



Universiteit
Leiden
The Netherlands

Resource allocation in networks via coalitional games

Shams, F.

Citation

Shams, F. (2016, September 21). *Resource allocation in networks via coalitional games*. Retrieved from <https://hdl.handle.net/1887/45667>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/45667>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/45667> holds various files of this Leiden University dissertation.

Author: Shams, F.

Title: Resource allocation in networks via coalitional games

Issue Date: 2016-09-21

Chapter 1

Coalitional game theory

Game theory is the study of decision making in an interactive environment. *Coalitional games* fulfill the promise of group efficient solutions to problems involving strategic actions. Formulation of optimal player behavior is a fundamental element in this theory. This chapter comprises a self-instructive didactic means indicating how cooperative game theory tools can provide a framework to tackle different network engineering problems. We show that coalitional game approaches achieve an improved performance compare to non-cooperative game theoretical solutions.

This chapter is divided into ten sections. After a brief motivation in the following section, Sect. 1.2 provides an introductory discussion of cooperative game theory. We systematically study fundamental definitions and conditions of cooperative games: superadditivity and convexity. Then, Sect. 1.3 and the sub-section inside discuss the core set solution as the most known solution for payoff distribution. Sect. 1.4 is devoted to a study of a strong payoff distribution, the so-called Shapley value. In Sect. 1.5 we present a systematic study of two others reward division called the *kernel* and *nucleolus*. Then, in Sect. 1.6, we extend the concept of Nash equilibria in cooperative games. Sect. 1.7 is an investigation of the concept of coordinated equilibria where players of game are admitted to pre-communicate among themselves at once. Finally, Sect. 1.8 helps a reader to understand the basic concepts and importance of dynamic learning in cooperative games. Every sections contain some motivation examples that are expedient to understand how different communication networks problems can be modeled as cooperative game. We discuss the features of mentioned approaches in Sect. 1.9 and finally we conclude this chapter in Sect. 1.10.

1.1 Motivation

The increase of the number of wireless services, combined with demand for high definition multimedia communications, have made the radio resources, and particularly the spectrum and power, a very precious and scarce resource, not because of their unavailability but because they are used inefficiently. For licensed spectrum, the measurements by Shared Spectrum Company [20] shows that the maximal usage of the spectrum is a low percentage of the whole licensed. While the number of users and the spectrum usage steadily increase, the amount of spectrum is still considered a limited resource. Beside, to differentiate between the true signal and background noise is complex for a radio equipment. Generally, this complex process enforces terminals to transmit strong version of signals, that wastes energy of a transmitter.

The modern wireless entities, i.e. wireless terminals and base stations, have considerable capacities to execute dynamic processes. This capability encourages wireless service providers to consider wireless entities as autonomous agents which could cooperate and negotiate with each other to achieve an efficient resource allocation in different situations. Cooperation among wireless terminals is usually intended to achieve a fair radio resource allocation. Cooperation between base stations can be devised to mitigate interference, and promote soft handover where channel gain is varying rapidly which is a challenge in LTE [21].

Game theory is the most prominent tool to analyze interaction issue in social sciences wherein often cooperation amongst autonomous agents is essential for successful task completion. In many settings, groups of competing agents are simultaneously concerned of both individual and overall benefits. In the game theory literature, this branch is known as cooperative game [10, 22]. The players, as the main decision making entities in the game, are considered to negotiate with each other to determine a binding agreement among them. If we assume that all users act rationally and we know what the behavior of the users are, it is possible to determine the overall performance of a system since the actions of one user becomes part of the circumstances for another user. Thus, we are interested in individual performance and overall system performance under a specific set of rules. To fully develop the different possibilities within a game for cooperation among players we have to address which groups the players can achieve collectively. Indeed, if a player assesses that within a certain group it does not receive what it is able to get by itself, then it might decide to abandon

the cooperation and pursue an alternative allocation by itself. Cooperative game theory offers the opportunity to extend and expand the treatment of the players in traditional non-cooperative games, especially where selfish players compete over a set of resources. The cooperative game theory is divided into two parts: coalitional game theory and bargaining games [10, 22]. In this contribution we focus on coalitional game theory.

Saad *et al.* in tutorial paper [23] classify coalitional games into three categories: canonical (coalitional) games, coalition formation games and coalitional graph games. In canonical games, no group of players can do worse by joining a coalition than by acting non-cooperatively. In coalition formation games, forming a coalition brings advantage to its members but the gains are limited by a cost for forming the coalition. In coalitional graph games, the coalitional game is in graph form and the interconnection between the players strongly affects the characteristics as well as the outcome of the game.

In the last few years, cooperative game theory has been successfully applied to communications and networking. Hossain *et al.* in [24] provides a guide to state-of-art which unifies the essential information, addressing both theoretical and practical aspects of cooperative communications and networking in the context of cellular design. The current literature is mainly focused on applying cooperative games in various applications such as distributed/centralized radio resource allocation [18,25,26], power control [27, 28], spectrum sharing in cognitive radio [29, 30], cooperative automatic repeat request (ARQ) mechanism [31], cooperative routing [32], and cooperative communications [33, 34]. These problems in wireless networks can be modeled as a cooperative game since it is highly likely that each wireless user can obtain a better utility value by forming groups and controlling resources cooperatively rather than individually. It has been shown that cooperation can result in an enhanced QoS in terms of throughput expansion, bit error rate reduction, or energy saving [24].

Cooperation can be realized at various layers of the network. At the physical layer, different separate antennas can constitute a cluster and then cooperate with each other to exploit multiple-input multiple-output (MIMO) gains. At the MAC sublayer, some wireless terminals can cooperate with each other to share a common wireless medium in an efficient manner and consequently mitigate the interference hazard. There is also the possibility of cooperation of physical and application layers among individual terminals to adapt channel and source codings in multimedia communications.

The altruistic decision of cooperation with others network entities may result in an improvement on overall network performance, and concurrently achieve an egoistic interest of self improvement.

1.2 Preliminaries

Game theory deals with the study, through mathematical models, of conflict situations in which two or more rational players make decisions that will influence each other's welfare. The theory of coalitional games [10,22] also assumes that binding agreements may be established among the players in the course of the conflict situation. In transferable utility (TU) games, agreement may be reached by any subset of the players, and the gain obtained from this agreement is a real number and it is transferable among these players. In non-transferable utility (NTU) games, agreement may be reached by any subset of the players, but the gain may be non-transferable. The main focus of this dissertation will be on the study of TU games.

