



A Cheat-Proof Power Control Policy for Self-Organizing Full-Duplex Small Cells

Prabodini Semasinghe

► To cite this version:

Prabodini Semasinghe. A Cheat-Proof Power Control Policy for Self-Organizing Full-Duplex Small Cells. 2015. hal-01121945

HAL Id: hal-01121945

<https://hal.science/hal-01121945>

Preprint submitted on 4 Mar 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Public Domain

Game Theory and Learning Techniques for Self-Organization in Small Cell Networks

Prabodini Semasinghe, Kun Zhu, Ekram Hossain, and Alagan
Anpalagan

Contents

<i>List of illustrations</i>	<i>page</i> 5
<i>List of tables</i>	6
<i>Notation</i>	1

1	Game Theory and Learning Techniques for Self-Organization in Small Cell Networks	3
1.1	Small Cell Networks	3
1.2	Self-Organization	5
1.2.1	Self-Organizing Functionalities	6
1.2.2	Characteristics of Self-Organizing Algorithms	9
1.3	Issues and Challenges in Self-Organizing Small Cell Networks	10
1.4	Game Theory for Self-Organizing Small Cell Networks	12
1.4.1	Fundamentals of Game Theory	12
1.4.2	Motivations of using Game Theory for Self-Organizing Networks	13
1.4.3	Types of Games	13
1.4.4	Price of Anarchy and Price of Stability	17
1.4.5	Design of Payoff Functions	17
1.5	Game Theory-Based Resource Management for Self-Organizing Small Cells	21
1.5.1	Non-cooperative Game-Based Decentralized Power Allocation	21
1.5.2	Non-cooperative game-based sub-carrier allocation	25
1.5.3	Complimentary TRi-Control Loop for Uplink Interference Mitigation	26
1.5.4	An Evolutionary Game Approach for Self-Organization with Reduced Information Exchange	28
1.6	Learning Techniques for Self-Organizing Small Cell Networks	35
1.6.1	Reinforcement Learning	35
1.6.2	Q-Learning	38
1.6.3	Regret-matching Learning	40
1.6.4	Learning by Cooperation	41
1.7	Conclusion	42

<i>References</i>	44
-------------------	----

Illustrations

1.1	A heterogeneous network.	3
1.2	SON functions defined by 3GPP.	7
1.3	SON related use cases defined by NGMN.	7
1.4	Uplink interference scenario in a two-tier network.	26
1.5	Complimentary TRi-control loop.	27
1.6	Small cell cluster underlaid with a macro network.	29

Tables

1.1	Game models for self organizing small cell networks	34
-----	---	----

Notation

3GPP	3 rd Generation Partnership Project
BeFemto	Broadband evolved Femto
CCE	Coarse Correlated Equilibrium
CE	Correlated Equilibrium
CCO	Coverage and Capacity Optimization
CSI	Channel State Information
CTRL	Complimentary TRi-Control Loop
DHCP	Dynamic Host Configuration Protocol
DSL	Digital Subscriber Line
EE	Evolutionary Equilibrium
EGT	Evolutionary Game Theory
FFR	Fractional Frequency Reuse
ICIC	Inter-Cell Interference Coordination
IP	Internet Protocol
LE	Logit Equilibrium
LTE-A	Long Term Evolution-Advanced
MBS	Macro Base Station
MDP	Markov Decision Process
MU	Macro User
NE	Nash Equilibrium
NGMN	Next Generation Mobile Networks
OAM	Operation, Administration and Management
OPEX	Operational Expenditure
PCI	Physical Cell ID
PoA	Price of Anarchy
PoS	Price of Stability
PPP	Poisson Point Process
QoS	Quality of Service
RACH	Random Access Channel
RB	Resource Block
RC	Recursive Core
RL	Reinforcement Learning
SBR	Smoothed Best Response
SBS	Small cell Base Station
SU	Small cell User
SINR	Signal to Interference and Noise Ratio
SO	Self Organization
SOCRATES	Self-Optimization and self-ConfigURATion in wirelEss networkS
SON	Self Organizing Network
UFR	Universal Frequency Reuse

1 Game Theory and Learning Techniques for Self-Organization in Small Cell Networks

1.1 Small Cell Networks

The tremendous increase of bandwidth craving mobile applications (e.g., video streaming, video chatting, and online game) has posed enormous challenges to the design of future wireless networks. Deploying small cells (e.g., pico, micro, and femto) has been shown to be an efficient and cost effective solution to support this constantly rising demand since the smaller cell size can provide higher link quality and more efficient spatial reuse [1]. Small cells could also deliver some other benefits such as offloading the macro network traffic, providing service to coverage holes and the regions with poor signal reception (e.g., macro cell edges). Following this trend, the evolving 5G networks [2] are expected to be composed of hundreds of interconnected heterogeneous small cells.

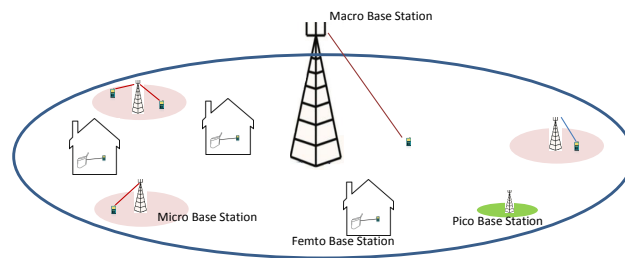


Figure 1.1 A heterogeneous network.

Fig. 1.1 gives an illustration of a heterogeneous network (HetNet) where a macrocell is underlaid with different types of small cells. Different from the cautiously planned traditional network, the architecture of a HetNet is more random and unpredictable due to the increased density of small cells and their impromptu way of deployment. In this case, the manual intervention and centralized control used in traditional network management will be highly inefficient, time consuming, and expensive, and therefore, will be not applicable for dense heterogeneous small cell networks. Instead, **self-organization** has been proposed as an essential feature for future small cell networks [3, 4].

The motivations for enabling self-organization in small cell networks are explained below.

- Numerous network devices with different characteristics are expected to be interconnected in future wireless networks. Also, these devices are expected to have ‘*plug and play*’ capability. Therefore, the initial pre-operational configuration has to be done with minimum expertise involvement.
- With the emergence of the small cells, the spatio-temporal dynamics of the networks has become more unpredictable than legacy systems due to the unplanned nature of the small cell deployment. Therefore, intelligent adaptation of the network nodes is necessary. That is, the self-organizing small cells need to learn from the environment and adapt with the network dynamics to achieve the desired performance.
- Improper or uncoordinated power and spectrum allocation paradigms can lead the small cells to cause severe inter-tier and intra-tier interference. Therefore, resource allocation is a key issue for interference management in heterogeneous small cell networks. Centralized control will be highly inefficient and time consuming for a dense network due to the high computational power and the huge amount of information exchange required. Instead, small cell base stations (SBSs) should be capable of taking individual decisions on resource allocation with local interactions.
- Self-organization of the network will also prevent possible human mistakes in configuration and network management which can drastically degrade the performance of the network and can result the extensively long recovery times. Also, enabling self-organization could reduce a considerable amount of operational and capital expenditure (OPEX/CAPEX).

The Small Cell Forum, which is an organization who supports and promotes the wide-scale adoption of small cell technologies claims that small cells are the first commercial example of a self-organizing network in practice [5].

There are ongoing projects which develops the self-organizing paradigms for small cell networks involving both academia and industry. *BeFemto* (Broadband evolved Femto network) is one such project which focuses on developing femtocell technologies for LTE-A systems [6]. They also plan to provide guidelines for standardization of the next generation femto-cell technologies. *SOCRATES* (*Self-Optimization and self-ConfiguRATION in wirelEss networkS*) also target on developing self-organizing paradigms for small cell networks in 3GPP LTE interface [7, 8]. SOCRATES project was partnered by several leading telecommunication companies in Europe including Nokia Siemens Networks (in Poland and Germany), Vodafone (United Kingdom) and Ericsson AB (Sweden). The End-to-End Efficiency (E^3) [9, 10] works on integrating the heterogeneous network infrastructures into a scalable and efficient cognitive framework with self-organizing capabilities. In addition to that they also focus on research, regulation, and standardization perspectives of cognitive radio networks.

1.2 Self-Organization

The concept of self-organization is not new and can be widely observed in many natural systems and phenomena (e.g., collective behaviors of ants and social insects, flocks of cranes, generation of laser light, and planetary systems). Extensive efforts have been taken by researchers to model the self-organizing behaviour of natural systems mathematically and these models can be borrowed and adapted to develop self-organizing algorithms for artificial systems [11]. First, it is essential to understand the basic properties, requirements, and design concepts of a self-organizing system.

As self-organization is a concept being used in many different fields, the term has been defined in many different ways based on the context. A globally accepted precise and concise definition of self-organizing networks (SONs) has not yet been presented. However, in the area of wireless communication, the standardization of technical specifications for self-organizing LTE and LTE-A networks has been initiated by the 3rd Generation Partnership Project (3GPP) in Release 8 and Release 9 [12, 13] and Next Generation Mobile Networks (NGMN) Alliance [14, 15]. In this section, we will illustrate the concept of self-organization and its basic cornerstones in the framework of cellular networks.

The basis of a self-organizing system is its *autonomous and intelligent adaptivity*, i.e., the ability to respond to external environmental changes. Many literature in the context of wireless networks also suggest that a self-organizing network should be capable of *learning* from environmental dynamics and adapt to it accordingly [16, 17, 18]. Specifically, for small cell networks, detecting the environmental dynamics can be done based on local interactions with other nodes and/or through spectrum sensing. In [19], the authors explain that the adaptive behavior of each member of a self-organizing set should also lead the whole system to form a global pattern which is denoted as the *emergent behavior*.

Based on the above notions, the basic cornerstones of a self-organizing small cell network are identified as follows:

- Autonomous and intelligent adaptivity
- Ability to learn from the environment
- Emergent behavior.

In addition to the aforementioned properties, researchers also discuss about *distributed control* where each node in the network has to take individual decisions on their own behavior. Distributed control is a desirable feature for self-organizing small cell networks. In 3GPP Release 11, the specifications divide self-organizing networks into three categories as given below.

- Centralized SON: Self-organizing algorithms are executed in the Operation, Administration, and Management (OAM) system.
- Distributed SON: Self-organizing algorithms are executed at the network node level.

- Hybrid SON: Algorithms are executed at both OAM and network node levels.

Distributed resource allocation is essential for the provision of distributed control in self-organizing networks. Several distributed resource allocation algorithms for small cell network have been proposed in the recent literature which will be discussed later in this chapter.

1.2.1 Self-Organizing Functionalities

In general, the self-organizing process of a small cell network can be splitted into three phases, i.e., pre-operational phase, operational phase, and failure recovery phase. These three phases commonly correspond to *Self-configuration*, *Self-optimization*, and *Self-healing*, which are also referred to as **Self-X** functionalities [3, 20, 21].

During the standardization process for LTE SON, 3GPP has defined a set of use cases and associated functions in Releases 9, 10 and 11 [22, 23, 24] which are described in Fig. 1.2. Next Generation Mobile Networks (NGMN) Alliance also highlights several operational use cases for the introduction of SON features for mobile networks. NGMN divides the SON related use cases into four categories i.e., planning, deployment, monitoring, and maintenance [25]. However, most of the steps in planing are not covered by SON functions, therefore, we only list the use cases of latter three categories in Fig. 1.3 which are similar to the 3GPP SON user cases. The network parameters such as neighborhood list and handover settings are considered as radio parameters while IP addresses and QoS requirements are considered as transport parameters.

A brief overview of the operation and associated functions of each phase of self-organization is given below.

- *Self-configuration*: Self-configuration is performed in the pre-operational process during which the small cell base stations (SBSs) connect to the network and execute their initialization algorithms automatically while providing plug-and-play capabilities to the network nodes. This functionality is composed of basic set-up of the base station and the initialization of network parameter settings.

Specifically, an SBS is expected to automatically configure its IP address once it is connected to the network. This can be done by using the Dynamic Host Configuration Protocol (DHCP). Then the SBS can communicate with the OAM center and small cell access gateway for authentication. This procedure is called automatic inventory. Once the SBS is connected to the core network it can download and install the required software (i.e., automatic software download). The SBSs are also expected to set transport parameters such as transport layer QoS setting and radio parameters (e.g., neighborhood list and handover settings). The assignment of a PCI (Physical Cell ID) is also done in the self-configuration phase.

As small cells are usually deployed in the coverage area of macro cells,

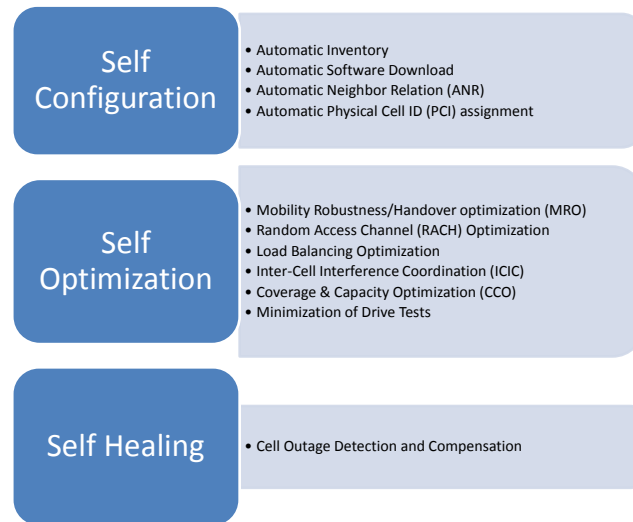


Figure 1.2 SON functions defined by 3GPP.

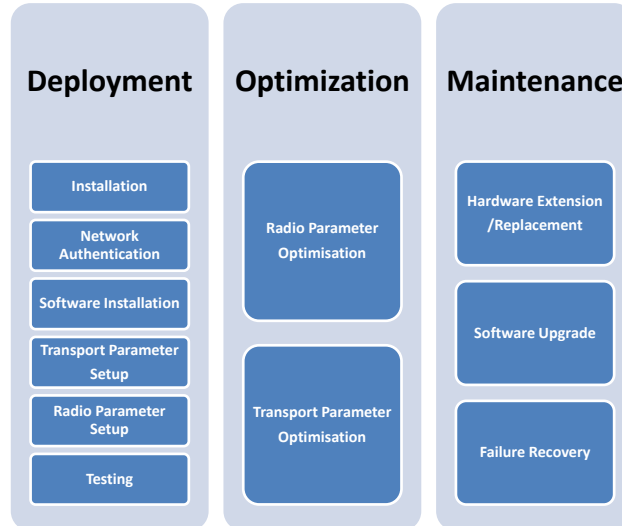


Figure 1.3 SON related use cases defined by NGMN.

frequency reuse scheme plays a major role in interference control and frequency selection is important which needs to be decided at the self-configuration phase. An SBS should identify its allowable frequency band

before entering into the operational phase. One option is to use Universal Frequency Reuse (UFR) with cross-tier interference constraints and another option is to split the existing bandwidth for each tier. In the latter case there will be no cross-tier interference; however, the spectral efficiency can be less than that in UFR [26]. A detailed description of frequency selection will be given in Section 1.3. Enabling self-configuration process lessens or avoids involvement of manual expertise during the installation phase.

