information processing in sensor networks*, page 65. IEEE Press, 2005.

[135] Ren-Shiou Liu, Kai-Wei Fan, Zizhan Zheng, and Prasun Sinha. Perpetual and fair data collection for environmental energy harvesting sensor networks. *IEEE/ACM Transactions on Networking (TON)*, 19(4):947–960, 2011.

[136] Zhaoyang Zhang, Jing Shi, Hsiao-Hwa Chen, Mohsen Guizani, and Peiliang Qiu. A cooperation strategy based on nash bargaining solution in cooperative relay networks. *IEEE Transactions on Vehicular Technology*, 57(4):2570–2577, 2008.

[137] MEMSIC Corporation, Inc. Memsic powerful sensing solutions. http://www.memsic.com/. Last accessed on July 01, 2017.

[138] MEMSIC Corporation, Inc. Memsic: eko outdoor wireless system datasheet. http://www.memsic.com/userfiles/files/Datasheets/WSN/eko_starter_system.pdf. Last accessed on July 01, 2017.

[139] Ioannis Hadjipaschalis, Andreas Poullikkas, and Venizelos Efthimiou. Overview of current and future energy storage technologies for electric power applications. *Renewable and sustainable energy reviews*, 13(6-7):1513–1522, 2009.

[140] Southampton Weather. About the southampton weather station. http://www.southamptonweather.co.uk/wxabout.php. Last accessed on March 01, 2018.

[141] Photovoltaic Geographical Information System. Pvgis solar data (eu). http://re.jrc.ec.europa.eu/pvgis/apps4/pvest.php. Last accessed on March 01, 2018.

[142] Weather Underground. https://api.weather.com/v1/geocode/50.83555603/-0.29722199/observations/historical.json. Accessed on April 01, 2019.

[143] Shad Roundy, Dan Steingart, Luc Frechette, Paul Wright, and Jan Rabaey. Power sources for wireless sensor networks. In *European workshop on wireless sensor networks*, pages 1–17. Springer, 2004.

[144] Davide Carli, Davide Brunelli, Davide Bertozzi, and Luca Benini. A high-efficiency wind-flow energy harvester using micro turbine. In *SPEEDAM 2010*, pages 778–783. IEEE, 2010.

[145] Christopher M Vigorito, Deepak Ganesan, and Andrew G Barto. Adaptive control of duty cycling in energy-harvesting wireless sensor networks. In *Sensor, Mesh and Ad Hoc Communications and Networks, 2007. SECON'07. 4th Annual IEEE Communications Society Conference on*, pages 21–30. IEEE, 2007.

[146] Prabal Dutta, Mike Grimmer, Anish Arora, Steven Bibyk, and David Culler. Design of a wireless sensor network platform for detecting rare, random, and ephemeral events. In *Proceedings of the 4th international symposium on Information processing in sensor networks*, page 70. IEEE Press, 2005.

[147] Teng Jiang et al. Opportunistic direct interconnection between co-located wireless sensor networks. In *Int. Conf. on Computer Communications and Networks*, pages 1–5, 2013.

[148] Andras Varga. Omnet++. In *Modeling and tools for network simulation*, pages 35–59. Springer, 2010.

[149] INET Framework. Inet framework. https://inet.omnetpp.org/. Last accessed on January 30, 2020.

[150] Seokho Son, Dong-Jae Kang, Seyoung Phillip Huh, Won-Young Kim, and Wan Choi. Adaptive trade-off strategy for bargaining-based multi-objective sla establishment under varying cloud workload. *The Journal of Supercomputing*, 72(4):1597–1622, 2016.

[151] Srikumar Venugopal, Xingchen Chu, and Rajkumar Buyya. A negotiation mechanism for advance resource reservations using the alternate offers protocol. In *2008 16th Interntional Workshop on Quality of Service*, pages 40–49. IEEE, 2008.

[152] Siqi Chen and Gerhard Weiss. An efficient automated negotiation strategy for complex environments. *Engineering Applications of Artificial Intelligence*, 26(10):2613–2623, 2013.

[153] Xianrong Zheng, Patrick Martin, Kathryn Brohman, and Li Da Xu. Cloud service negotiation in internet of things environment: A mixed approach. *IEEE Transactions on Industrial Informatics*, 10(2):1506–1515, 2014.

[154] Bo An, Victor Lesser, David Irwin, and Michael Zink. Automated negotiation with decommitment for dynamic resource allocation in cloud computing. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, pages 981–988. International Foundation for Autonomous Agents and Multiagent Systems, 2010.

[155] Kwang Mong Sim. Grid resource negotiation: survey and new directions. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(3):245–257, 2010.

[156] Jiadao Li and Ramin Yahyapour. Negotiation model supporting co-allocation for grid scheduling. In *Proceedings of the 7th IEEE/ACM International Conference on Grid Computing*, pages 254–261. IEEE Computer Society, 2006.

[157] Reyhan Aydogan, Tim Baarslag, Catholijn M Jonker, Katsuhide Fujita, Takayuki Ito, Rafik Hadfi, and Kohei Hayakawa. A baseline for non-linear bilateral negotiations: the full results of the agents competing in anac 2014. 2016.

[158] Gerhard Weiss. A modern approach to distributed artificial intelligence. *IEEE transactions on systems man & cybernetics-part c applications & reviews*, 22(2), 1999.

[159] Herbert Robbins. Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, 58(5):527–535, 1952.

[160] Haifeng Xu, Long Tran-Thanh, and Nicholas R Jennings. Playing repeated security games with no prior knowledge. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pages 104–112. International Foundation for Autonomous Agents and Multiagent Systems, 2016.

[161] Nicolo Cesa-Bianchi and Gábor Lugosi. Combinatorial bandits. *Journal of Computer and System Sciences*, 78(5):1404–1422, 2012.

[162] Adam Kalai and Santosh Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005.

[163] Gergely Neu and Gábor Bartók. An efficient algorithm for learning with semi-bandit feedback. In *International Conference on Algorithmic Learning Theory*, pages 234–248. Springer, 2013.

[164] Ketema Adere Gemeda, Gabriele Gianini, and Mulugeta Libsie. Collaborative packets forwarding to extend lifetime of multi-authority wireless sensor networks. In *2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)*, pages 513–519. IEEE, 2017.

[165] Marc Barcelo, Alejandro Correa, Jose Lopez Vicario, and Antoni Morell. Cooperative interaction among multiple rpl instances in wireless sensor networks. *Computer Communications*, 81:61–71, 2016.

[166] Kamarul Zaman Panatik, Kamilia Kamardin, Sya Azmeela Shariff, Siti Sophiayati Yuhaniz, Noor Azurati Ahmad, Othman Mohd Yusop, and SaifulAdli Ismail. Energy harvesting in wireless sensor networks: A survey. In *2016 IEEE 3rd International Symposium on Telecommunication Technologies (ISTT)*, pages 53–58. IEEE, 2016.

[167] Jongdeog Lee, Krasimira Kapitanova, and Sang H Son. The price of security in wireless sensor networks. *Computer Networks*, 54(17):2967–2978, 2010.

[168] Amruta More, Sheetal Vij, and Debajyoti Mukhopadhyay. Agent based negotiation using cloud–an approach in e-commerce. In *ICT and Critical Infrastructure: Proceedings of the 48th Annual Convention of Computer Society of India-Vol I*, pages 489–496. Springer, 2014.