A TU game is a pair $\mathcal{G} = (\mathcal{K}, \nu)$, where $\mathcal{K} = [1, \dots, K]$ denotes the set of players and ν the *coalition (characteristic) function* which is interpreted as the maximum outcome (a real number) to each coalition (subset of \mathcal{K}) whose players can jointly produce. An NTU game is a pair $\mathcal{G} = (\mathcal{K}, V)$ where V is a mapping which for each coalition \mathcal{A} , defines a characteristic set, $V(\mathcal{A})$, satisfying:

1. $V(\mathcal{A})$ is non-empty and closed subset of $\mathbb{R}^{|\mathcal{A}|}$,
2. For each $k \in \mathcal{A}$ there is a $V_k \in \mathbb{R}$ such that $V(\{k\}) = (-\infty, V_k]$,
3. $V(\mathcal{A})$ is comprehensive, i.e. for all $\mathbf{u} \in V(\mathcal{A})$ and for all $\mathbf{u}' \in \mathbb{R}^{|\mathcal{A}|}$, if $\mathbf{u}'[k] \leq \mathbf{u}[k] \forall k \in \mathcal{A}$ then $\mathbf{u}' \in V(\mathcal{A})$,
4. The set $V(\mathcal{A}) \cap \{\mathbf{u}' \in \mathbb{R}^{|\mathcal{A}|} \mid \mathbf{u}'[k] \geq V_k \forall k \in \mathcal{A}\}$ is bounded.

The characteristic set, $V(\mathcal{A})$, is interpreted as the set of achievable outcomes the players in \mathcal{A} can guarantee themselves without cooperating with the players in $\mathcal{K} \setminus \mathcal{A}$. In particular, an NTU-game $\mathcal{G} = (\mathcal{K}, V)$ is called a TU game when the characteristic set for each coalition \mathcal{A} , takes the form:

$$V(\mathcal{A}) = \left\{ \mathbf{u} \in \mathbb{R}^{|\mathcal{A}|} : \sum_{k \in \mathcal{A}} u_k \leq \nu(\mathcal{A}) \right\} \quad (1.1)$$

where $\mathbf{u} = [u_1, \dots, u_{|\mathcal{A}|}] \in \mathbb{R}^{|\mathcal{A}|}$ and u_k is the payoff of player k in \mathcal{A} and $\nu : 2^{\mathcal{K}} \rightarrow \mathbb{R}$. If \mathcal{A} is a coalition (subset) of \mathcal{K} formed in \mathcal{G} , then its members get an overall payoff $\nu(\mathcal{A})$, zero for the empty set. Each coalition can be represented as a pure strategy in non-cooperative game theory. There exist only few works on NTU games applications to problems in communications [35]. This is because defining an utility function which meets all conditions of a character set in NTU game is not always feasible.

An important property of interest in characteristic form TU games is *superadditivity*, which, if present, implies that the value of the unite of any two disjoint coalitions is at least as big as the sum of their values.

Definition 1 A TU game \mathcal{G} is superadditive if

$$\nu(\mathcal{A}_i \cup \mathcal{A}_j) \geq \nu(\mathcal{A}_i) + \nu(\mathcal{A}_j) \quad \forall \mathcal{A}_i, \mathcal{A}_j \subset \mathcal{K} \text{ s.t. } \mathcal{A}_i \cap \mathcal{A}_j = \emptyset \quad (1.2)$$

In a superadditive TU game there are positive synergies and the players prefer to join each other rather than act alone. Under superadditivity condition, the players are willing to form the *grand coalition* (the set \mathcal{K}).

Convex, or alternatively *supermodular* coalitional games were introduced by L. Shapley [36]. They model coalitional situations where the marginal contribution of a player to a coalition increases as the coalition becomes larger.

Definition 2 A TU game \mathcal{G} is convex or supermodular if for all $k \in \mathcal{K}$:

$$\nu(\mathcal{A}_i \cup \{k\}) - \nu(\mathcal{A}_i) \leq \nu(\mathcal{A}_j \cup \{k\}) - \nu(\mathcal{A}_j) \quad \forall \mathcal{A}_i \subseteq \mathcal{A}_j \subset \mathcal{K} \setminus \{k\} \quad (1.3)$$

Equivalently:

Definition 3 A TU game \mathcal{G} is convex or supermodular if:

$$\nu(\mathcal{A}_i) + \nu(\mathcal{A}_j) \leq \nu(\mathcal{A}_i \cap \mathcal{A}_j) + \nu(\mathcal{A}_i \cup \mathcal{A}_j) \quad \forall \mathcal{A}_i, \mathcal{A}_j \subseteq \mathcal{K} \quad (1.4)$$

Convexity means that there are increasing returns to scale. Note that a convex game is superadditive. To better understand the importance of convexity approach in network problems, we verify the convexity condition in a K -user channel access game. The payoff of each coalition of players (transmitters) is defined as the outer MAC capacity region. [37, Lemma 1] shows that in a multiple access channel scenario, the inequality (1.4) is not met. This means the game is not convex, and thus adding a new player does not give benefit to others transmitters.

1.3 The core solution

A central question in a coalitional game is how to divide the extra earnings (or cost savings) among the members of the formed coalition. In a TU game, an allocation is a function \mathbf{u} from \mathcal{K} to \mathbb{R} that specifies for each player $k \in \mathcal{K}$ the payoff $u_k \in \mathbb{R}$ that this player can expect when it cooperates with the other players. The payoff of each player can show the cost borne by the player, the power of influence, and so on depending on the problem setting.

Definition 4 Let \mathcal{K} be the set of K players of the superadditive TU game \mathcal{G} , and let ν be the payoff of the game. The set of all “imputations” of \mathcal{G} is the set:

$$\mathcal{I}(\mathcal{K}, \nu) = \left\{ \mathbf{u} \in \mathbb{R}^K : \left. \begin{array}{l} i) \quad \sum_{k \in \mathcal{K}} u_k = \nu(\mathcal{K}) \\ ii) \quad u_k \geq \nu(\{k\}) \quad \forall k \in \mathcal{K} \end{array} \right\} \quad (1.5)$$

where $\mathbf{u} = [u_1, \dots, u_k, \dots, u_K] \in \mathbb{R}^K$ is the imputation vector of the players. The former condition is called the feasibility, and the latter individually rational condition.