- *Self-optimization*: The main task of self-optimization is to automatically adjust certain parameter settings to adapt with the network dynamics for the optimal performance. In order to perform self-optimization, the network nodes need to measure certain network parameters (e.g., number of users, traffic patterns, and traffic load) and collect the information about the network conditions (e.g., channel gains). Then these information can be used to optimize the network performance.

In recent literature, many approaches have been proposed to realize self-optimization in small cell networks. Some of the prominent game theory based approaches will be discussed in the latter parts of this chapter.

Resource allocation-based inter-cell interference coordination is one of the mostly targeted issues in self-organizing networks. Resource allocation settings (e.g., channel allocation and power allocation) and scheduling are significant in inter-cell interference coordination. Different criteria can be used for performance optimization depending on the objectives. Several commonly used optimization objectives are as follows: throughput/data rate/ SINR maximization, coverage maximization, load balancing, power minimization.

Note that multiple objectives can also be merged together by defining a proper payoff function [27, 28].

These objective can be further categorized as system centric objectives and user centric objectives. System centric objectives focus on optimizing the total network performance rather than individual performance. This type of approaches generally rely on a considerable amount of information exchange among network nodes and a centralized controller is usually required. An example is the maximization of the total network throughput with the constraint of a maximum transmit power. In comparison, user centric objectives focus on individual performance at each node (e.g., maximizing the individual rate) rather than the overall performance. This type of objectives are common for self-organizing network since they are more likely to rely on local interactions among the nodes.

- *Self-healing*: Self-healing enables the network to have the ability to detect, diagnosis, compensate, and recovery from failures and abnormal status. Accordingly, the self-healing process is mainly composed of three functions [29]: fault detection, fault diagnosis, and fault recovery.

Traditional healing approaches may not be feasible due to the existence

of a large number of heterogeneous base stations and their random nature of deployment. Instead, methods for self-healing would be required. Firstly, the problems should be detected from performance measurement (e.g., abrupt performance degradation) or event driven report. In this case, periodic monitoring should be performed. Then a diagnosis process can be performed to determine the cause for the failure (e.g., software or hardware) according to which the corresponding compensation and recovery schemes can be performed. In the case of software faults, the base station may try several actions such as reloading of a backup of software, activation of a fallback software load, and downloading a software unit and reconfiguration. In the case of hardware faults, the base station may use redundant resources [30].

1.2.2 Characteristics of Self-Organizing Algorithms

Comprehending the significant and necessary features of a self-organizing algorithm is important and essential for the design of self-organizing small cell networks. In this section, we summarize the important characteristics of self-organizing algorithms as follows.

1. *Stability*: The stability in the context of self-organizing networks is defined as [3]: “An algorithm or adaptation mechanism that is able to consistently traverse a finite number of states within an acceptable finite time.”. That is, a self-organizing algorithm should be able to converge within acceptable iterations. Note that for game theory-based algorithms, there could exist more than one equilibrium points. Certain conditions and initial points may be required for the algorithm to converge to the desired equilibrium point. Also, the delay in information exchange may result in delayed convergence or oscillations around an equilibrium point [28].
2. *Robustness*: Robustness is the ability of an algorithm to reach back to a stable state within a bounded duration of time in case of an unexpected change in the system or environment which makes the system deviate from a stable state. Small cell base stations may be more vulnerable to failures than cautiously planned macro base stations. Self-organizing algorithms should be capable of bringing the system back to an equilibrium state. In this regard, robustness can also be viewed as a part of self-healing functionality.
3. *Scalability*: The complexity of self-organizing algorithms should not increase in an unbounded manner with the increase of network size. The scalability poses certain complexity requirements on the algorithms. Specifically, less complex algorithms which occupy less computation resource (e.g., CPU and memory) could make the network more scalable. Also, the amount of information exchange should not increase unbounded with increase in the number of network nodes. Learning through local interactions can prevent the system from extensive information exchange.

4. *Agility*: The network should respond to the environmental changes within a reasonable duration of time. Agility depends on backhaul constraints of the nodes as the information has to be exchanged prior to the decision making. It also depends on the computational power of the network nodes. Global information exchange can make the system respond too sluggishly. While responding to temporary changes may also result in oscillations between states. Therefore, perfect agility is considered as one of the most difficult conditions to be fulfilled for a self-organizing network.

1.3 Issues and Challenges in Self-Organizing Small Cell Networks

Enabling self-organization for small cell networks poses a number of issues and challenges which should be fully understood. In this section, we identify the main design issues and challenges for self-organizing small cell networks.

1. *Interference mitigation*: Due to the scarcity of the available bandwidth allocated for wireless networks, small cells have to share the same transmission bandwidth with the existing macro network which results in both cross-tier and co-tier interferences. With the increasing density of the small cell networks, interference mitigation, which is essential for self-organizing small cell networks, becomes more challenging.
2. *Resource management*: Guaranteeing the efficient coexistence of a large number of small cells with traditional macro cells from the perspective of resource allocation is a fundamental issue [31]. Self-organizing algorithms should be capable of performing resource allocation to achieve optimal performance. Note that resource allocation objectives may vary depending on the requirement. For example, cross-tier and intra-tier interferences can be mitigated through proper power and sub-channel allocation. In addition to that most of other use cases categorized under self-optimization phase (e.g., load balancing, coverage and capacity optimization, and handover optimization) can also be achieved by using appropriate resource allocation. It is desirable for SON entities to take independent decisions on resource allocation without any centralized control. Therefore, developing distributed or semi-distributed resource management techniques for self-organizing small cell networks is one of the key issues.
3. *Access control*: A mobile user in a multi-tier network is capable of connecting to either macro base station or a small cell base station provided that the user is in the coverage area of both cells. This decision can be taken by the users based on the receive power of the pilot signal. On the other hand, the base stations can also decide how many users and which users should be accepted to be served in order to meet their own requirements (e.g., maximizing the total capacity and load balancing). Decisions on access control are expected to be taken distributively in SONs.

4. *Learning and reasoning*: Devising suitable learning techniques for self-organizing small cell networks is one of the major challenges. Self-organizing entities are expected to collect network information during the learning process. The learning technique should be strong enough to develop a sufficient knowledge base that can be used by the self-organizing entities to exploit the available resources efficiently. This also involves issues such as deciding the information collection rate and achieving a balance between the exploration and exploitation trade-off. Reasoning refers to the decision process to achieve optimal or desired network performance according to the knowledge base obtained during the learning process.
5. *Computation cost*: The SBSs may not have high processing power as that of traditional macro base stations. In this case, complex algorithms which require high computation power may not be suitable for small cells. Designing low-complexity self-organizing algorithms for small cells is a major challenge.
6. *Imperfect information*: With certain self-organizing algorithms, the SBSs are expected to exchange information with nearby nodes (i.e., local interactions). However, this information can be distorted due to the noisy backhaul and can be delayed due to the time taken in processing and transmission. In addition to local interactions, many algorithms also rely on Channel State Information (CSI). While CSI can also be distorted or temporally unavailable due to the fading experienced by feedback channels. Also, if the status of each channel is estimated by spectrum sensing, the sensing result can be inaccurate.

These imperfect information could affect the self-organizing algorithms from two aspects. First, the performance of the algorithms could degrade due to the use of inaccurate information. Secondly, the stability of the algorithms may not be guaranteed due to the delayed information [28, 32]. Therefore, dealing with imperfect information also poses a significant challenge to the design of self-organizing algorithms for small cell networks. The consideration of imperfect information and quantification of its effect can be found in several works such as [32, 33, 34].

7. *Limited backhaul*: Unlike macro base stations which have a separate backhaul, SBSs such as femto base stations connect to the core network via a IP-based backhaul such as DSL. The same backhaul link may also be used for inter-cell coordination and periodic information exchange required by self-organizing algorithms. The limited capacity of backhaul and the possible latency and errors introduced are considerable issues in the context of self-organizing small cell networks. Also note that the backhaul can be hybrid (e.g., coexistence of both wired and wireless backhaul) with different constraints [35]. These backhaul limitations and constraints should be taken into account when developing self-organizing algorithms for small cell networks. Security is also a significant issue since the backhaul may not be owned by the same operator.

1.4 Game Theory for Self-Organizing Small Cell Networks

1.4.1 Fundamentals of Game Theory

Game theory provides a rich set of mathematical tools for modeling and analyzing interactive decision making problems in which the interests of agents (i.e., players) may conflict with each other. It is a well developed area in applied mathematics and has been used primarily in economics to model competitions in markets.

In recent years, game theory has also been widely adopted to solve many problems in the area of wireless communications [36, 37]. A number of works have explored the applications of game theory for the analysis and optimization of various issues in wireless systems, in most cases to solve resource allocation problems in a competitive environment.

A non-technical definition of a game is given as follow. A game is a process in which the agents select certain strategies from their own strategy sets and obtain payoffs according to the strategies of all agents. The choice of a strategy can be made both simultaneously and non-simultaneously. In addition, an agent may make decisions multiple times according to the game rule. A game consists of a set of players, a set of strategies available to those players, and a specification of payoffs for each combination of strategies.

1. *Set of players* \mathcal{N} : The set of decision makers involved in the game. The players are assumed to be rational or bounded rational depending on the type of the game.
2. *Set of strategies* $(\mathcal{S}_i)_{i \in \mathcal{N}}$: Strategies are the options that a player can select depending on the state of the game. Here \mathcal{S}_i denotes the set of strategies of player $i \in \mathcal{N}$. A player's strategy could contain a single action, multiple actions, or probability distribution over multiple actions. As common in game theory, \mathcal{S}_{-i} denotes the strategies of all players other than i . The state of a game depends on the strategies taken by all the players (i.e., $[\mathbf{s}_i, \mathbf{s}_{-i}]$). Note that different players could have different strategy sets.
3. *Payoff* π_i : The payoff represents the preference of each player under the current strategy profile. The payoff could be modeled as a cost function $c_i(\mathbf{s}_i, \mathbf{s}_{-i})$, a utility function $u_i(\mathbf{s}_i, \mathbf{s}_{-i})$, or a combination of both (e.g., in the form of equation (1.1)), where the cost function represents the cost of performing certain strategies (e.g., transmit power) which needs to be minimized, the utility function represents the gain (e.g., profit of service providers) which needs to be maximized.

$$\pi_i(\mathbf{s}_i, \mathbf{s}_{-i}) = u(\mathbf{s}_i, \mathbf{s}_{-i}) - c(\mathbf{s}_i, \mathbf{s}_{-i}). \quad (1.1)$$

It is straightforward to see that a player's payoff depends not only on her own strategy but also on the strategies of all other players.

1.4.2 Motivations of using Game Theory for Self-Organizing Networks

The motivations of using game theory for self-organizing small cell networks are summarized below.

- The heterogeneous network nodes in small cell networks can be deployed by different operators/users. The performance of one network could be easily affected by the behavior of other networks. In this case, modeling of interactive behavior would be required. Different from optimization models in which the mutual impact among different entities during the decision making process cannot be accurately taken into account, game theory models provide a mathematical framework to analyze the competitive or cooperative interactions among the players in a multi-player system.
- Different network nodes could have different QoS requirements and can be self-interested. Each node takes individual decisions (e.g., on resource allocation and scheduling) to meet her own requirements rather than optimizing the system-wide performance. In this case, these nodes may have conflicting interests. Such self-interested behavior can be easily modeled by using game theory (e.g., by formulating a non-cooperative game). The “self-interest” of the nodes can be modeled in terms of performance metrics such as capacity, delay, throughput, interference and signal-to-interference-plus-noise ratio (SINR), in the in the payoff.
- The basic keystones of a self-organizing network as defined in Section 1.2 are *ability to learn from environment*, *autonomous adaptivity* and *emergent behavior capability*. In the context of game theory, the players could adapt their decisions to obtain a better payoff (i.e., learning and adaptation). Also, after several adaptation iterations, the game could reach the equilibrium (emergent behavior). The above mentioned properties of a self-organizing network can be attained by devising self-organizing algorithms based on game theory.
- Centralized algorithms could be highly inefficient for a dense heterogeneous wireless networks due to the complexity of the algorithms and the amount of information exchange. Accordingly, distributed control is a desirable feature for self-organizing small cell networks as explained in Section 1.2. Game theory provides a natural tool to develop distributed self-organizing algorithms as it allows local interactions and individual decision making. Local interactions will reduce the amount information exchange among the nodes and as a result the network becomes more scalable and more capable of operating with limited backhaul conditions.

1.4.3 Types of Games

Different game models (e.g., non-cooperative/cooperative, static/dynamic) have been used to address self-organizing problems in small cell networks the choice

of which depends on the characteristics of the network, applications, and also the objectives.

Different game theory models may differ considerably in structure from many aspects, e.g., number of players, number of strategies, and payoffs. The number of players may vary in different games. If a game has only one player, the game becomes an optimization problem. We call a two person game or multiple person game if the game has two or more players, respectively. In different games, the number of strategies for players can be either finite (e.g., in a rock-scissor-paper game) or infinite (e.g., in a pricing game). The analysis of a finite strategy game and an infinite strategy game are different. The summation of payoffs of all players may also differ in different models. In general, this summation can be zero, a non-zero constant number, or any arbitrary value. The game process is an important aspect in the game structure. The players in a game may take actions simultaneously, in a certain order, or in a repeated fashion, according to which the game can be referred to as a *static game*, a *dynamic game*, and a *repeated game*, respectively. In addition, the assumptions of players' rationality are different. Most of the game theory models assume perfect rationality of players, while some models consider that the players are with limited rationality (i.e., bounded rationality). According to the above analysis, game models can be divided into the following categories.

Non-cooperative vs. cooperative games

Non-cooperative games are the most popular games. In non-cooperative games, the players are commonly considered to be rational and self-interested who have fully or partially conflicting interests. Each player selects the strategy to optimize her own payoff function. For non-cooperative games, the most commonly used solution concept is Nash Equilibrium the definition of which is given as follows.