The core concept was introduced in [38] and is the most attractive and natural way to define a payoff distribution: if a payoff distribution is in the core, no agent has any incentive to be in a different coalition. The core of a TU game is the subset of all imputations $\mathbf{u} \in \mathcal{I}(\mathcal{K}, \nu)$ that no other imputation *directly dominates*, that is $\nexists \mathbf{u}' \in \mathcal{I}(\mathcal{K}, \nu)$ s.t. $u'_k > u_k \quad \forall k \in \mathcal{K}$. As can be seen, for coalitional games as well as non-cooperative games, the notion of dominance is essentially equivalent; the payoffs under the various situations are compared and one situation dominates the others if these payoffs are higher. The core actually presents a condition stronger than Nash equilibrium in non-cooperative game: no group of agents should be able to profitably deviate from a configuration in the core. Equivalently, no set of players can benefit from forming a new coalition, which corresponds to the group rationality assumption.

In an NTU game $\mathcal{G} = (\mathcal{K}, V)$, the core apportionment is defined as [22, Ch. 12]:

Definition 5 Let \mathcal{K} be the set of K players of the superadditive NTU-game \mathcal{G} , and let V be the payoff of the game. The core of \mathcal{G} is the set

$$\mathcal{S}(\mathcal{K}, \mathbf{V}) = \{ \mathbf{u} \in V(\mathcal{K}) : \forall \mathbf{u}' \in V(\mathcal{A}) \quad \exists k \in \mathcal{A} \text{ s.t. } u_k \geq u'_k \} \quad (1.6)$$

where \mathbf{u} is the payoff distribution across players, and $u_k \in \mathbf{u}$ if and only if no coalition can improve upon u_k .

In a TU game $\mathcal{G} = (\mathcal{K}, \nu)$, the core apportionment is defined as follows:

Definition 6 Let \mathcal{K} be the set of K players of the superadditive TU game \mathcal{G} , and let ν be the payoff of the game. The core of \mathcal{G} is the set

$$\mathcal{S}(\mathcal{K}, \nu) = \left\{ \mathbf{u} \in \mathbb{R}^K : \left. \begin{array}{l} i) \quad \sum_{k \in \mathcal{K}} u_k = \nu(\mathcal{K}) \\ ii) \quad \sum_{k \in \mathcal{A}} u_k \geq \nu(\mathcal{A}) \quad \forall \mathcal{A} \subset \mathcal{K} \end{array} \right\} \quad (1.7)$$

where $\mathbf{u} = [u_1, \dots, u_k, \dots, u_K] \in \mathbb{R}^K$ is the payoff distribution across players, and $u_k \in \mathbf{u}$ if and only if no coalition can improve upon u_k . The second condition is called non-blocking condition.

The core consists of the set of allocations that can be blocked by any coalition of agents. If for some set of agents \mathcal{A} , the non-blocking condition does not hold, then the agents in \mathcal{A} have an incentive to collectively deviate from the coalition structure and to divide $\nu(\mathcal{A})$ among themselves. In general, the core of a given TU game (\mathcal{K}, ν) is found by linear programming (LP) as:

$$\min_{\mathbf{u} \in \mathbb{R}^K} \sum_{k \in \mathcal{K}} u_k ; \quad \text{s.t.} \quad \sum_{k \in \mathcal{A}} u_k \geq \nu(\mathcal{A}) \quad \forall \mathcal{A} \subseteq \mathcal{K} \quad (1.8)$$

Madiman in [39] introduces some intuitive applications of core solution to information theory contexts e.g. source coding and multiple-access channel, and summarize some of its limitations in multi user scenarios. Li *et al.* in [40] show that the cooperation among wireless nodes and core apportionment can increase spectrum efficiency in a TDMA cooperative communication. In [41], Niyato *et al.* applies the core solution in a coalition among different wireless access networks to offer a stable and efficient bandwidth allocation.

Indeed, there is a number of realistic application scenarios, in which the emergence of the grand coalition is either not guaranteed, or might be perceivably harmful, or is plainly impossible [42]. For a non-superadditive coalitional game, the coalition formation process does not lead the players to form the grand coalition. In this case, Def. 6 does not apply. Let us redefine the core set in a general (not necessarily superadditive) coalitional formation TU game [26]. Let $\psi = [\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_m]$ denote a partition of the set \mathcal{K} wherein $\mathcal{A}_i \cap \mathcal{A}_j = \emptyset$ for $i \neq j$, $\bigcup_{i=1}^m \mathcal{A}_i = \mathcal{K}$ and $\mathcal{A}_i \neq \emptyset$ for $i = 1, \dots, m$, and let Ψ denote the set of all possible partitions ψ . Let us also

define $\mathcal{F} = [\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n]$, such that $\bigcup_{i=1}^n \mathcal{A}_i = \mathcal{K}$ and $\mathcal{A}_i \neq \emptyset$ for $i = 1, \dots, n$, as a family of (not necessarily disjoint) coalitions.

Definition 7 A “core apportionment” $\mathbf{u} \in \mathbb{R}^K$ is a payoff distribution with the following property:

$$S(\mathcal{K}, \nu) = \left\{ \mathbf{u} \in \mathbb{R}^K : \left. \begin{array}{l} i) \quad \sum_{k \in \mathcal{K}} u_k = \max_{\psi \in \Psi} \sum_{\mathcal{A} \in \psi} \nu(\mathcal{A}) \\ ii) \quad \sum_{k \in \mathcal{A}} u_k \geq \nu(\mathcal{A}) \quad \forall \mathcal{A} \subset \mathcal{K} \end{array} \right\} \quad (1.9)$$

Note that, if \mathcal{G} is superadditive, then $\max_{\psi \in \Psi} \sum_{\mathcal{A} \in \psi} \nu(\mathcal{A}) = \nu(\mathcal{K})$.

The core allocation set can be found through linear programming and its existence, in general, depends upon the feasibility of (1.8). Unfortunately, the core is a strong notion, and there exist many games where it is empty. We can study the non-emptiness of the core without explicitly solving the core equation. The following notation helps to simplify the dual of (1.8).