DEFINITION 1.1 *Nash Equilibrium*: Let $\mathbf{s}_i \in \mathcal{S}_i$ and $\mathbf{s}_{-i} \in \mathcal{S}_{-i}$. Then the NE strategy profile $(\mathbf{s}_i^*, \mathbf{s}_{-i}^*)$ is defined as,

$$\pi_i(\mathbf{s}_i^*, \mathbf{s}_{-i}^*) \geq \pi_i(\mathbf{s}_i, \mathbf{s}_{-i}^*) \quad (1.2)$$

for all $\mathbf{s}_i \in \mathcal{S}_i$ and for all $i \in \mathcal{N}$.

When the game reaches a Nash equilibrium, none of the players can improve her payoff by changing strategy unilaterally. There are also other solution concepts such as correlated equilibrium which can be considered as a generalized version of NE [38], evolutionary equilibrium and dominant-strategy equilibrium. We also discuss some of the other solution concepts that have been applied in the context of self-organizing small cell networks later in this chapter.

Recently, cooperation among network nodes for improving both individual and system wide performance has attracted much attention. The players can make agreements and cooperate. Cooperative game provides analytical tools to model and analyze the cooperative behavior of rational players who may form coalitions. In this case, the members of each coalition cooperate to maximize

the coalition payoff and the competition is among coalitions instead of among individual players.

Static vs. dynamic games

A static game is one in which a single decision (time irrelevant but may contain multiple actions) is made by each player, and each player has no knowledge of the decisions made by other players before making her own decision. Decisions are made simultaneously (or their order is irrelevant). A game is dynamic if the order in which the decisions are made is important or the strategy itself is time-dependent. For dynamic games, the dynamics can be abstracted from different aspects which lead to different types of dynamic games listed as follows:

(i) *Dynamic nature in games' play order*: The dynamic nature in games' play (decision) order leads to the development of multi-stage game (e.g., Stackelberg game). In this case, the decisions are made asynchronously and the games' play order is important. The players who move later can observe the decisions of the players who move first and then make the decisions accordingly. Note that if multiple players exist in one stage, the competition within this stage is usually formulated as a stage game.

(ii) *Dynamic nature in time dependency*: The dynamic nature in the time dependency leads to the development of differential game and evolutionary game. For differential game, the strategy of a player is time-dependent (i.e., function of time t). That is, the player seeks a best response strategy considering the entire time horizon. For evolutionary game, the players adapt their strategies according to the time-varying system state.

Games with special structures

Note that in general the existence and uniqueness of equilibrium as well as the convergence of best response dynamics to the equilibrium cannot be guaranteed. However, based on the characteristics of the game formulation, some special structures of games can be identified with which the games show remarkable properties in terms of the existence and convergence of pure strategy NE. Two of those special structure games which are useful in deriving self-organizing solutions are discussed below.

(i) *Supermodular games*: Supermodular games are characterized as the games with strategic complementarities. The 'increment' of strategy of one player will be unprofitable to other players. Therefore, the best response of other players would also be an increment of their strategies. The technical definition is given as follows.

DEFINITION 1.2 *Supermodular game*: Let $\mathcal{G} = (\mathcal{N}, (\mathcal{S}_i)_{i \in \mathcal{N}}, (\pi_i)_{i \in \mathcal{N}})$, where \mathcal{N} is the player set, $(\mathcal{S}_i)_{i \in \mathcal{N}}$ is the strategy set which is a subset of Euclidean space, and (π_i) is the payoff of i^{th} player. The game \mathcal{G} is said to be supermodular if following conditions are satisfied [39]:

1. $(\mathcal{S}_i)_{i \in \mathcal{N}}$ is a compact subset of \mathbb{R} .

2. $(\pi_i)_{i \in \mathcal{N}}$ is continuous.
3. s_i and s_{-i} show increasing differences which is equivalent to the condition $\frac{\partial^2 \pi_i(s_i, s_{-i})}{\partial s_{ih} \partial s_{ik}} \geq 0$ for all $k \neq h$.

The following properties can be observed in supermodular games [40, 41].

1. Best responses are monotonically increasing.
2. Pure strategy NE exists.
3. NE can be attained using greedy best-response algorithms.
4. If the NE is unique, it is also globally stable.

(ii) Potential games: A game is categorized as a potential game if the motivation of all players to change their strategy can be expressed using a single global function (i.e., the potential function). In such games, obtaining the NE is equivalent to the maximization of the potential function.

DEFINITION 1.3 *Exact potential game*: Let $\mathcal{G} = (\mathcal{N}, (\mathcal{S}_i)_{i \in \mathcal{N}}, (\pi_i)_{i \in \mathcal{N}})$ be a non-zero sum non-cooperative game where \mathcal{N} is the player set, $(\mathcal{S}_i)_{i \in \mathcal{N}}$ strategy set, and (π_i) is the payoff of the i^{th} player. The game \mathcal{G} is an exact potential game if there exists an exact potential function $\Phi : \mathcal{S} \rightarrow \mathbb{R}$ for all $i \in \mathcal{N}$ such that

$$\Phi(s'_i, s_{-i}) - \Phi(s''_i, s_{-i}) = \pi(s'_i, s_{-i}) - \pi(s''_i, s_{-i}), \quad (1.3)$$

where $s_{-i} \in \mathcal{S}_{-i}$ and $s'_i, s''_i \in \mathcal{S}_i$.

In other words, the change in individual payoff gained by any player by unilaterally deviating to another strategy is same as the difference in the corresponding values of the potential function. In ordinal potential games the signs of the differences are similar.

DEFINITION 1.4 *Ordinal potential game*: Let $\mathcal{G} = (\mathcal{N}, (\mathcal{S}_i)_{i \in \mathcal{N}}, (\pi_i)_{i \in \mathcal{N}})$, is a non-zero sum non-cooperative game where \mathcal{N} is the player set, $(\mathcal{S}_i)_{i \in \mathcal{N}}$ strategy set and (π_i) is the payoff of i^{th} player. The game \mathcal{G} is an exact potential game if there exists an exact potential function $\Phi : \mathcal{S} \rightarrow \mathbb{R}$ for all $i \in \mathcal{N}$ such that

$$\text{sgn}[\Phi(s'_i, s_{-i}) - \Phi(s''_i, s_{-i})] = \text{sgn}[\pi(s'_i, s_{-i}) - \pi(s''_i, s_{-i})], \quad (1.4)$$

where $s_{-i} \in \mathcal{S}_{-i}$, $s'_i, s''_i \in \mathcal{S}_i$, and sgn denotes the sign function.

Note that the above definitions are only valid for static potential games. Potential games have following remarkable properties [36, 42].

1. Every finite exact or ordinal potential game has at least one pure strategy NE.
2. Both best response dynamics and better response dynamics converges to the pure NE.
3. The NE is unique if

- \mathcal{S} is compact and convex.
- Φ is continuously differentiable on the interior of \mathcal{S} .
- Φ is strictly concave on \mathcal{S} .

1.4.4 Price of Anarchy and Price of Stability

Game theoretic approaches may not guarantee the optimal performance, i.e., an equilibrium solution of the game may not be the optimal solution for the problem. This inefficiency of the game theoretic solutions may occur due to the selfish behavior of the players. To measure the inefficiency of equilibrium solutions of a game, two popular concepts, i.e., the price of anarchy (PoA) and the price of stability (PoS) [43] can be defined as follows.

Price of Anarchy

Price of anarchy is defined as the ratio between the payoffs at the worst equilibrium (i.e., the equilibrium point which gives the least payoff) and the optimal centralized solution to the problem. PoA can vary for different payoff functions.

Price of Stability

Price of anarchy can be significantly small for the games with multiple equilibria even if only one equilibrium point is inefficient. Hence, price of stability is defined as the ratio between payoff received at the best equilibrium and the optimal (best possible) payoff.

Note that PoA and PoS are both equal for the games with unique equilibrium.

1.4.5 Design of Payoff Functions

Game theory was initially proposed and developed for economics and social sciences. Therefore, properly fitting those game models in the context of communication engineering is challenging. Specifically, defining the payoff functions based on the network performance metrics (e.g., achievable data rate, delay, and transmit power), modeling the network dynamics (e.g., randomness of the wireless channel, randomness of the user locations and base station deployment, and mobility of the users), meeting the requirements defined by the standards and realizing of the SON characteristics have to be considered within the scope of the game. Among all these, defining a proper payoff function is one of the key challenges. The payoff function quantifies the perceived preference or the satisfaction level of a player. In the context of self-organizing small cells, the user satisfaction level may depend on one or multiple performance metrics given as follows:

- Individual performance (e.g., rate, SINR, and delay)
- Global network performance
- Interference level caused to other network nodes

- Power/energy consumption
- User fairness.

As self-organizing small cell technologies are still in its infancy, there is no well-defined framework for designing the payoff functions. To this end, we will introduce some general approaches and guidelines on how payoff functions can be designed for various applications and objectives in the context of self-organizing small cell networks.

A payoff function $\pi(x)$ is expected to satisfy the following criteria.

1. The non-stationary property: $\frac{d\pi(x)}{dx} > 0$, which states that the payoff increases with the preference or satisfaction.
2. The risk aversion property : $\frac{d^2\pi(x)}{dx^2} < 0$, which states that the payoff function is concave. In other words, the marginal payoff of satisfaction decreases with increasing level of satisfaction.

Depending on the objective, behavior, and rationality of the network nodes, different payoff functions are defined in the wireless communications literature. The payoff/utility functions which can be applied in the context of small cell networks are discussed below.

Payoff functions for power consumption

Power/energy conservation is crucial in small cell networks as they might be operated in energy-limited environment (e.g power supplied by a battery). [44] defines a simple energy aware payoff function as follows:

$$\pi_i(e) = \frac{E_{tot}}{e_i}, \quad (1.5)$$

where E_{tot} is the total energy available for each player and e_i is the energy required by player i for transmission. Players would try to achieve a higher payoff by reducing the transmission power.

Payoff functions for individual performance

Instead of direct power minimization as that in equation (1.5), it is more appropriate for self-organizing algorithms to perform power control in such a way that the desired performance can be satisfied. The following logarithmic payoff function with individually perceived SINR as the input parameter can capture the self-interest of network nodes and is used for power control in [45, 46]:

$$\pi_i(s_i, \mathbf{s}_{-i}) = \log(\gamma_i(s_i, \mathbf{s}_{-i})), \quad (1.6)$$

where γ_i is the SINR of the i^{th} player. Such a logarithmic payoff function and its extensions are most popular payoff functions used in the context of resource allocation due to its simplicity and mathematical tractability [47]. For example, such form of payoff can be used for subcarrier allocation (in OFDMA networks) and joint power-subcarrier allocation as well.

Another widely used payoff function is the Shannon capacity or the maximum

achievable rate which can be considered as an extended version of logarithmic function of SINR as shown below:

$$\pi_i(s_i, \mathbf{s}_{-i}) = \ln(1 + \gamma_i(s_i, \mathbf{s}_{-i})). \quad (1.7)$$

Fairness utility function

One of the desired objectives of resource allocation is to provide fairness among users instead of obtaining the optimum performance. The most widely used payoff function which guarantees fairness is given below:

$$u(x) = \begin{cases} \frac{x^a}{a}, & \text{if } a < 0, \\ \log x, & \text{if } a = 0, \end{cases} \quad (1.8)$$

where $a \leq 0$. By twice differentiation of (1.8) with respect to x we obtain

$$\frac{du(x)}{dx} = \begin{cases} x^{a-1}, & \text{if } a \neq 0, \\ \frac{1}{x}, & \text{if } a = 0, \end{cases} \quad (1.9)$$

and

$$\frac{d^2u(x)}{dx^2} = \begin{cases} (a-1)x^{a-2}, & \text{if } a \neq 0, \\ -\frac{1}{x^2}, & \text{if } a = 0. \end{cases} \quad (1.10)$$

It can be observed that the function given in equation (1.8) has both non-stationary and risk aversion properties for all $x > 0$ since $\frac{du(x)}{dx} > 0$ and $\frac{d^2u(x)}{dx^2} < 0$.

System payoff functions

In self-organizing enabled small cell networks, a group of densely deployed small cells could form a cluster and cooperate with each other to enhance the performance of the cluster [48]. In addition to that, cooperative games can also be formulated to design self-organizing algorithms for small cells. Accordingly, cooperative payoff functions, which reflect the overall network/cluster performance, are required.

The simplest and most intuitive cooperative payoff function would be the sum capacity/rate of the cluster/network as shown below:

$$\pi_i(\mathbf{s}) = \sum_{j \in \mathcal{N}_i} C_j(\mathbf{s}), \quad (1.11)$$

where \mathcal{N} is the set of players in the i^{th} cluster who cooperates with each other and C_j is the capacity of the j^{th} player.

Multi-dimensional payoff function

The payoff function can be designed considering multiple performance metrics. In such cases, these multiple metrics could appear in the payoff function (most case in a product form). One typical example is given as follows:

$$\pi_i = \pi_i^{\text{rate}} \pi_i^{\text{delay}}. \quad (1.12)$$

Payoff function with cost

For a strategy adopted by a player, there could be a cost associated with it (e.g., cost of using bandwidth, power consumption) or it may affect the performance of other players (e.g., cause interference). This issue can be modeled by introducing certain cost functions into the payoff function. In particular, the payoff function (some may refer to this as *net utility*) can be defined to reflect both the satisfaction of the player (modeled by utility function) and the cost (e.g., price per unit resource) as follows:

$$\pi_i(s_i, \mathbf{s}_{-i}) = u_i(s_i, \mathbf{s}_{-i}) - mx, \quad (1.13)$$

where $u_i(s_i, \mathbf{s}_{-i})$ is the utility based on the user satisfaction and m is the price paid for each resource x .

[49] uses a net utility function with logarithmic payoff as given below:

$$\pi_i(s_i, \mathbf{s}_{-i}) = a_i \log(1 + \gamma_i(s_i, \mathbf{s}_{-i})) - b_i m \gamma_i(s_i, \mathbf{s}_{-i}), \quad (1.14)$$

where γ_i is the SINR of the i^{th} user, a_i and b_i are weighting parameters and m is the cost for the received SINR. The gain of maximizing γ_i could be neutralized by the cost associated with the received SINR.

The following form of payoff function (equation 1.15) is used in [28] to limit the interference caused to the macro users by the downlink transmission of small cells:

$$\pi_i(s_i, \mathbf{s}_{-i}) = w_1 (\pi(\gamma(s_i, \mathbf{s}_{-i}))) - w_2 (I_m - T), \quad (1.15)$$

where w_1 and w_2 are biasing factors which can be determined based on which network (i.e., macro or small cell network) should be given priority in resource allocation. I_m is the interference caused to the nearest macro user and T is the macro user interference threshold. When the interference caused to the nearest macro user (I_m) exceeds a certain threshold, small cell base stations are demotivated to allocate resources to its user even if it increases the individual payoff. At the same time, such payoff function encourages the SBSs to use resources (i.e., transmit power and OFDMA subcarrier) as long as it does not exceed the interference threshold of the macro users.

Guaranteeing the existence of equilibrium is one of the essential features of any game formulation. It is straightforward that the existence of equilibrium, convergence, and stability of the equilibrium is highly related to the payoff function and the structure of the game. Therefore, special payoff function can also be designed to fit the game model into special structures (e.g., super-modular, potential). Polynomial time computability is another important feature of a payoff function. Besides, when it comes to self-organizing small cell networks, the ability to compute with local information or with reduced information exchange is also highly desirable.

1.5 Game Theory-Based Resource Management for Self-Organizing Small Cells

Resource management aims for efficient usage of scarce resources (e.g., power and spectrum) as well as for interference management when it comes to the underlying small cell networks. Besides, some other issues such as load balancing and coverage optimization can also be eventually modeled as resource allocation problems. In general, the resource allocation in orthogonal frequency-division multiple access (OFDMA)-based small cell networks can be categorized into three classes: subcarrier allocation, power allocation, and joint subcarrier-power allocation.

Game theory-based resource management is one of the mostly addressed issues in the context of self-organizing small cell networks. Different types of games are used to address the above issues depending on the objective and the network characteristics. In the following, formulations of the selected game models for devising self-organizing distributed resource management algorithms are discussed.

1.5.1 Non-cooperative Game-Based Decentralized Power Allocation

Power allocation problem of a self-organizing small cell network can be modeled by a non-cooperative game in which the players are the small cell nodes (e.g., SBSs or users). The strategy of a player is the allocation of transmit power. The strategy selection of a player will impact the payoff of other players. Specifically, the transmission power selection of a player creates a positive or negative impact on the payoff of other players due to the possible increase or decrease of interference. The payoff function of the players can be chosen appropriately according to the design objective.

Non-cooperative game-based downlink power control

In [39], a non-cooperative game is used to model the downlink transmission power allocation problem among the SBSs. A system with one central macro cell and several underlaid closed access small cells is considered. It is also assumed that the distance between an SBS and its associated users are almost the same, hence all users served by this SBS have equal rate and the rate of a user of small cell i (R_i) is given by

$$R_i = \frac{1}{N_i} \log \left(1 + \frac{h_{i,i} P_i}{I + \sum_{j \neq i} h_{i,j} P_j} \right), \quad (1.16)$$

where N_i is the number of users associated with base station i , $h_{i,j}$ is the average channel gain from small cell base station i to users in small cell j , I is the noise power plus the interference from the macro base station and P_i is the transmit power of small cell base station i . The small cell base stations are the players of the game each of which is self-interested and tries to increase its own capacity.

The strategy set for base station i is defined as $\mathcal{S}_i = [0, P_{max}]$, where P_{max} is the maximum allowable transmit power for a SBS. The payoff function is then defined considering three factors:

- Achievable average rate of the small cell base station
- Fairness of the system
- Transmit power.

The payoff increases with the increasing average rate, while it decreases with the increasing transmit power due to the increased interference caused to neighboring cells. Also, fairness should also be considered among the base stations. Therefore, the payoff function is defined as follows:

$$\pi(P_i) = N_i \log(R_i) - \beta P_i, \quad (1.17)$$

where β is a positive constant.

This game is shown to be a supermodular game and accordingly the NE can be achieved by using best response dynamics (see Sec.1.4.3). Based on best response dynamics, the following power control algorithm is derived. Small cell base stations update their transmit power periodically to best response to the current strategy profile of other base stations. Eventually, the algorithm converges to the NE.

Algorithm 1 Non-cooperative game-based downlink power allocation algorithm

- 1: Initialize
 - 2: **repeat**
 - 3: Measure noise and interference from other SBSs
 - 4: Calculate the payoff by substituting in equation (1.17)
 - 5: Find P_i which maximizes equation (1.17)
 - 6: Update P_i
 - 7: Wait until the next update time
 - 8: **until** SBS turns off
-

The performance of the above algorithm is evaluated numerically in [39] which proves the capability of the algorithm to be implemented in a real environment while providing fairness to the small cell users. However, the algorithm shows slightly degraded performance than the centralized system which delivers optimal performance. This is due to the selfish decentralized behavior of the users.

Non-cooperative game-based uplink power control

Uplink power allocation in small cell networks can also be modeled as a non-cooperative game [50, 51]. Specifically, a non-cooperative game based distributed uplink power control algorithm is proposed in [51]. The power control is performed distributively based on SINR adaptation while mitigating the interference caused to the macro base station. A single macro cell and a set of N underlaid

small cells are considered. Each base station serves only one user at a time with a guaranteed SINR requirement.

In order to protect the macro base station from interference due to the uplink transmission of the small cell users, the macro user is also considered as a player in the game. In this case, the player set consists of the macro user and the small cell users denoted by, $\mathcal{N} = \{0, 1, \dots, N\}$, where index 0 denotes the macro user and the indices $1, 2, \dots, N$ denote small cell users. The strategy of each player i is its transmit power denoted by p_i . The payoff function for macro user is given by

$$\pi_0(p_0, \mathbf{p}_{-0}) = -(\gamma_0 - \Gamma_0)^2, \quad (1.18)$$

where Γ_0 is the target SINR and γ_0 is the received SINR of the macro user. The received SINR of any user is given by

$$\gamma_i = \frac{p_i h_{i,i}}{\sigma^2 + \sum_{j \neq i} p_j h_{i,j}}, \quad (1.19)$$

where σ^2 is the noise power and $h_{i,j}$ is the channel gain between users i and j .

Each small cell user also tries to maximize her own individual SINR while meeting the minimum SINR requirement, Γ_i . The payoff of a small cell user is given by

$$\pi_i(p_i, \mathbf{p}_{-0}) = R(\gamma_i, \Gamma_i) + b_i \frac{C(p_i, \mathbf{p}_{-0})}{\sigma^2 + \sum_{j \neq i} p_j h_{i,j}}, \quad (1.20)$$

where b_i is a weighting factor. The reward function, $R(\gamma_i, \Gamma_i)$ and the penalty function $C(p_i, \mathbf{p}_{-0})$ are defined as follows:

$$R(\gamma_i, \Gamma_i) = 1 - \exp(-a_i(\gamma_i - \Gamma_i)), \quad (1.21)$$

where a_i is a constant and

$$C(p_i, \mathbf{p}_{-0}) = -p_i h_{0,i}. \quad (1.22)$$

The reward increases with γ_i until the threshold Γ_i is met. Once the received SINR exceeds the minimum requirement, the reward decreases with γ_i , which discourages the small cell users to increase their power. By equation (1.22), the small cell users are given a penalty with the increase of transmit power. The penalty is scaled by interference and noise ($\sigma^2 + \sum_{j \neq i} p_j h_{i,j}$) in equation (1.33) to ensure that small cells experiencing higher interference are less penalized. Note that the payoff function is a monotonically increasing concave function of γ_i for fixed p_i . Also, for fixed γ_i , the payoff is a monotonically decreasing concave function of p_i .

The existence of Nash equilibrium for the above uplink power control game can be proven by employing the following theorem from [52, 53, 54].

THEOREM 1.5 *A Nash equilibrium exists in game $\mathcal{G} = (\mathcal{N}, (\mathcal{S}_i)_{i \in \mathcal{N}}, \pi_i(\cdot)_{i \in \mathcal{N}})$ if, for all $i = 0, 1, \dots, N$,*

1. $(\mathcal{S}_i)_{i \in \mathcal{N}}$ is a nonempty, convex, and compact subset of some Euclidean space \mathbb{R}^{N+1} .
2. $\pi_i(\mathbf{s})$ is continuous in \mathbf{s} and quasi-concave in p_i .

The uplink transmit power at the NE (denoted by \mathbf{p}^*) is given by following two equations [51]:

$$p_0^* = \min \left(\frac{I_0(\mathbf{p}_{-0}^*)}{g_{0,0}} \Gamma_0, p_{max} \right), \text{ when } i = 0, \quad (1.23)$$

$$p_i^* = \min \left(\frac{I_i(\mathbf{p}_{-i}^*)}{g_{i,i}} \left[\Gamma_i + \frac{1}{a_i} \ln \left(\frac{a_i g_{i,i}}{b_i g_{0,i}} \right) \right]^+, p_{max} \right), \text{ when } i \neq 0, \quad (1.24)$$

where $[x]^+ = \max(x, 0)$ and $I_i(\mathbf{p}_{-i}^*) = \sigma^2 + \sum_{j \neq i} p_j^* h_{i,j}$.

In order to devise a distributed power control algorithm which converges to the NE, [51] uses the *standard interference* function defined in [55].

DEFINITION 1.6 *Standard interference function:* $f(\mathbf{p})$ is a standard interference function if the following conditions are satisfied for all $p \geq 0$:

1. Positivity, $f(\mathbf{p}) > 0$.
2. Monotonicity, if $p' > p$ then $f(p') > f(p)$.
3. Scalability, for all $\alpha > 1$, $\alpha f(p) > f(\alpha p)$.

Yates [55] showed that an iterative power control algorithm which calculates the power at next iteration $k+1$ according to the rule $\mathbf{p}^{k+1} = f(\mathbf{p}^k)$ converges to a unique fixed point if $f(\mathbf{p})$ is a standard interference function.

The received SINR is $\gamma_i = \frac{p_i}{I_i}$, according to which the equations (1.23) and (1.23) can be modified to form a distributed iterative power control algorithm. The individual uplink transmit power is updated as follows:

$$p_0^{k+1} = \min \left(\frac{p_0^k}{\gamma_0^k} \Gamma_0, p_{max} \right), \text{ when } i = 0, \quad (1.25)$$

$$p_i^{k+1} = \min \left(\frac{p_i^k}{\gamma_i^k} \left[\Gamma_i + \frac{1}{a_i} \ln \left(\frac{a_i g_{i,i}}{b_i g_{0,i}} \right) \right]^+, p_{max} \right), \text{ when } i \neq 0. \quad (1.26)$$

Both equations (1.25) and (1.26) are standard interference functions. Therefore, the power control algorithm converges to a unique fixed point which is the NE defined in (1.23) and (1.23).

More importantly the algorithm can be executed distributively with minimal network overhead, and therefore, would be suitable for resource allocation in a self-organizing small cell network.

1.5.2 Non-cooperative game-based sub-carrier allocation

In addition to power allocation, sub-carrier allocation (we consider OFDMA networks) is also an essential part for self-organizing resource allocation. Distributed subcarrier allocation problem can also be modeled as a non-cooperative game in a similar way as that for power control. In this case, the strategy set should represent the selection of available subcarriers for each node of the network. In the following, we give a descriptive example for uplink distributed subcarrier allocation in OFDMA-based small cell networks based on non-cooperative games.

[56] proposes a decentralized method for small cells in a two-tier network to individually select the most appropriate subset of resource blocks in order to mitigate both cross-tier and co-tier interferences. In the model, the macro network consists of 19 macrocell sites, each of which has three hexagonal sectors. Small cells are deployed inside the macrocell according to the 5×5 grid model specified in the 3GPP simulation scenario given for urban deployment in [57]. 25 apartments are arranged according to a 5×5 grid and each of these apartments would have a small cell (femto cells in this case) with a probability of p_d . U_M number of macro users are randomly and uniformly located in each sector and U_S number of small cell users are randomly and uniformly located in each apartment. It is also assumed that the total bandwidth W is divided into K resource blocks.

The resource block (RB) allocation is modeled as a non-cooperative game $\mathcal{G} = (\mathcal{N}, (\mathcal{S}_i)_{i \in \mathcal{N}}, \pi_i(\cdot)_{i \in \mathcal{N}})$, where \mathcal{N} is the set of small cell users. The strategy \mathcal{S} of each player i is the selection of a subset of RBs. Two payoff functions π_1 and π_2 are considered (given in equations (1.27) and (1.28)), where π_1 takes into account only the co-tier interference between SBSs while π_2 considers both co-tier and cross-tier interferences. Let each small cell user select H number of resource blocks for transmission. S and M are the total number of small cell users and macro users, respectively. The selected set of resource blocks by small cell user i is given by $\mathcal{R}_i = k_i^1, k_i^2, \dots, k_i^H$. For all $i, j \in \mathcal{N}$ and $x, y \in 1, 2, \dots, H$, $\delta_{k_i^x, k_j^y}$ is an interference indicator function, where $\delta_{k_i^x, k_j^y} = 1$ if RBs k_i^x and k_j^y are the same and $\delta_{k_i^x, k_j^y} = 0$, otherwise. The two payoff functions are given below.

$$\pi_i^1(s_i, s_{-i}) = \sum_{x=1}^H \sum_{y=1}^H \left(- \sum_{j=1, j \neq i}^S g_j^{b_i} p_j^{k_j^y} \delta_{k_i^x, k_j^y} - \sum_{j=1, j \neq i}^S g_j^{b_j} p_i^{k_i^y} \delta_{k_j^x, k_i^y} \right), \quad (1.27)$$

and

$$\begin{aligned} \pi_i^2(s_i, s_{-i}) = & \sum_{x=1}^H \left[\sum_{y=1}^H \left(- \sum_{j=1, j \neq i}^S g_j^{b_i} p_j^{k_j^y} \delta_{k_i^x, k_j^y} - \sum_{j=1, j \neq i}^S g_j^{b_j} p_i^{k_i^y} \delta_{k_j^x, k_i^y} \right) \right. \\ & \left. + \sum_{z=1}^L \left(- \sum_{m=1, m \neq i}^M g_m^{b_i} p_m^{k_m^z} \delta_{k_i^x, k_m^z} - \sum_{m=1, m \neq i}^M g_m^{b_m} p_i^{k_i^z} \delta_{k_j^x, k_m^z} \right) \right], \end{aligned} \quad (1.28)$$

where $g_j^{b_i}$ is the channel gain between the small cell j and the SBS b_i , $g_j^{b_m}$

denotes the channel gain between the small cell j and the MBS b_m , and $p_j^{k_j^y}$ is the transmit power of user j on its selected RB k_j^y .

The two terms in equation (1.27) represent the co-tier interference caused at the base station i by other base stations and the co-tier interference caused to other SBSs by SBS i , respectively. The additional two terms in equation (1.28) measure the cross-tier interference.