Definition 8 A superadditive TU game \mathcal{G} for a family \mathcal{F} of coalitions is totally balanced if, for any $\mathcal{A} \in \mathcal{F}$, the inequality

$$\sum_{\mathcal{A} \in \mathcal{F}} \mu_{\mathcal{A}} \cdot \nu(\mathcal{A}) \leq \nu(\mathcal{K}) \quad (1.10)$$

holds, where $\mu_{\mathcal{A}}$ is a collection of numbers in $[0, 1]$ (balanced collection of weights) such that

$$\sum_{\mathcal{A} \in \mathcal{F}} \mu_{\mathcal{A}} \cdot \mathbf{1}_{\mathcal{A}} = \mathbf{1}_{\mathcal{K}} \quad (1.11)$$

with $\mathbf{1}_{\mathcal{A}} \in \mathbb{R}^K$ denoting the characteristic vector whose elements are

$$(\mathbf{1}_{\mathcal{A}})[i] = \begin{cases} 1, & i \in \mathcal{A} \\ 0, & \text{otherwise} \end{cases} \quad (1.12)$$

The following pathbreaking result in the theory of TU games was independently gave by Bondareva [43] and L. Shapley [44].

Lemma 1 ([10]) A totally balanced TU game has a non-empty core set.

Where forming the grand coalition is not guaranteed, the following notation is applied.

Definition 9 *A (not necessarily superadditive) TU game \mathcal{G} for a family \mathcal{F} of coalitions is totally balanced if, for every balanced collection of weights $\mu_{\mathcal{A}}$, and for any $\mathcal{A} \in \mathcal{F}$,*

$$\sum_{\mathcal{A} \in \mathcal{F}} \mu_{\mathcal{A}} \cdot \nu(\mathcal{A}) \leq \max_{\psi \in \Psi} \sum_{\mathcal{A} \in \psi} \nu(\mathcal{A}) \quad (1.13)$$

So, if a TU game is totally balanced, then the core is non empty and therefore it is a convenient solution concept on the class of totally balanced TU games. There is an interesting relation between convex and balanced games.

Lemma 2 ([22]) *A convex game is totally balanced, but the converse is not necessarily true.*

The other key feature of coalitional convex games is

Lemma 3 (L. Shapley [36]) *The core set of a convex game is unique.*

Now, we illustrate an intuitive example of *power distribution* based on core set solution. This example is an extended form of the example established by [45, Ch. 12]. The network sketched in Fig. 1.1 wishes to allocate power among three players $\mathcal{K} = \{k_1, k_2, k_3\}$, according to their will to cooperate with each other. A power of 1 mW is provided to the network if three players decide to cooperate, or equivalently if the grand coalition will form. If only one player refuses to cooperate, a power of 0.8 mW will be assigned to the pair of cooperating nodes. The coalition game of Fig. 1.1 is defined by:

$$\nu(\mathcal{A}) = \begin{cases} 0 & \text{if } |\mathcal{A}| = 1; \\ 0.8 & \text{if } |\mathcal{A}| = 2; \\ 1 & \text{if } |\mathcal{A}| = 3. \end{cases} \quad (1.14)$$

The players of each coalition will cooperate with each other. The player of a singleton coalition will be isolated.

Each player receives a positive payoff if it decides to cooperate, whereas all players receive zero if no agreement is bound. To divide the total payoff (power) in some appropriate way, we rest on the core set definition. It is straightforward to show

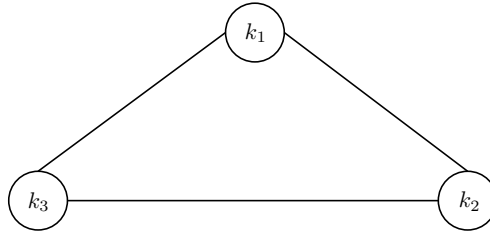


Fig. 1.1: The network allocates power among three players according to their will to cooperate with each other. A selfish player receives zero, a pair of cooperative players receive 0.8 mW, and the network supply 1 mW to the grand coalition.

that the coalitional TU game defined by (1.14) is superadditive. From Eqs. 1.3 and 1.4, it is easy to show that TU game (1.14) is not convex (supermodular). To check whether the core set of TU game (1.14) is empty or not, we resort to the balanced solution. TU game (1.14) is not balanced even though assigning the balanced weights as $\mu_{\mathcal{A}} = 1$ for singleton coalitions, and $\mu_{\mathcal{A}} = 0$ otherwise, inequality (1.10) holds. By using the fact that there exists other balanced collection of weights in which $\mu_{\mathcal{A}} = \frac{1}{2}$ for $|\mathcal{A}| = 2$, and $\mu_{\mathcal{A}} = 0$ otherwise, the game is not balanced, and its core set may be empty. Note that, this result *does not* mean that the core set of the game is *surely* empty.

Now, we heuristically find a core apportionment studying various possible networks. When there is no cooperation among players, the players are not provided with any power. That is, $\mathcal{F} = [\{k_1\}, \{k_2\}, \{k_3\}]$ with payoff distribution:

$$u_{k_1} = u_{k_2} = u_{k_3} = 0$$

If only one player decides to stay alone, the payoff 0.8 is equally divided between the two cooperative players and the isolated player gets zero. That is, for instance, $\mathcal{F} = [\{k_1, k_2\}, \{k_3\}]$ with payoff distribution:

$$\begin{cases} u_{k_3} = 0 \\ u_{k_1} = u_{k_2} = 0.4 \end{cases}$$

Now, we suppose a player, for example, k_2 decides to cooperate with both k_1 and k_3 , but the two players k_1 and k_3 do not bind an agreement to mutually cooperate. It is reasonable to suppose that the player k_2 can act as a relay between k_1 and k_3

and it must be provided more power. That is $\mathcal{F} = [\{k_1, k_2\}, \{k_2, k_3\}]$ with payoff distribution:

$$\begin{cases} u_{k_1} = u_{k_3} = 0.2 \\ u_{k_2} = 0.6 \end{cases}$$

Finally, in the complete network each player receives the same payoff. That is $\mathcal{F} = [\{k_1, k_2, k_3\}]$ with payoff distribution:

$$u_{k_1} = u_{k_2} = u_{k_3} = 1/3$$

As can be easily seen, the above argument satisfies feasibility and non-blocking conditions of the core set apportionment in Def. 6. It is worthwhile to note that the core set definition does not imply an even division of the whole payoff across players. Thus, it is clear that this game consists of multiple core sets. The power distribution problem can also be solved by cooperative game-theoretic bargaining solutions; e.g. Nash bargaining game and Auction [10].