The existence and the uniqueness of the NE are proved by showing that the above game is an exact potential game. The corresponding potential functions and the proof can be found in [56]. In potential games, the best response dynamics always converges to a pure strategy NE. The best response strategy is given by

$$s_i^{t+1} = \arg \max_{s' \in \mathcal{S}} \pi_i(s', s_{-i}^t). \quad (1.29)$$

The small cell users are assumed to be able to sense the spectrum in order to select the best set of RBs as s_i^{t+1} .

Giupponi in [58] also formulates the downlink resource (joint subcarrier and power) allocation problem as a potential game. The payoff function is designed to model both co-tier and cross-tier interferences and a distributed resource allocation algorithm is proposed.

1.5.3 Complimentary TRi-Control Loop for Uplink Interference Mitigation

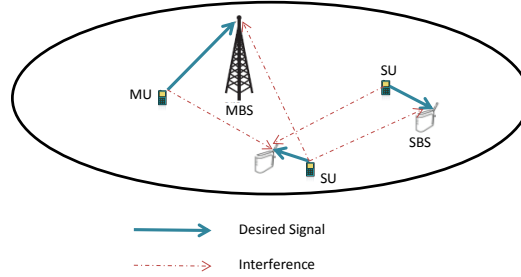


Figure 1.4 Uplink interference scenario in a two-tier network.

Small cell users may cause considerable amount of interference to the macro base station in uplink and vice versa (as shown in Fig. 1.4). [59] proposes a self-organizing uplink interference management architecture called Complimentary TRi-Control Loop (CTRL) which is composed of three control loops as shown in Fig. 1.5, which are explained below.

- *MTXPC - maximum transmit power control loop*: The maximum transmit power control loop determines the maximum possible transmit power for

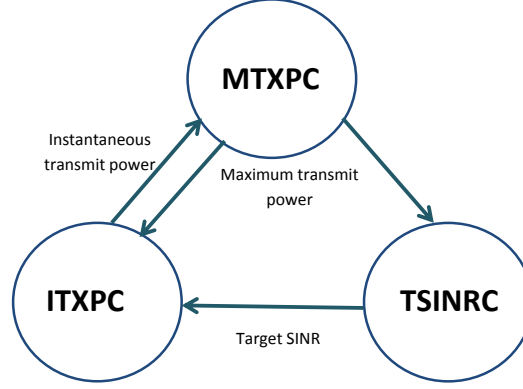


Figure 1.5 Complimentary TRi-control loop.

small cells in order to provide the uplink protection for the macro network. The decision is taken based on the feedback uplink load margin information of the macro base station.

- *TSINRC - target SINR control loop*: The duty of target SINR control loop is to determine the required uplink SINR for each SBS with minimum possible information exchange among them. A non-cooperative game is formulated and the decisions are taken considering the maximum transmit power decided at MTXPC loop.
- *ITXPC - instantaneous transmit power control loop*: The actual transmit power allocation is done at this loop. The transmit power is allocated to achieve the target SINR calculated at TSINRC loop with the constraint on maximum transmit power.

The maximum transmit power control loop is modeled as a Q estimation problem in adaptive control theory. MTXPC loop calculates the maximum transmit power for each SBS user in a self-organizing manner. The complete model can be found in [59].

The target SINR control loop is formulated as a non-cooperative game in which NE is obtained as the solution. The players are the set of small cell base stations. The strategy set (\mathcal{S}) is composed of the set of possible transmit powers on each resource block. The transmit power vector of each user i is given by $\mathbf{p}_i = (p_{i,1}, p_{i,2}, \dots, p_{i,K})$, where $p_{i,k}$ is the transmit power of user i on subcarrier k . $p_{i,k}$ must be less than the maximum transmit power $P_{i,k}$ obtained at the MTXPC loop. The vector \mathbf{p} is composed of the transmit powers of all the users in all SBSs. Let \mathcal{K} and \mathcal{N}_f denote the set of subcarriers and the set of users connected to SBS i , respectively. $b_{i,k}$ is the normalized time period that the user

i transmits on resource block k . The payoff function of each player f is given by

$$\pi_f(\mathbf{p}, \mathbf{b}) = \sum_{i \in \mathcal{N}_f} \sum_{k \in \mathcal{K}} b_{i,k} W \log_2 \left(1 + \frac{\gamma_{i,k}}{c} \right) - \sum_{i \in \mathcal{N}_f} \sum_{k \in \mathcal{K}} b_{i,k} \mu_{i,k} p_{i,k}, \quad (1.30)$$

where W is the size of a resource block, $\omega = -\ln(5BER)/1.6$ is a constant to achieve a given bit error rate (BER) $\mu_{i,k}$ is the price paid for the interference caused and $\gamma_{i,k}$ is the target SINR of user i on subcarrier k .

There exists a NE (given in equation 1.31) for the above game if $\mu_{i,k}$ is large enough.

$$\gamma_{i,k}^* = \max \left(\left[\frac{Wh_i^k}{(\ln 2)I_{i,k}(\mathbf{p}_{-i})\mu_i^k} - \omega \right]^+, \frac{h_{i,k}P_{i,j}}{I_{i,k}(\mathbf{p}_{-i})} \right), \quad (1.31)$$

where $I_{i,k}$ is the interference (including macro cell interference) plus the thermal noise at user i on resource block k and h_i^k is the channel gain from user i to its own base station on subcarrier k . The proof of the existence of an NE is based on the fact that the payoff function is continuous and quasi-concave and the strategy set is a non-empty, compact, and convex subset in the Euclidean space. The complete proof can be read from [59].

Based on the interactions among the above proposed three control loops, the spatial reuse of spectrum within small cells is enabled without degrading the performance of the macro tier. The operation of CTRL does not require any changes in the resource management of the macro network and converges distributively to a stable solution.

1.5.4 An Evolutionary Game Approach for Self-Organization with Reduced Information Exchange

As we have seen in the previous examples, traditional game theory (e.g., Nash equilibrium problem) relies on the rational decisions of the players. The players are expected to choose their strategies rationally as the best responses to the strategies of other players. The rationality implies complete information and strong computation capability of each player to calculate the best response to other players strategies. This assumption may be too strong for densely deployed nodes in a self organizing small cell network. As a solution, the distributed resource allocation problem can be formulated as an evolutionary game [28, 60, 61].

Evolutionary game theory (EGT) was originally developed to analyze the evolution of populations of biological species with bounded-rationality. Instead of selecting the strategy which gives the best response to the other users' strategies, in EGT, each player selects a strategy by replication and can adapt its selection for a better payoff (i.e., evolution). Accordingly, EGT focuses on the dynamics of the strategy adaptation in the population. A population is the set of players involved in the game. The behavior of the population can be described by the number of its members choosing each pure strategy. The success of a strategy is

reflected by the proportion of members in the population using it. In the following, we provide an example on formulating the downlink subcarrier selection and transmit power allocation of a small cell network as an evolutionary game [28].

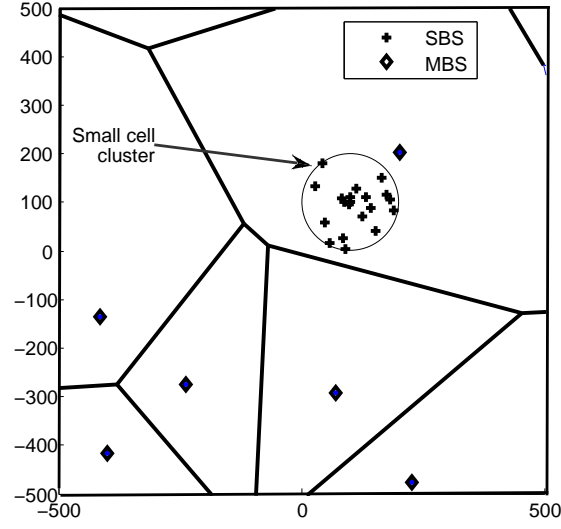


Figure 1.6 Small cell cluster underlaid with a macro network.

The downlink transmission of an OFDMA-based two-tier cellular network composed of macrocells and an underlying self-organizing small cell cluster is considered here (as shown in figure 1.6). The spatial distribution of the small cell base stations and macro base stations follow two independent point processes in \mathbb{R}^2 with densities λ_f and λ_m , respectively. Each macro user is attached to the nearest macro base station and each small cell user is located at a distance r_f from its serving base station. Each SBS serves only one user at a time and selects one subcarrier to serve that user. The macrocell and the small cells share the same set of orthogonal subcarriers denoted by $\mathcal{K} = \{1, 2, \dots, K\}$. They are also capable of selecting a transmit power level from a finite set of values which is denoted by $\mathcal{L} = \{1, 2, \dots, L\}$. Each SBS should select a suitable subcarrier-power combination which is referred to as the “transmission alignment” of that SBS. The set of transmission configurations (i.e., strategy set) is denoted by \mathcal{S} . For each subcarrier k , there is a maximum aggregate interference threshold that can be caused by the entire small cell cluster to the macro users which is denoted by $T^{(k)}$.

The small cell base stations form the player set of the game, denoted by \mathcal{N} . The strategies \mathcal{S} available for each player is the set of transmission alignments. In the context of an evolutionary game, the set of players also constitutes the population. Denote by n_s the number of SBSs selecting pure strategy $s \in \mathcal{S}$.

Then the frequency of strategy s used in the population is given by

$$x_s = \frac{n_s}{N}, \quad (1.32)$$

where the frequency x_s is also referred to as the population share of pure strategy s . The population shares of all strategies add to 1. The payoff is a function of the utility of an SBS when certain transmission alignment is used and the interference caused to the nearest macro user and is given by

$$\pi_s = \pi_l^{(k)} = w_1 \left(\mathcal{U}(\text{SINR}_l^{(k)}) \right) - w_2 \left(I_m^{(k)} - T^{(k)} \right), \quad (1.33)$$

where $\text{SINR}_l^{(k)}$ is the received SINR of a small cell user served by subcarrier k and power level l , w_1 and w_2 are biasing factors and $I_m^{(k)}$ is the aggregate interference created by the small cell cluster on subcarrier k at the nearest macro user.

Sepcifically, in [28], two utility functions are considered which are given as follows:

$$\mathcal{U}_1(\text{SINR}_l^{(k)}) = \mathbf{E} \left[\text{SINR}_l^{(k)} \right], \quad (1.34)$$

and

$$\mathcal{U}_2(\text{SINR}_l^{(k)}) = \mathbf{E} \left[\ln \left(1 + \text{SINR}_l^{(k)} \right) \right]. \quad (1.35)$$

Based on two utility functions, two payoff functions can be defined $(\pi_s^{(1)}, \pi_s^{(2)})$ and hence two games are formulated.

$$\mathcal{G}^1 = \left(\mathcal{N}, \mathcal{S}, \pi_s^{(1)} \right), \quad (1.36)$$

and

$$\mathcal{G}^2 = \left(\mathcal{N}, \mathcal{S}, \pi_s^{(2)} \right). \quad (1.37)$$

The Evolutionary Equilibrium (EE) is the solution concept for both \mathcal{G}^1 and \mathcal{G}^2 . In the context of the evolutionary game for transmission alignment selection, each SBS will adapt its strategy according to its received payoff. This is referred to as the evolution of the game during which the strategy adaptation of SBSs will change the population share, and therefore, the population state will evolve over time. The strategy adaptation process and the corresponding population state evolution can be modeled and analyzed by *replicator dynamics* [62] which is a set of ordinary differential equations defined as follows:

$$\dot{x}_s(t) = x_s(t) (\pi_s(t) - \bar{\pi}(t)), \quad (1.38)$$

for all $s \in \mathcal{S}$, with initial population state $\mathbf{x}(0) = \mathbf{x}_0 \in \mathbb{X}$, where \mathbb{X} is the state space which contains all possible population distributions. Here π_s is the payoff of each SBS choosing transmission alignment s and $\bar{\pi}$ is the average payoff of the entire population. The equilibrium point of the game can be obtained by solving the replicator dynamics. Evolutionary equilibrium is the point where the replicator dynamics is equal to zero. In other words, when the system is

at equilibrium, the fractions of the population choosing each strategy remain constant.

As the game (either \mathcal{G}^1 or \mathcal{G}^2) is repeated, each SBS observes its own payoff and compares it with the average payoff of the system. Then, if its payoff is less than the average, in the next period, the SBS randomly selects another strategy. The proposed distributed resource allocation algorithm is given in **Algorithm 2**.

Algorithm 2 Evolutionary game-based distributed resource allocation

- 1: **Initialize:** The SBSs choose a transmission alignment randomly and set $i = 1$.
 - 2: **repeat**
 - 3: **Exploitation** : Each SBS transmits on the selected transmission configuration and observes the received utility. The utility and the transmission alignment information are then sent to the central controller.
 - 4: **Learning**: A central controller calculates the average payoff of the population and the population state and broadcasts it to all SBSs.
 - 5: **Update**: Each SBS compares its own payoff with the average payoff of the population. If the payoff is less than the average, the SBS randomly selects another subcarrier for transmission.
 - 6: $i = i + 1$
 - 7: **until** $i \geq \text{Max}_i$ (maximum number of iterations that the algorithm can execute)
-

The stability of the of the equilibrium point can be analyzed by using a stochastic geometry approach. The spatial distribution of macro base stations and small cell base stations are approximated by Poisson point processes (PPP). Then the expected SINR for any population distribution can be derived in terms of the population shares (see [28] for the stochastic geometry based derivation). The expressions obtained for the above mentioned utility functions (when path-loss exponents equals 4) are as given below:

$$\mathbf{E} [\text{SINR}_l^{(k)}] = \frac{8p_l}{A^2 \left(\lambda_m \sqrt{p_m} + \lambda_f^{(k)} \mathbf{E} [\sqrt{p_f}] \right)^2}, \quad (1.39)$$

and

$$\mathbf{E} [r_l^{(k)}] = \int_{t=0}^{\infty} \exp \left(\frac{-A}{2\sqrt{p_l}} \left(\lambda_m \sqrt{p_m} + \lambda_f^{(k)} \mathbf{E} [\sqrt{p_f}] \right) \sqrt{e^t - 1} \right) dt, \quad (1.40)$$

where p_l is the transmit power of level l , p_m is the transmit power of MBSs, $p_f \in \{p_1, p_2, \dots, p_L\}$ denotes the transmit power of SBSs, $A = \pi^2 r_s^2$, and $\lambda_f^{(k)}$ is the density of SBSs transmitting on subcarrier k which is also assumed to be uniformly randomly distributed.

The probability mass function of the transmit power of any interferer (i.e.,

p_f in (1.39) and (1.40)) can be directly obtained from the proportions of the population selecting each strategy. For transmission alignment corresponding to subcarrier k and power level l , the PMF (which can be used to find $\mathbf{E}[\sqrt{p_f}]$) of the transmit power of a generic interferer is given as follows:

$$\Pr(p_f = p_j) = \begin{cases} \frac{n_j^{(k)}}{\sum_{t=1}^L n_t^{(k)} - 1}, & \text{if } j \neq l, \\ \frac{n_j^{(k)} - 1}{\sum_{t=1}^L n_t^{(k)} - 1}, & \text{if } j = l, \end{cases}$$

or equivalently,

$$\Pr(p_f = p_j) = \begin{cases} \frac{x_j^{(k)}}{\sum_{t=1}^L x_t^{(k)} - \frac{1}{N}}, & \text{if } j \neq l, \\ \frac{x_j^{(k)} - \frac{1}{N}}{\sum_{t=1}^L x_t^{(k)} - \frac{1}{N}}, & \text{if } j = l, \end{cases}$$

where $n_j^{(k)}$ is the number of players selecting subcarrier k and power level j and $x_j^{(k)} = \frac{n_j^{(k)}}{N}$. For a network with two orthogonal subcarriers and one transmit power level, the interior evolutionary equilibrium in game \mathcal{G}^1 can be shown to be asymptotically stable [28].

Simulations show that \mathcal{G}^2 converges faster than \mathcal{G}^1 . The impact of delay in information exchange also analyzed numerically which shows that the system converges to the equilibrium under small delays. However, when the delay is larger than a certain bifurcation point, the system will diverge. Also there is no guarantee that the system will converge to the same equilibrium point as the delay-free system. The key features of the above discussed evolutionary game based algorithm are as follows: simplicity, reduced information exchange than other non-cooperative game based algorithms, and fairness.

In [28], the performance of the above algorithm is compared with the optimal performance obtained by a centralized resource allocation which acts as a benchmark. A gap exists between the maximum payoff and the payoff obtained by the EGT-based algorithm. Also the gap increases with the number of base stations in both \mathcal{G}^1 and \mathcal{G}^2 .

Up to this end, we have studied some basic examples of using different game models in order to devise self-organizing algorithms for small cells. Other customized algorithms which satisfy various constraints can be built on top of these basic examples. For example, in [63] a self-organizing interference management paradigm is proposed while taking into account the constraints due to the presence of heterogeneous backhauls. The problem is formulated as a non-cooperative game and a fully distributed learning algorithm is devised based on reinforcement learning (RL) which converges to an equilibrium solution.

Cognitive Radio (CR) enabled small cells ([64, 65]) which can sense the spectrum are also proposed as a solution for interference mitigation. CR enabled SBSs can opportunistically allocate both licensed and unlicensed frequency bands to the users in order to avoid interference. [66] uses a correlated equilibrium-based

approach ([67]) to mitigate co-tier interference among cognitive femto access points for the downlink OFDMA LTE networks. Correlated equilibrium is preferred for a self-organizing network than NE as it allows devising decentralized and adaptive algorithms. In [66], the spectrum allocation competition among cognitive base stations is formulated as a non-cooperative game. The spectrum allocation is done using two payoff functions i.e., global payoff and local payoff. The global function provides fairness among players considering the total network performances while the local payoff function is based on individual performance measures.

A summary of the game models discussed in this chapter is given in Table 1.1.

Table 1.1 Game models for self organizing small cell networks

Objective	Game type	Player set	Strategy	Solution
Power allocation [39]	Non-cooperative, Supernodular	SBSs	Downlink transmit power	NE
Power allocation [50]	Non-cooperative	SUs	Uplink transmit power	NE
Power allocation [51]	Non-cooperative	MUs & SUs	Uplink transmit power	NE
Resource block allocation [56]	Non-cooperative, Potential	SUs	OFDMA RBs for uplink	NE
Joint power-subchannel allocation [58]	Non-cooperative, Potential	SBSs	A composite of downlink transmit power and subcarrier	NE
Power allocation [59]	Non-cooperative	SBSs	Uplink transmit power	NE
Joint power-subcarrier allocation [28, 61, 68]	Evolutionary	SBSs	Downlink power/subcarrier combination	EE
Forming a group for interference draining [69]	Cooperative coalition formation	SBSs	Coalitions	RC
Joint power-subcarrier allocation [70]	Non-cooperative	SBSs	A composite of downlink transmit power and subcarrier (mixed strategy)	LE
Forming a cluster for joint beamforming [34]	Cooperative coalition formation	SBSs	Coalitions	RC
Joint power-spectrum allocation [38]	Non-cooperative	SBSs	Downlink power level and frequency band	ϵ -CCE
Coverage optimization [48]	Non-cooperative	SBSs	Downlink transmit power	NE
Resource block (RB) allocation in cognitive SBSs [66]	Non-cooperative	SBSs	Downlink RBs	Correlated equilibrium

1.6 Learning Techniques for Self-Organizing Small Cell Networks

In the previous sections of this chapter, we have observed that in order to achieve the equilibrium, small cells have to be aware of the environment (e.g., by sensing or local interaction) and react accordingly by adjusting their resource allocation policies. Also, due to the dynamics of wireless environment, most of the system parameters need to be adjusted as well. Such adaptation for obtaining better performance and/or for reaching the equilibrium can be viewed as a learning process which is crucial for self-organization in small cells. Accordingly, different learning techniques can be applied. Specifically, distributed learning techniques such as RL and Q-Learning have recently gained significant attentions of the research community. Also, in the past few years there has been a growing interest in applying machine learning techniques for wireless networks.

In this section, we will introduce the basics of some commonly used learning techniques and provide examples on how those techniques can be applied for self-organizing small cell networks. Specially, we focus on the adaptations of different learning techniques to learn the equilibria in game-based self-organizing systems.

1.6.1 Reinforcement Learning

RL [71] can be simply explained as mapping situations into actions in such a way that the cumulative reward is maximized. A learning agent has to *explore* the environment and *exploit* what it has already explored to get a better payoff. The basic elements of RL are explained below.

- *Policy* (Φ): A policy is a mapping of the set of states given by \mathcal{S} into actions (\mathcal{A}). Generally the policies are stochastic.
- *Reward function* (u): Similar to the payoff/utility function in game theory, the reward function reflects an agent's preference of that state. Simply, the reward function maps the state-action pair into a numerical value. The agent's objective is to maximize its total reward in the long run.
- *Model* of the environment : The model reflects the behavior of the environment.

RL has been widely used in the field of cognitive radio networks [72]. Here we introduce an example (based on [70] and [73]) of applying RL to learn and reach equilibrium in a self-organizing small cell network.

Let us consider the downlink transmission of a macro base station and a set of N underlaid small cell base stations (denoted by \mathcal{N}) each of which transmits to one user at a time. $\mathcal{K} = \{1, \dots, K\}$ is the set of orthogonal sub-carriers shared by both tiers. At each time slot, the macro base station serves one macro user over each sub-carrier and also each SBS selects one subcarrier to transmit. Each SBS is capable of selecting a transmission power level from a finite set of power levels. The combination of the power level and the subcarrier is termed as a *transmission*

alignment. The problem is to select suitable transmission alignments for the downlink transmission of SBSs while protecting the macro users from interference.

The above problem can be modeled as a mixed-strategy non-cooperative game. SBSs form the player set \mathcal{N} and the available transmission alignments form the action set \mathcal{A}_i . The mixed strategy (\mathbf{s}_i) vector of user i is given by

$$\mathbf{s}_i = (\alpha_{i,a_1}, \alpha_{i,a_2}, \dots, \alpha_{i,a_L}) \in \Delta(\mathcal{A}_i), \quad (1.41)$$

where L is the total number of transmission alignments, α_{i,a_l} is the long term probability of the SBS taking action a_l .

Two games are formulated (\mathcal{G}^1 and \mathcal{G}^2) based on two payoff functions as follows:

$$\pi_i^1(a_i(n), \mathbf{a}_{-i}(n)) = \sum_{k \in \mathcal{K}} \log_2(1 + \gamma_i^k(n))_{\mathbb{I}_{\{\gamma_0^k(n) > \Gamma_0^k\}}}, \quad (1.42)$$

$$\pi_i^2(a_i(n), \mathbf{a}_{-i}(n)) = \sum_{i \in \mathcal{N}} \sum_{k \in \mathcal{K}} \log_2(1 + \gamma_i^k(n))_{\mathbb{I}_{\{\gamma_0^k(n) > \Gamma_0^k\}}}, \quad (1.43)$$

where γ_i^k is the SINR at the user served by SBS i on subcarrier k , γ_0^k is the SINR at macro user receiving on subcarrier k , Γ_0^k is the SINR threshold at macro user on subcarrier k , and n denotes the time step.

The long term average value of the payoff (for both games) is given by

$$\bar{\pi}_i(s_i, \mathbf{s}_{-i}) = \sum_{a \in \mathcal{A}} \pi_i(a_i, \mathbf{a}_{-i}) \prod_{j=1}^N \alpha_{j,a_j}. \quad (1.44)$$

The solutions obtained for the above games are in the notion of logit equilibrium [74]. Before defining the logit equilibrium, it is necessary to understand the concept of Smoothed Best Response (SBR).

DEFINITION 1.7 *Smoothed best response*: The smoothed best response of player i with parameter m_i is given by

$$\beta_i^{m_i}(\mathbf{s}_{-i}) = \left(\beta_{i,i}^{m_i}(\mathbf{s}_{-i}), \dots, \beta_{i,L}^{m_i}(\mathbf{s}_{-i}) \right), \quad (1.45)$$

where

$$\beta_{i,l}^{m_i}(\mathbf{s}_{-i}) = \frac{\exp\left(m_i \bar{u}_i(\mathbf{e}_l^{(L)}, \mathbf{s}_{-i})\right)}{\sum_{t=1}^L \exp\left(m_i \bar{u}_i(\mathbf{e}_t^{(L)}, \mathbf{s}_{-i})\right)}, \quad (1.46)$$

in which the vector $\mathbf{e}_l^{(L)} = (e_{l,1}^{(L)}, e_{l,2}^{(L)}, \dots, e_{l,L}^{(L)}) \in \mathbb{R}^L$ denotes the s^{th} vector of the canonical base spanning the space of real vectors of dimension S , (i.e., $e_{l,t}^{(L)} = 0$ for $t \in \{1, 2, \dots, L\} \setminus \{l\}$ and $e_{l,l}^{(L)} = 1$). Note that SBR is equivalent to best response when $m_i \rightarrow \infty$. For finite $m_i > 0$, SBR assigns high probabilities to the actions associated with high average payoffs.

Using the above definition, the logit equilibrium is defined as follows:

DEFINITION 1.8 *Logit equilibrium*: A strategy profile $\mathbf{s}^* = (\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_N) \in \Delta(\mathcal{A}_1) \times \dots \times \Delta(\mathcal{A}_N)$ is logit equilibrium with parameters $m_i (\forall i \in \mathcal{N}) > 0$ of the \mathcal{G}^1 or \mathcal{G}^2 if

$$\mathbf{s}_i^* = \beta_i^{m_i}(\mathbf{s}_{-i}^*), \quad (1.47)$$

where $\beta_i^{m_i}$ is the smoothed best response of player i with parameter m_i .

Note that \mathbf{s}^* is an ϵ -equilibrium with $\epsilon = \max_{i \in \mathcal{N}} \left(\frac{1}{m_i} \ln(L) \right)$. Also, it can be observed that $\epsilon \rightarrow 0$ for large m_i which means ϵ -equilibrium reaches a pure strategy NE when m_i is large enough.

As both the above games are finite games, the existence of the LE can be proved following the Theorem 1 in [75]. Due to the fact that \mathcal{G}^2 is a potential game, it can be proved that the convergence of the SBR dynamics is guaranteed for \mathcal{G}^2 if each SBS possesses the complete information of the strategies of other SBSs [76]. For self-organizing small cell networks, [70] proposes an RL-based technique in order to learn and reach the equilibrium. Each SBS makes an estimation (given by equation (1.48) on their own instantaneous payoff ($\pi(a_k(n), \mathbf{a}_{-k}(n))$)) based on user feedback as follows:

$$\tilde{\pi}_i(n) = \pi(a_i(n), \mathbf{a}_{-i}(n)) + \epsilon_{i,a_i(n)}(n), \quad (1.48)$$

where $\epsilon_{i,a_i(n)}(n)$ represents the error of the estimation due to thermal noise and it is also assumed that $E[\epsilon_{i,a_i(n)}(n)] = 0, \forall i$.

Each SBS should estimate the expected utility it achieves with each of its actions in order to build the SBR. Two coupled RL processes are proposed in [70] to achieve the LE. The first RL process allows SBSs to build an estimate of the vector of average payoffs $\bar{\pi}_i(\cdot, \boldsymbol{\pi}_{-i}(n))$ using observations $\tilde{\pi}_i(n)$, where $\bar{\pi}_i(\cdot, \boldsymbol{\pi}_{-i}(n)) = \left(\bar{\pi}_i(\mathbf{e}_1^{(L)}, \boldsymbol{\pi}_{-i}(n)), \dots, \bar{\pi}_i(\mathbf{e}_L^{(L)}, \boldsymbol{\pi}_{-i}(n)) \right)$. The first process is given by equation (1.49). The second RL process (given in equation (1.50)) uses the vector of estimated average payoffs at time n to update the transmission probability vector $\mathbf{s}_i(n)$, where w_i^1 and w_i^2 are learning parameters.

$$\hat{\pi}_{i,l}(n) = \hat{\pi}_{i,l}(n-1) + w_i^1(n) \mathbb{1}_{\{a_i(n)=l\}} (\tilde{\pi}_i(n) - \hat{\pi}_{i,l}(n-1)). \quad (1.49)$$

$$s_{i,l}(n) = s_{i,l}(n-1) + w_i^2(n) \left(\tilde{\beta}_{i,l}^{(m_i)}(\hat{\mathbf{u}}_i(n)) - s_{i,l}(n-1) \right). \quad (1.50)$$

The parameters should satisfy the following conditions:

$$\begin{aligned}
\lim_{T \rightarrow \infty} \sum_{t=1}^T w_i^1(t) &= +\infty, \\
\lim_{T \rightarrow \infty} \sum_{t=1}^T (w_i^1(t))^2 &< +\infty, \\
\lim_{T \rightarrow \infty} \sum_{t=1}^T w_i^2(t) &= +\infty, \\
\lim_{T \rightarrow \infty} \sum_{t=1}^T (w_i^2(t))^2 &< +\infty, \\
\lim_{T \rightarrow \infty} \frac{w_i^1(t)}{w_i^2(t)} &= 0,
\end{aligned}$$

and

$$\forall i \in \mathcal{N}, w_i^2 = w^2, \quad (1.51)$$

or

$$\forall i \in \mathcal{N}/\{N\} \lim_{T \rightarrow \infty} \frac{w_i^2(t)}{w_i^2(t+1)} = 0. \quad (1.52)$$

[70] proves that the convergence point (if there is a one) of the above given RL algorithm is an LE. In addition to that it is also proved that the convergence of \mathcal{G}^2 is always guaranteed as it is a potential game and all players share identical interests.