1.3.1 On core stability

The goal of the network Fig. 1.1 is to allocate power among players in order to stimulate all of them to cooperate. Obviously, each player tries to get the highest possible payoff. Let us predict the behavior of the players after having known the definition of the game. Suppose that the players k_1 and k_2 find an opportunity to meet each other. Obviously, they quickly take advantage to cooperate and achieve payoff distribution $\mathbf{u} = [0.4, 0.4, 0]$. Then, it is profitable for player k_1 to invite player k_3 to join and therefore, improving its own payoff from 0.4 to 0.6 and that of player k_3 from zero to 0.2. On the other hand, this new agreement causes a decreasing payoff of player k_2 from 0.4 to 0.2, and now the players k_2 and k_3 have an incentive to cooperate and increase their proper payoff from 0.2 to 1/3. Note that this agreement makes the player k_1 's payoff decrease from 0.6 to 1/3. The unfavorable decision of player k_2 would tempt player k_1 to retaliate. A negotiation between k_1 and k_3 to release cooperation with k_2 results increasing their payoffs and boiling down k_2 's payoff to zero. The result of above argument concerns that the network is sustained by only one pair cooperation under the threat of: "If you cooperate with the third player, then I will do the same".¹ It is fairly clear that the players would seek to

¹Two is cooperation, three is a crowd.

cooperate only as pairs for the purpose of negotiation, and not cooperate in the grand coalition framework, even though the game is superadditive. This is due to fact of being superadditive but not balanced. The pairs can be changed as time goes on. In fact, the core apportionment suffers the lack of “farsighted” (i.e., long-term) stability.

A coalition structure based on core set, is not adequately farsighted to avoid the elusiveness of negotiation structure. At first sight, the core appears to be an extremely myopic notion, requiring the stability of a proposed allocation to deviations or blocks by coalitions, but not examining the stability of the deviations themselves. In general, the stability requirement is that the outcome be immune to deviations of a certain sort by coalitions. To provide the formal definition of farsighted stability, we need some additional notation.

Definition 10 ([46]) *For $\mathbf{u}, \mathbf{u}' \in \mathcal{I}(\mathcal{K}, \nu)$, \mathbf{u} indirectly dominates \mathbf{u}' , which is denoted by $\mathbf{u}' \ll \mathbf{u}$, if there exist a finite sequence of imputations $\mathbf{u}' = \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m = \mathbf{u}$ and a finite sequence of nonempty coalitions $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_m$ such that for each $j = 1, 2, \dots, m - 1$: i) by the deviation of \mathcal{A}_j , the imputation of \mathbf{u}_j is replaced to \mathbf{u}_{j+1} , and ii) $\mathbf{u}_j[k] < \mathbf{u}[k]$ for all $k \in \mathcal{A}_j$.*

Condition (i) says that each coalition \mathcal{A}_j has the power to replace imputation \mathbf{u}_j by imputation \mathbf{u}_{j+1} , and the condition (ii) says that each player in \mathcal{A}_j strictly prefers imputation \mathbf{u} to imputation \mathbf{u}_j . It is clear that the indirect dominance relation contains the direct dominance relation.

Definition 11 ([46, 47]) *Let $\mathcal{G} = (\mathcal{K}, \nu)$ be a TU game. A subset \mathcal{J} of $\mathcal{I}(\mathcal{K}, \nu)$ is a farsighted stable set if: i) for all $\mathbf{u}, \mathbf{u}' \in \mathcal{J}$, neither $\mathbf{u} \ll \mathbf{u}'$ nor $\mathbf{u}' \ll \mathbf{u}$, and ii) for all $\mathbf{u}' \in \mathcal{I}(\mathcal{K}, \nu) \setminus \mathcal{J}$ there exists $\mathbf{u} \in \mathcal{J}$ such that $\mathbf{u}' \ll \mathbf{u}$. Conditions i) and ii) are called internal stability and external stability, respectively.*

By internal stability, there is no imputation in \mathcal{J} that is dominated by another imputation in \mathcal{J} . By external stability, an imputation outside a stable set \mathcal{J} is unlikely to be attained. Let us introduce three other different payoff distribution concepts which capture foresight of the players.

1.4 Shapley value

The Shapley value is an alternative solution for the payoff distribution in TU games. The Shapley value has long been a central solution concept in coalitional game theory. It was introduced by L. S. Shapley in the seminal paper [48] and it was seen as a reasonable way of distributing the gains of cooperation, in a fair and unique way, among the players in the game. In the Shapley solution, those who contribute more to the groups that include them are paid more. Let us denote $\phi_k(\nu)$ as the Shapley value of player k in the TU game defined by ν . The surprising result due to Shapley is the following theorem.

Theorem 1 *There is a unique single-valued solution to TU games satisfying efficiency, symmetry, additivity and dummy. It is the well-known Shapley value, the function that assigns to each player k the payoff:*

$$\phi_k(\nu) = \sum_{\substack{\forall \mathcal{A} \subseteq \mathcal{K} \\ \text{s.t. } k \in \mathcal{A}}} \frac{(|\mathcal{A}| - 1)! \cdot (K - |\mathcal{A}|)!}{K!} (\nu(\mathcal{A}) - \nu(\mathcal{A} \setminus \{k\})) \quad (1.15)$$

The expression $\nu(\mathcal{A}) - \nu(\mathcal{A} \setminus \{k\})$ is the marginal payoff of player k to the coalition \mathcal{A} . The Shapley value can be interpreted as the expected marginal contribution made by a player to the value of a coalition, where the distribution of coalitions is such that any ordering of the players is equally likely. The Shapley value can be interpreted as follows [48]. Suppose the grand coalition of all players is being formed in a sequence by including one player at a time. In general, the marginal contribution of a particular player depends on its position in the sequence, which is determined randomly. There are $K!$ possible orderings of the players in total. Consider any coalition \mathcal{A} containing k and observe that the probability that player k enters the coalition to find precisely the players in $\mathcal{A} \setminus \{k\}$ already there is $\frac{(|\mathcal{A}| - 1)! \cdot (K - |\mathcal{A}|)!}{K!}$. That is, out of $K!$ permutations of \mathcal{K} , there are $(|\mathcal{A}| - 1)!$ different orders in which the players in $\mathcal{A} \setminus \{k\}$ can precede k and $(K - |\mathcal{A}|)!$ orders in which the remaining $\mathcal{K} \setminus \mathcal{A}$ can follow. When leaving a coalition \mathcal{A} , a player k receives the amount by which his exclusion changes (increase or decreases) the profits of the group, $\nu(\mathcal{A}) - \nu(\mathcal{A} \setminus \{k\})$. That makes the Shapley value exponentially hard to compute. Shapley characterized such value as the unique solution that satisfies the following four axioms:

1. *Efficiency:* The payoffs must add up to $\nu(\mathcal{K})$, which means that all the grand

coalition surplus is allocated. That is:

$$\sum_{k \in \mathcal{K}} \phi_k(\nu) = \nu(\mathcal{K})$$

In the absence of superadditivity, instead we use: $\max_{\psi \in \Psi} \sum_{\mathcal{A} \in \psi} \nu(\mathcal{A})$.