1.6.2 Q-Learning

Q-learning proposed by Watkins in [77] is also a form of reinforcement learning technique which can be used to find an optimal decision policy for any given finite Markov decision process (MDP) problem without knowledge of the transition probabilities. It has also been shown that Q-learning algorithm converges to the optimal policy for the systems with centralized control [71, 77]. Q-learning has recently been applied in the field of cognitive radio and wireless communications. For example, [78] investigates the problem of network selection in a heterogenous network and [65] applies Q-learning based learning technique for channel selection in multi-user cognitive radios. Q-learning based distributed resource allocation algorithm is devised in [79] in order to reduce interference in a network where small cells coexist with the macro network. In the following, we provide an example of the use of Q-learning for downlink resource allocation in a two-tier small cell network [60].

Small cell base stations (denoted by the set \mathcal{N}) form the player set and universal frequency reuse with K subcarriers is considered. It is also assumed that there is only one macro user receiving on each subcarrier at a given time. Each of

these macro users has with a minimum SINR requirement. Similar to the system model of the example given in Section 1.6.1, each SBS is able to select its transmission power from a finite set of values. Therefore, the set of subcarrier-power level combinations (i.e., transmission alignments) defines possible set of actions for each SBS.

The states for each player i at time t is defined as follows.

$$\mathbf{s}_i(t) = \left(s_i^{(1)}(t), s_i^{(2)}(t), \dots, s_i^{(K)}(t) \right), \quad (1.53)$$

where $s_i^{(k)}(t)$ takes the value 0 if the SBS i violates the QoS constraint for macro user on subcarrier k and $s_i^{(k)}(t) = 1$, otherwise. The action and utility vectors of each SBS at time t are given by

$$\mathbf{a}_i(t) = \left(a_i^{(1)}(t), a_i^{(2)}(t), \dots, a_i^{(K)}(t) \right), \quad (1.54)$$

and

$$\mathbf{u}_i(t) = \left(u_i^{(1)}(t), u_i^{(2)}(t), \dots, u_i^{(K)}(t) \right), \quad (1.55)$$

where $a_i^{(1)}(t) \in \{0, 1\}$.

Each SBS i observes its current state $\mathbf{s}_i(t)$ and takes an action $\mathbf{a}_i(t)$ based on the decision policy $\Phi : \mathbf{s} \rightarrow \mathbf{a}$. Our objective is to find an optimal decision policy Φ^* .

For each SBS, a Q -function maintains the knowledge of other players based on which the decisions can be taken individually without interacting with other players.

The expected discounted reward over a finite horizon is given by

$$V^\Phi(s) = E\{\gamma^t \times r(s_t, \Phi^*(s_t)) | s_0 = s\}, \quad (1.56)$$

where $0 \leq \gamma \leq 1$ is the discount factor at time t and r is the reward. The above equation can be re-written as

$$V^\Phi(s) = R(s, \Phi^*(s)) + \gamma \sum_{s' \in \mathcal{S}} p_{s,s'}(\Phi(s)) V^\Phi(s'), \quad (1.57)$$

where $R(s, \Phi^*(s))$ is the mean value of the reward $r(s, \Phi(s))$, $p_{s,s'}$ is the transition probability from state s to s' and $0 \leq \gamma \leq 1$ is the discount factor. The optimal policy Φ^* gives the optimal discounted reward $V^*(s)$. Hence,

$$V^*(s) = V^{\Phi^*}(s) = \max_{\forall a} \left(R(s, a) + \gamma \sum_{s' \in \mathcal{S}} p_{s,s'}(a) V^*(s') \right). \quad (1.58)$$

A Q value is maintained to learn the expected discounted reward. For an agent who takes action a when it is at state s and then follows the policy Φ , the expected discounted utility (which is the Q value) is given by

$$Q^*(s, a) = R(s, a) + \gamma \sum_{s' \in \mathcal{S}} P_{s,s'} V^\Phi(s'). \quad (1.59)$$

Each agent keeps trying all action-state combinations with a non-zero probability. The Q -learning algorithm utilizes the obtained reward at each step to update the Q values according to the following equation:

$$Q_t((s), (a)) = (1 - \alpha)Q_{t-1}((s), (a)) + \alpha \left[r_t((s), (a)) + \gamma \max_{b \neq a} Q_{t-1}(v, b) \right], \quad (1.60)$$

where α is the learning rate.

Simulation results show that this technique reach the convergence after few iterations for the above mentioned system model. However, the evolutionary game based learning algorithm in Section 1.5.4 shows faster convergence than the Q -learning based algorithm. The faster convergence is achieved at the expense of more information exchange among base stations.

1.6.3 Regret-matching Learning

Regret-matching [80] is a learning technique which can converge to a Correlated Equilibrium (CE) in finite games. The notion of CE is based on having a correlating mechanism for the players. This correlating mechanism provides a probability distribution over the set of actions of each player which provides an assignment *recommendation* for each action. Such assignment recommendation is said to be in CE if none of the players would benefit by deviating from the recommendation. For a more formal definition of CE, we first define the game $\mathcal{G} = \{\mathcal{N}, (\mathcal{S}_i)_{i \in \mathcal{N}}, (\mathcal{P}_i)_{i \in \mathcal{N}}\}$ where \mathcal{N} is the set of players, \mathcal{S}_i is the action set of i^{th} player and \mathcal{P}_i gives the set of payoffs that can be obtained by the i^{th} player. $\mathcal{S} = \prod_{i \in \mathcal{N}} (\mathcal{S}_i)$ is the set of N -tuples of the strategies. Let s denote any element in \mathcal{S} , s^i denotes an element of \mathcal{S}_i and $\pi_i \in \mathcal{P}$ is the payoff of the player i .

DEFINITION 1.9 *Correlated Equilibrium*: A probability distribution Ψ over \mathcal{S} gives a correlated equilibrium for the game \mathcal{G} , if $\forall i \in \mathcal{N}$, $\forall k \in \mathcal{S}_i$ and $\forall \mathbf{s}_{-i} \in \mathcal{S}_{-i}$,

$$\sum_{\mathbf{s}_{-i} \in \mathcal{S}_{-i}} \Psi(s) (\pi^i(k, \mathbf{s}_{-i}) - \pi^i(s, \mathbf{s}_{-i})) \leq 0, \quad (1.61)$$

where \mathcal{S}_{-i} is the set of actions played by the opponents of player i . Every Nash equilibrium is also a correlated equilibrium which corresponds to the case where the recommendations are not correlated at all.

In regret matching algorithm, a player would take decisions in order to minimize the regret. The regret of a player playing action s' is defined as the difference between the average payoff that the player would have achieved if she played the action s' all the time and the average current payoff. Regret of player i playing action s at n^{th} step is defined as follows.

$$r_i^{(s')}(n) = \frac{1}{n-1} \sum_{t=1}^{n-1} (\pi_i(s', \mathbf{s}_{-i}(t)) - \pi_i(s(t), \mathbf{s}_{-i}(t))), \quad (1.62)$$

where $s(t)$ denotes the action played by the corresponding player (k in the above

equation) at time step t . The steps of the regret matching algorithm are given in Algorithm 1.6.3.

Algorithm 3 Regret Matching Algorithm

- 1: For $t = 1, 2, 3, \dots$
 - 2: Calculate the regret for each user using equation (1.62).
 - 3: Obtain the regret vector for each player, i.e., $\mathbf{R}_i(n) = (\forall s \in \mathcal{S} : r_i^{(s)}(n))$.
 - 4: Obtain the probability distribution $\Psi_i(n)$ by normalizing $\mathbf{R}_i(n)$.
 - 5: The action played at time step n is chosen according to the probability distribution $\Psi_i(n)$.
-

It is also known that regret-matching learning can converge towards pure strategy NE points of exact potential games [81]. However, regret-based learning algorithm assumes that each player can calculate the expected payoff that it would have achieved by playing any action other than the current action. Therefore, a considerable amount of information exchange might be needed for this algorithm to be implemented.

In the context of self-organizing small cells, the aforementioned regret-based learning algorithm can be modified to implement a fully decentralized algorithm which only based on the SINR feedback of the users to the base station [38]. This modified algorithm converges to an ϵ -coarse correlated equilibrium.

1.6.4 Learning by Cooperation

The performance of learning mechanisms can be significantly improved by enabling cooperation among the learners [82, 83]. Small cell base stations who cooperate with neighboring base stations in order to speed up and improve their learning process are called *docitive base stations*. Cooperation is generally done via the backhaul. Network nodes are expected to select other nodes which operate under similar conditions to learn from. The similarity between two base stations are captured by a *gradient* which is defined based on the network architecture.

Several different cases of docition can be identified based on the degree of docition [84, 6].

- Startup docition: When a small cell base station connects to the network for the first time, it can learn the policies from other SBSs with similar gradients by exchanging Q tables.
- IQ-driven docition: SBSs with similar gradients share their policies periodically.
- Performance-driven docition: Base stations share their policies with less expert nodes, based on their ability to meet a pre-defined QoS targets.

In addition to the above discussed techniques, there are a number of other

learning techniques that are potentially applicable for self-organizing small cell networks. For instance, logit learning algorithm and its variants (i.e., max-logit algorithm and binary logit algorithm) converge to NE in potential games [85]. Learning automata [86] which is a branch of adaptive control theory is also another potential technique to implement distributed learning in self-organizing small cell networks. Stochastic learning automata-based channel selection algorithm has been proposed for opportunistic spectrum access in cognitive radio networks in [87]. This stochastic learning automata-based algorithm can converge to a pure strategy NE point for any exact potential game. Since development of distributed learning techniques for small cell networks has attracted a significant attention from the research community recently, new learning techniques are still being emerged.

1.7 Conclusion

In this chapter, we have discussed game theory approaches and learning techniques for self-organization in small cell networks. First we have given an overview of self-organizing networks including the motivations of enabling self-organizing functionalities in densely deployed small cells. Then a brief introduction to game theory has been given and the motivations of using game theory in self-organizing small cell networks have been discussed. Also, some widely used game models have been explained. Then a few examples have been given to explain how game theory can be used to solve the problem of self-organization in small cell networks. This chapter has also discussed learning techniques that can be used in small cell networks.

Some of the future research issues in designing self-organizing small cell networks are outlined below.

1. *Incomplete information games*: In order to address the issue of incomplete information, most of the existing algorithms use RL based techniques as we discussed in this chapter. However, models to address partial information can also be developed using Bayesian games. By using the Bayesian theorem, a belief on the parameters of other players can be constructed. The solution concept obtained in such games is the Bayesian Nash equilibrium.
2. *Multi radio access technology (Multi-RAT)*: In future networks, different radio access technologies (e.g., Wi-Fi, small cells) are expected to be integrated in order to provide seamless service to the user. Access control between different technologies should be done in a self-organizing way to achieve the optimal performance.
3. *Signaling overhead-optimal performance trade-off*: There is always a trade-off between the signaling overhead and optimal performance of a network. A network may deliver optimal performance with complete information but the signaling cost for implementing such algorithms would be higher. On

the other hand, the algorithms that rely on less information or incomplete information may deliver slightly degraded performance. Addressing this issue and quantifying the trade-off is significant in order to achieve near-optimal or optimal performance in self-organizing networks.

4. *Context-awareness*: Context awareness which is a powerful feature in many intelligent systems is recently proposed to be applied for enhancing self-organizing features in small cell networks. The idea is to utilize the context information, i.e., information from the users' environment, behavior, and social media, to enhance the provision of services and applications. The algorithms should be devised considering the efficient exploitation of context aware information taken from different sources. The reliability of the different information sources would also be an important issue.

References

- [1] V. Chandrasekhar, "Femtocell networks: a survey," *IEEE Communications Magazine*, pp. 59–67, September 2008.
- [2] P. Demestichas and A. Georgakopoulos, "5G on the Horizon: Key Challenges for the Radio-Access Network," *IEEE Vehicular Technology Magazine*, pp. 47–53, september 2013.
- [3] O. Aliu, A. Imran, M. Imran, and B. Evans, "A survey of self organisation in future cellular networks," *IEEE Communications Surveys and Tutorials*, vol. 15, pp. 336–361, 2012.
- [4] H. Claussen, L. T. Ho, and L. G. Samuel, "An overview of the femtocell concept," *Bell Labs Technical Journal*, vol. 13, pp. 221–245, 2008.
- [5] N. Johnson, "Small cell forum and lte small cells," The Small Cell Forum, Tech. Rep., 2011. [Online]. Available: http://www.smallcellforum.org/smallcellforum_resources/pdfs7/
- [6] M. Bennis and L. Giupponi, "Interference management in self-organized femtocell networks: the BeFEMTO approach," in *2nd International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace Electronic Systems Technology (Wireless VITAE)*, Chennai, 2011, pp. 5–10.
- [7] J. V. den Berg and R. Litjens, "SOCRATES: Self-Optimisation and self-ConfiguRation in wirelEss networkS," *Fourth ERCIM Workshop on eMobility*, 2008.
- [8] T. Jansen, M. Amirijoo, U. Turke, L. Jorguseski, K. Zetterberg, R. Nascimento, L.-C. Schmelz, J. Turk, and I. Balan, "Embedding multiple self-organisation functionalities in future radio access networks," in *IEEE 69th Vehicular Technology Conference*. IEEE, 2009, pp. 1–5.
- [9] EU FP7 project. End-to-End Efficiency (E3). [Online]. Available: <https://ict-e3.eu/>
- [10] E. Bogenfeld and I. Gaspard, "Self-x in radio access networks," *E3 White Paper v1. 0*, vol. 22, 2008.
- [11] E. Bonabeau, M. Dorigo, and G. Theraulaz, *Swarm intelligence: from natural to artificial systems*. Oxford university press New York, 1999, vol. 4.
- [12] 3GPP TS 32.500, "Telecommunication management; self-organizing networks (son); concepts and requirements."
- [13] 3GPP TR 36.902, "Evolved universal terrestrial radio access network (e-utran); self-configuring and self-optimizing network (son) use cases and solutions."
- [14] NGMN Alliance, "Next generation mobile networks beyond HSPA & EVDO," *White PaperDec*, vol. 5, 2006.