2. *Symmetry*: This axiom requires that the names of the players play no role in determining the value. If two players are substitutes because they contribute the same to each coalition, the solution should treat them equally. That is:

$$\nu(\mathcal{A} \cup \{k\}) = \nu(\mathcal{A} \cup \{i\}) \implies \phi_k(\nu) = \phi_i(\nu).$$

3. *Additivity*: The solution to the sum of two TU games must be the sum of what it awards to each of the two games. That is:

$$\phi_k(\nu + \omega) = \phi_k(\nu) + \phi_k(\omega) \quad \forall k \in \mathcal{K}.$$

4. *Dummy player*: The player k is dummy (null) if $\nu(\mathcal{A} \cup \{k\}) = \nu(\mathcal{A})$ for all \mathcal{A} not containing k . If a player k is dummy, the solution should pay it nothing; i.e. $\phi_k(\nu) = 0$.

The Shapley value is a feasible allocation, but need not be individually rational. Whenever the TU game is superadditive, the Shapley value is feasible and individually rational, but need not be in the core and hence can be directly dominated by another imputation. Reference [36] shows that the Shapley value of a supermodular TU-game is a core imputation, that is, the Shapley value is not dominated. For a superadditive TU game The Shapley value is an internal and external stable imputation, and for NTU games, it is formulated in [49, 50]. To make an example, let us calculate the Shapley value of the players in the power distribution game of Fig. 1.1:

$$\nu = \begin{cases} 0 & \{k_1\}, \{k_2\}, \{k_3\}; \\ 0.8 & \{k_1, k_2\}, \{k_1, k_3\}, \{k_2, k_3\}; \\ 1 & \{k_1, k_2, k_3\}. \end{cases} \implies$$

$$\begin{aligned} \phi_{k_1}(\nu) = \phi_{k_2}(\nu) = \phi_{k_3}(\nu) = \\ 0 + \frac{1! \cdot 1!}{3!} (0.8 - 0) + \frac{1! \cdot 1!}{3!} (0.8 - 0) + \frac{2! \cdot 0!}{3!} (1 - 0.8) = 1/3. \end{aligned}$$

Young in [51] defines an equivalent definition for Shapley value. He withdraws the additivity axiom, and instead, adds an axiom of marginality.

1. *Marginality*: If the marginal contribution to coalitions of a player in two games is the same, then the the award of the player must be the same. That is, if:

$$\nu(\mathcal{A}_i) - \nu(\mathcal{A}_i \setminus \{k\}) = \omega(\mathcal{A}_j) - \omega(\mathcal{A}_j \setminus \{k\}) \quad \forall \mathcal{A}_i \in \nu \text{ and } \forall \mathcal{A}_j \in \omega,$$

then $\phi_k(\nu) = \phi_k(\omega)$.

Marginality is an idea with a strong tradition in economic theory. In Young's definition, marginality is assumed and additivity is dropped. Young in [51] shows that the Shapley value is unique.

Theorem 2 (Young [51]) *There exists a unique single-valued solution to TU games satisfying efficiency, symmetry and marginality, and this solution is the Shapley value.*

In the network engineering literature, S. Kim in [52] proposes an energy efficient routing protocol based on the Shapley value. The concept of Shapley value is used by Khouzani *et al.* [53] to achieve a fair aggregate cost of link sharing, among primary and secondary users in a cognitive network. Using the Shapley value, a suitable network resource sharing among multimedia users is fairly achievable, as Park *et al.* propose in [54].

1.5 The kernel and nucleolus

Let $\mathcal{G} = (\mathcal{K}, \nu)$ be a coalitional game with transferable payoff. The *excess* of the coalition \mathcal{A} with respect to the payoff vector $\mathbf{u} \in \mathbb{R}^K$ is defined as

$$e(\mathcal{A}, \mathbf{u}) = \nu(\mathcal{A}) - \sum_{k \in \mathcal{A}} u_k \quad (1.16)$$

A positive excess can be interpreted as an incentive for a coalition to generate more utility. Using the excess notion, the core apportionment in a TU game can be redefined as:

$$\{ \mathbf{u} \in \mathbb{R}^K : e(\mathcal{K}, \mathbf{u}) = 0 \text{ , and } e(\mathcal{A}, \mathbf{u}) \leq 0 \quad \forall \mathcal{A} \subset \mathcal{K} \} \quad (1.17)$$

The *maximum excess* of player k against i is defined as

$$s_{ki}(\mathbf{u}) = \max \{ e(\mathcal{A}, \mathbf{u}) \mid \mathcal{A} \subset \mathcal{K}, k \in \mathcal{A}, i \in \mathcal{K} \setminus \mathcal{A} \} \quad (1.18)$$

If player k departs from \mathbf{u} , the most it can hope to gain (the least to lose) without the consent of player i is the amount of maximum excess. Extensions of the excess for NTU games is formalized in [55].

As defined by Osborne and Rubinstein [10, Ch. 14], a coalition \mathcal{A}_i is an *objection* of k against i to \mathbf{u} , if \mathcal{A}_i includes k but not i and $u_i > \nu(\{i\})$. Equivalently, \mathcal{A}_i is a coalition that contains k , excludes i and which gains too little. A coalition \mathcal{A}_j is a *counter-objection* to the objection \mathcal{A}_i of k against i , if \mathcal{A}_j includes i but not k and $e(\mathcal{A}_j, \mathbf{u}) \geq e(\mathcal{A}_i, \mathbf{u})$. Equivalently, \mathcal{A}_j is a coalition that contains i and excludes k and that gains even less. Objections and counter-objections are exchanged between members of the same coalition in \mathcal{A}_i .

The idea captured by the *kernel* is that if at a non empty imputation \mathbf{u} the maximum excess of player k against any other player i is less than the maximum excess of player i against the player k , then player k should get less. Of course, the players cannot get less than their individual worths if \mathbf{u} is an imputation. The definition of the kernel follows:

Definition 12 *The kernel is the set of all imputations \mathbf{u} with the property that for every objection \mathcal{A}_i of any player k against any other player i to \mathbf{u} there is a counter-objection of i to \mathcal{A}_i such that:*

- a) $s_{ki}(\mathbf{u}) = s_{ik}(\mathbf{u});$ *or*
- b) $s_{ki}(\mathbf{u}) < s_{ik}(\mathbf{u})$ *and* $u_k = \nu(\{k\});$ *or*
- c) $s_{ki}(\mathbf{u}) > s_{ik}(\mathbf{u})$ *and* $u_i = \nu(\{i\}).$

The kernel is the set of imputations \mathbf{u} such that for any coalition \mathcal{A}_i , for each objection \mathcal{A}_j of a user $k \in \mathcal{A}_i$ over any other member $i \in \mathcal{A}_i$, there is a counter-objection of i to \mathcal{A}_j . The kernel is contained in the (nonempty) core in any assignment game ν [56, Th. 1]. In Fig. 1.1, the unique kernel element is the equal split $\mathbf{u} = [1/3, 1/3, 1/3]$, otherwise for the single player coalition objection of the player with the minimum payoff, there is no any counter-objection.

The last type of a stable imputation we will study is the *nucleolus*. With the nucleolus no confusion regarding the player set can arise. The basic motivation behind the nucleolus is that one can provide an allocation that minimizes the excess of the coalitions in a given coalitional game $\mathcal{G} = (\mathcal{K}, \nu)$. For a TU game $\mathcal{G} = (\mathcal{K}, \nu)$ and the

payoff vector $\mathbf{u} \in \mathbb{R}^K$, let us denote $\mathbf{E}(\mathbf{u}) = [\cdots \geq e(\mathcal{A}, \mathbf{u}) \geq \cdots : \emptyset \neq \mathcal{A} \neq \mathcal{K}]$ as a $2^K - 2$ dimensional vector whose components are the values of the excess function for all $\mathcal{A} \subset \mathcal{K}$, arranged in a non-increasing order. The nucleolus of a game is the imputation which minimizes the excess with respect to the lexicographic order² over the set of imputations. The nucleolus of \mathcal{G} with respect to $\mathcal{I}(\mathcal{K}, \nu)$ is given by:

$$\{\mathbf{u} \in \mathcal{I}(\mathcal{K}, \nu) \mid \mathbf{E}(\mathbf{u}) \preceq_{lex} \mathbf{E}(\mathbf{u}') \quad \forall \mathbf{u}' \in \mathcal{I}(\mathcal{K}, \nu)\} \quad (1.19)$$

The definition of the nucleolus of a coalitional game in characteristic function form entails comparisons between vectors of exponential length. Thus, if one attempts to compute the nucleolus by simply following its definition, it would take an exponential time. In the network engineering literature, Han and Poor in [57] apply the Shapley value, excess and nucleolus solutions to study a possible cooperative transmission among intermediate nodes to help relay the information of wireless users.

This defining property makes the nucleolus appealing as a fair single-valued solution. It is easy to see that, whenever the core of a game is nonempty, the nucleolus lies in it [22]. Moreover, the nucleolus always belongs to the kernel and satisfies the symmetry and dummy axioms of Shapley: dummy players receive zero payoffs. If a null player is removed from the game, the payoff allocation of the remaining players is uninfluenced by its departure. Because of these desirable properties, the nucleolus solution has found a lot of applications in cost sharing and resource allocation as Maschler in [12] reports. However, the nucleolus possesses certain features that makes it less agreeable. The original definition treats the excesses of any two coalitions as equally important, regardless of coalition sizes and coalition composition. Some unappealing features of utility distribution, derived with the nucleolus are listed in [51]. For instance, the nucleolus lacks many monotonicity properties. That is, if a game changes so that some player's contribution to all coalitions increases then the player's allocation should not decrease. Monotonicity states that as the underlying data of game change, the utility must change in a parallel fashion.

²The lexicographic order between two vectors \mathbf{u} and \mathbf{u}' is defined by $\mathbf{u} \preceq_{lex} \mathbf{u}'$, if there exists an index k such that $\mathbf{u}[l] = \mathbf{u}'[l]$ for all $l < k$, and $\mathbf{u}[k] < \mathbf{u}'[k]$.

1.6 Cooperative Nash equilibria

Coalitional games aim at identifying the best coalitions of the agents and a fair distribution of the payoff among the agents. The classic core solution is an extension of the Nash equilibrium, since the coalitions bind agreements of agents with each other and earns a vector value rather than a real number. In [58, Sec. 7.6] it is shown that the core set of an underlying coalitional game, if it exists, asymptotically coincides with the set of Nash equilibria of the repeated game, in the long run. The result of the Nash equilibrium is not always a satisfactory outcome for an external observer (e.g., prisoner's dilemma game). R. Aumann³ in [59] and Bernheim *et al.* in [60] introduce a stronger notion of Nash equilibria based on coalitional game theory. First, let us review the definition of the Nash equilibrium where each pure strategy in a static game is presented as a coalition in a coalitional game. Thus, each player belongs to only one coalition.

Definition 13 *A pure strategy (coalition) combination $\psi = [\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_m]$ wherein $\mathcal{A}_i \cap \mathcal{A}_{j \neq i} = \emptyset$, $\bigcup_{i=1}^m \mathcal{A}_i = \mathcal{K}$ and a payoff distribution $\mathbf{u} = [u_1, \dots, u_K]$ is a pure Nash equilibrium if there does not exist a player $k \in \mathcal{K}$ whose unilateral deviation to a different coalition (pure strategy) yields a new distribution $\mathbf{u}' = [u'_1, \dots, u'_K]$ such that $u'_k > u_k$.*

In other words, in a Nash equilibrium no agent is motivated to deviate from its coalition (strategy) given that the others do not deviate. As an example, we study the forwarder's dilemma game [14] presented in Fig. 1.2. This game is intended to represent a basic wireless relay operation between two different wireless terminals. These two agents, represented by players k_1 and k_2 , are supposed to operate a direct link that enables them to communicate without intermediaries. Each players wants to send a packet to its destination, d_1 and d_2 respectively, in each time step using the other player as a forwarder. We assume that each forwarding has a energy cost $0 < c \ll 1$. If player k_1 forwards (F) the player's k_2 packet, player k_2 gets a reward 1 and vice versa. Each player's utility is its reward minus the cost. Each player is allured to drop (D) the received packet for saving energy. The strategic form of this game is depicted in Tab. 1.1. In the cooperative representation of the forwarder's dilemma

³Robert Aumann has received in 2005 the Nobel prize in economy for his contributions to game theory, together with Thomas Schelling.