-
- [15] NGMN Alliance, "Next generation mobile networks recommendation on son and o&m requirements," *Req. Spec. v1*, vol. 23, 2008.
 - [16] S. Haykin, "Cognitive dynamic systems," in *IEEE International Conference on Acoustics, Speech and Signal Processing, 2007.*, vol. 4, 2007, pp. IV-1369-IV-1372.
 - [17] Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 201-220, 2005.
 - [18] A. Spilling, A. Nix, M. Beach, and T. Harrold, "Self-organisation in future mobile communications," *Electronics and Communication Engineering Journal*, vol. 12, pp. 133-147, 2000.
 - [19] C. Prehofer and C. Bettstetter, "Self-organization in communication networks: principles and design paradigms," *IEEE Communications Magazine*, vol. 43, no. 7, pp. 78-85, 2005.
 - [20] S. Hämmäläinen, H. Sanneck, C. Sartori *et al.*, *LTE Self-Organising Networks (SON): Network Management Automation for Operational Efficiency*. John Wiley & Sons, 2012.
 - [21] I. Viering, M. Döttling, and A. Lobinger, "A mathematical perspective of self-optimizing wireless networks," in *IEEE International Conference on Communications*, 2009, pp. 1-6.
 - [22] H. Hu, J. Zhang, and X. Zheng, "Self-configuration and self-optimization for LTE networks," *IEEE Communications Magazine*, pp. 94-100, February 2010.
 - [23] H. Sanneck, Y. Bouwen, and E. Troch, "Context based configuration management of plug & play LTE base stations," *IEEE Network Operations and Management Symposium - NOMS 2010*, pp. 946-949, 2010.
 - [24] "Self-Optimizing Networks: The Benefits of SON in LTE." 4G Americas, 2013. [Online]. Available: <http://www.4gamericas.org>
 - [25] NGMN Alliance, "Use cases related to self organising network: Overall description," 2007.
 - [26] Y. Choi, C. Kim, and S. Bahk, "Flexible design of frequency reuse factor in OFDMA cellular networks," in *IEEE International Conference on Communications*, vol. 4, 2006, pp. 1784-1788.
 - [27] O. Ileri, S.-C. Mau, and N. B. Mandayam, "Pricing for enabling forwarding in self-configuring ad hoc networks," *Journal on Selected Areas in Communications*, vol. 23, pp. 151-162, 2005.
 - [28] P. Semasinghe, E. Hossain, and K. Zhu, "An Evolutionary Game Approach for Distributed Resource Allocation for Self-Organizing Small Cells," *IEEE Transactions of Mobile Computing*, 2013, "to appear".
 - [29] M. Dirani and Z. Altman, "Self-organizing networks in next generation radio access networks: Application to fractional power control," *Computer Networks*, vol. 55, no. 2, pp. 431-438, 2011.
 - [30] 3GPP TS 32.541, "Evolved universal terrestrial radio access network (e-utran); self-organizing networks (son); self-healing concepts and requirements (release 11)."
 - [31] A. Attar, V. Krishnamurthy, and O. N. Gharehshiran, "Interference management using cognitive base-stations for umts lte," *IEEE Communications Magazine*, vol. 49, pp. 152-159, 2011.

-
- [32] K. Zhu, E. Hossain, and D. Niyato, "Pricing, spectrum sharing, and service selection in two-tier small cell networks: A hierarchical dynamic game approach," *IEEE Transactions on Mobile Computing*, 2013.
 - [33] X. Huang and B. Beferull-Lozano, "Non-cooperative power allocation game with imperfect sensing information for cognitive radio," in *IEEE International Conference on Communications (ICC)*. IEEE, 2012, pp. 1666–1671.
 - [34] S. Guruacharya, D. Niyato, M. Bennis, and D. Kim, "Dynamic coalition formation for network mimo in small cell networks," *IEEE Transactions on Wireless Communications*, 2013.
 - [35] R. Madan, A. Sampath, N. Bhushan, A. Khandekar, J. Borran, and T. Ji, "Impact of coordination delay on distributed scheduling in lte-a femtocell networks," in *IEEE Global Telecommunications Conference (GLOBECOM 2010)*, 2010, pp. 1–5.
 - [36] Z. Han, D. Niyato, W. Saad, T. Basar, and A. Hjorungnes, *Game theory in wireless and communication networks*. Cambridge University Press, 2012.
 - [37] K. Akkarajitsakul, E. Hossain, D. Niyato, and D. I. Kim, "Game theoretic approaches for multiple access in wireless networks: A survey," *IEEE Communications Surveys and Tutorials*, vol. 13, pp. 372–395, 2011.
 - [38] M. Bennis, S. M. Perlaza, and M. Debbah, "Learning coarse correlated equilibria in two-tier wireless networks," in *IEEE International Conference on Communications (ICC)*, 2012, pp. 1592–1596.
 - [39] E. J. Hong, S. Y. Yun, and D.-H. Cho, "Decentralized power control scheme in femtocell networks: A game theoretic approach," in *IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, 2009, pp. 415–419.
 - [40] D. M. Topkis, *Supermodularity and complementarity*. Princeton University Press, 1998.
 - [41] E. Altman and Z. Altman, "S-modular games and power control in wireless networks," *IEEE Transactions on Automatic Control*, vol. 48, pp. 839–842, 2003.
 - [42] A. B. MacKenzie and L. A. DaSilva, "Game theory for wireless engineers," *Synthesis Lectures on Communications*, vol. 1, pp. 1–86, 2006.
 - [43] N. Nisan, *Algorithmic game theory*. Cambridge University Press, 2007.
 - [44] M. K. H. Yeung and Y.-K. Kwok, "A game theoretic approach to power aware wireless data access," *IEEE Transactions on Mobile Computing*, vol. 5, pp. 1057–1073, 2006.
 - [45] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 1074–1084, 2006.
 - [46] H. Li, Y. Gai, Z. He, K. Niu, and W. Wu, "Optimal power control game algorithm for cognitive radio networks with multiple interference temperature limits," in *Vehicular Technology Conference, Spring 2008*. IEEE, 2008, pp. 1554–1558.
 - [47] P. Reichl, B. Tuffin, and R. Schatz, "Logarithmic laws in service quality perception: where microeconomics meets psychophysics and quality of experience," *Telecommunication Systems*, pp. 1–14, 2011.
 - [48] L. Huang, Y. Zhou, X. Han, Y. Wang, M. Qian, and J. Shi, "Distributed coverage optimization for small cell clusters using game theory," in *IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2013, pp. 2289–2293.

-
- [49] S. Subramani, T. Basar, S. Armour, D. Kaleshi, and Z. Fan, "Noncooperative equilibrium solutions for spectrum access in distributed cognitive radio networks," in *3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN 2008*. IEEE, 2008, pp. 1–5.
 - [50] W. Ma, H. Zhang, W. Zheng, and X. Wen, "Differentiated-pricing based power allocation in dense femtocell networks," in *15th International Symposium on Wireless Personal Multimedia Communications (WPMC)*. IEEE, 2012, pp. 599–603.
 - [51] V. Chandrasekhar, J. G. Andrews, T. Muharemovic, Z. Shen, and A. Gatherer, "Power control in two-tier femtocell networks," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 4316–4328, 2009.
 - [52] I. L. Glicksberg, "A further generalization of the kakutani fixed point theorem, with application to nash equilibrium points," *Proceedings of the American Mathematical Society*, vol. 3, pp. 170–174, 1952.
 - [53] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave n-person games," *Econometrica: Journal of the Econometric Society*, pp. 520–534, 1965.
 - [54] G. Debreu, "A social equilibrium existence theorem," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 38, p. 886, 1952.
 - [55] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 1341–1347, 1995.
 - [56] I. W. Mustika, K. Yamamoto, H. Murata, and S. Yoshida, "Potential game approach for self-organized interference management in closed access femtocell networks," in *IEEE 73rd Vehicular Technology Conference (VTC Spring)*. IEEE, 2011, pp. 1–5.
 - [57] 3GPP TR 36.814 v9.0.0, "Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancement of E-UTRA physical layer aspects (Release 9)," March 2010.
 - [58] L. Giupponi and C. Ibars, "Distributed interference control in ofdma-based femtocells," in *IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*. IEEE, 2010, pp. 1201–1206.
 - [59] J.-H. Yun and K. G. Shin, "Adaptive interference management of OFDMA femtocells for co-channel deployment," *IEEE Journal on Selected Areas in Communications*, vol. 29, pp. 1225–1241, 2011.
 - [60] M. Bennis, S. Guruacharya, and D. Niyato, "Distributed learning strategies for interference mitigation in femtocell networks," in *IEEE Global Telecommunications Conference (GLOBECOM 2011)*. IEEE, 2011, pp. 1–5.
 - [61] P. Semasinghe, K. Zhu, and E. Hossain, "Distributed Resource Allocation for Self-Organizing Small Cell Networks: An Evolutionary Game Approach," *IEEE GLOBECOM Workshops (GC Wkshps)*, 2013.
 - [62] J. W. Weibull, *Evolutionary game theory*. MIT press, 1997.
 - [63] S. Samarakoon, M. Bennis, W. Saad, and M. Latva-aho, "Backhaul-Aware Interference Management in the Uplink of Wireless Small Cell Networks," *IEEE Transactions on Wireless Communications*, 2013, "to appear".
 - [64] G. Gur, S. Bayhan, and F. Alagoz, "Cognitive femtocell networks: an overlay architecture for localized dynamic spectrum access [dynamic spectrum management]," *IEEE Wireless Communications*, vol. 17, pp. 62–70, 2010.
 - [65] H. Li, "Multi-agent q-learning of channel selection in multi-user cognitive radio systems: a two by two case," in *IEEE International Conference on Systems, Man and Cybernetics, SMC*. IEEE, 2009, pp. 1893–1898.

-
- [66] J. W. Huang and V. Krishnamurthy, "Cognitive base stations in LTE/3GPP femtocells: a correlated equilibrium game-theoretic approach," *IEEE Transactions on Communications*, vol. 59, pp. 3485–3493, 2011.
 - [67] R. J. Aumann, "Correlated equilibrium as an expression of bayesian rationality," *Econometrica: Journal of the Econometric Society*, pp. 1–18, 1987.
 - [68] M. Nazir, M. Bennis, K. Ghaboosi, A. B. MacKenzie, and M. Latva-aho, "Learning based mechanisms for interference mitigation in self-organized femtocell networks," in *Forty Fourth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, 2010, pp. 1886–1890.
 - [69] F. Pantisano, M. Bennis, W. Saad, M. Latva-aho, and R. Verdone, "Enabling macrocell-femtocell coexistence through interference draining," in *Wireless Communications and Networking Conference Workshops (WCNCW)*, 2012, pp. 81–86.
 - [70] M. Bennis, S. M. Perlaza, P. Blasco, Z. Han, and H. V. Poor, "Self-organization in small cell networks: A reinforcement learning approach," *IEEE Transactions on Wireless Communications*, vol. 12, pp. 3202–3212, 2013.
 - [71] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. Cambridge Univ Press, 1998, vol. 1, no. 1.
 - [72] M. Bkassiny, Y. Li, and S. Jayaweera, "A survey on machine-learning techniques in cognitive radios," *IEEE Communications Surveys Tutorials*, vol. 15, pp. 1136–1159, 2013.
 - [73] M. Bennis and S. M. Perlaza, "Decentralized cross-tier interference mitigation in cognitive femtocell networks," in *IEEE International Conference on Communications (ICC)*. IEEE, 2011, pp. 1–5.
 - [74] D. L. McFadden, "Quantal choice analysis: A survey," in *Annals of Economic and Social Measurement, Volume 5, number 4*. NBER, 1976, pp. 363–390.
 - [75] R. D. McKelvey and T. R. Palfrey, "Quantal response equilibria for normal form games," *Games and economic behavior*, vol. 10, pp. 6–38, 1995.
 - [76] J. Hofbauer and W. H. Sandholm, "On the global convergence of stochastic fictitious play," *Econometrica*, vol. 70, pp. 2265–2294, 2002.
 - [77] C. J. C. H. Watkins, "Learning from delayed rewards." Ph.D. dissertation, University of Cambridge, 1989.
 - [78] D. Niyato and E. Hossain, "Dynamics of network selection in heterogeneous wireless networks: an evolutionary game approach," *IEEE Transactions on Vehicular Technology*, vol. 58, pp. 2008–2017, 2009.
 - [79] M. Bennis and D. Niyato, "A q-learning based approach to interference avoidance in self-organized femtocell networks," in *IEEE GLOBECOM Workshops*. IEEE, 2010, pp. 706–710.
 - [80] S. Hart and A. Mas-Colell, "A simple adaptive procedure leading to correlated equilibrium," *Econometrica*, vol. 68, pp. 1127–1150, 2000.
 - [81] N. Nie and C. Comaniciu, "Adaptive channel allocation spectrum etiquette for cognitive radio networks," *Mobile networks and applications*, vol. 11, pp. 779–797, 2006.
 - [82] M. N. Ahmadabadi and M. Asadpour, "Expertness based cooperative q-learning," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 32, pp. 66–76, 2002.

-
- [83] M. Tan, "Multi-agent reinforcement learning: Independent vs. cooperative agents," in *Proceedings of the tenth international conference on machine learning*, vol. 337. Amherst, MA, 1993.
 - [84] A. G. Serrano, L. Giupponi, and M. Dohler, "Befemto's self-organized and docitive femtocells," in *Future Network and Mobile Summit*. IEEE, 2010, pp. 1–8.
 - [85] Y. Song, S. H. Wong, and K.-W. Lee, "Optimal gateway selection in multi-domain wireless networks: a potential game perspective," in *Proceedings of the 17th annual international conference on Mobile computing and networking*. ACM, 2011, pp. 325–336.
 - [86] K. S. Narendra and M. Thathachar, "Learning automata-a survey," *IEEE Transactions on Systems, Man and Cybernetics*, pp. 323–334, 1974.
 - [87] Y. Xu, J. Wang, Q. Wu, A. Anpalagan, and Y.-D. Yao, "Opportunistic Spectrum Access in Unknown Dynamic Environment: A Game-Theoretic Stochastic Learning Solution," *IEEE Transactions of Wireless Communications*, pp. 1380–1391, 2012.