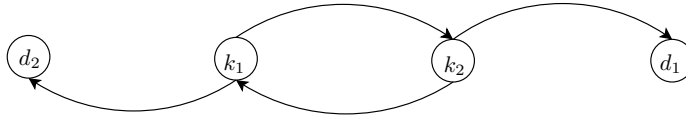


Fig. 1.2: *The network scenario of the forwarder's dilemma game.*

		k_2	
		F	D
	F	$1 - c, 1 - c$	$-c, 1$
k_1	D	$1, -c$	$0, 0$

$\mathbf{u} = [u_{k_1}, u_{k_2}]$

Tab. 1.1: *The strategic form in the forwarder's dilemma game. In each cell, the first value is the payoff of player k_1 , whereas the second is that of k_2 .*

game there are two coalitions $\psi = [\mathcal{A}_F, \mathcal{A}_D]$ and each player in $\mathcal{K} = \{k_1, k_2\}$ must choose one coalition. For instance, $\psi = [\mathcal{A}_F = \{k_1, k_2\}, \mathcal{A}_D = \emptyset]$ is equivalent to the strategy profile (F, F) and $\psi = [\mathcal{A}_F = \{k_2\}, \mathcal{A}_D = \{k_1\}]$ corresponds to the strategy profile (D, F) , and so on.

Unilateral deviation of player k_1 from $\psi = [\mathcal{A}_F = \{k_1, k_2\}, \mathcal{A}_D = \emptyset]$ to the profile $\psi = [\mathcal{A}_F = \{k_2\}, \mathcal{A}_D = \{k_1\}]$ increases its own payoff, and therefore the pure strategy profile (F, F) is not a Nash equilibrium point. The same applies to the departure of player k_2 from $\psi = [\mathcal{A}_F = \{k_1, k_2\}, \mathcal{A}_D = \emptyset]$ to the pure strategy $\psi = [\mathcal{A}_F = \{k_1\}, \mathcal{A}_D = \{k_2\}]$. We can easily check the different combinations of $\psi = [\mathcal{A}_F = \{k_1\}, \mathcal{A}_D = \{k_2\}]$, $\psi = [\mathcal{A}_F = \{k_2\}, \mathcal{A}_D = \{k_1\}]$, and finally $\psi = [\mathcal{A}_F = \emptyset, \mathcal{A}_D = \{k_1, k_2\}]$. The unilateral move of user k_1 (resp. k_2) from the strategy profile $\psi = [\mathcal{A}_F = \emptyset, \mathcal{A}_D = \{k_1, k_2\}]$ to $\psi = [\mathcal{A}_F = \{k_1\}, \mathcal{A}_D = \{k_2\}]$ (resp. to $\psi = [\mathcal{A}_F = \{k_2\}, \mathcal{A}_D = \{k_1\}]$), does not yield any benefit. This game has a unique Nash equilibrium at the pure joint strategy $\psi = [\mathcal{A}_F = \emptyset, \mathcal{A}_D = \{k_1, k_2\}]$ with unsatisfactory payoff distribution $\mathbf{u} = [0, 0]$. At the Nash equilibrium point either players

choose the “competitive” and “egoistic” strategy D .

In many games, there are opportunities for joint deviations that are mutually beneficial for a subset of players. This led Aumann [59] to propose the idea of *strong Nash equilibrium* which ensures a more restrictive stability than the conventional Nash equilibrium. Strong Nash equilibrium reflects the unprofitability of coalition deviations. It is a strategy profile that is stable against deviations not only by single players but by all coalitions of players. A strong equilibrium is defined as a strategic profile for which no subset of players has a joint deviation that strictly benefits all of them, while all other players (in the subset) are expected to maintain their equilibrium strategies.

Definition 14 *A strategy (coalition) combination $\psi = [\mathcal{A}_1, \dots, \mathcal{A}_m]$ where $\mathcal{A}_i \cap \mathcal{A}_{j \neq i} = \emptyset$ and $\bigcup_{i=1}^m \mathcal{A}_i = \mathcal{K}$ with payoff distribution $\mathbf{u} = [u_1, \dots, u_K]$ is a strong Nash equilibrium if there do not exist a coalition $\mathcal{A}_i \in \psi$ whose deviation yields a new distribution $\mathbf{u}' = [u'_1, \dots, u'_K]$ such that $u'_k \geq u_k \quad \forall k \in \mathcal{A}_i$ and $\exists k \in \mathcal{A}_i$ such that $u'_k > u_k$.*

This definition of strong equilibrium is actually slightly different from those of [59] and [60]. Def. 14 allows a coalition to deviate from a strategy profile that strictly increases the payoffs of some of its members without decreasing those of the other members, whereas the original definition allows only deviations that strictly increase the payoffs of all members of a deviating coalition. We note that if a game implements a strategy for strong equilibrium, it does not necessarily implement it for Nash equilibrium. Both interpretations of strong Nash equilibrium are prominent in the literature, and in most games the two definitions lead to the same sets of strong Nash equilibria; however, the one that we use here is slightly more appealing in the context of network-formation games (see, e.g., [61]). Network formation games involve a number of independent players that interact with each other in order to form a suited graph that connects them.

Now, we restudy the forwarder’s dilemma game and try to find strong Nash equilibria profile. We will show that the game possesses strong Nash equilibria which are not equivalent to the Nash equilibrium. We pick different coalition combination and test whether there exist any coalition whose deviation satisfies its own members or not.

1. $\psi = [\mathcal{A}_F = \{k_1\}, \mathcal{A}_D = \{k_2\}]$ is not strong Nash equilibrium because the deviation of \mathcal{A}_F increases its member's payoff.
2. $\psi = [\mathcal{A}_F = \{k_2\}, \mathcal{A}_D = \{k_1\}]$ is not strong Nash equilibrium because the deviation of \mathcal{A}_F renders its member's payoff higher.
3. $\psi = [\mathcal{A}_F = \emptyset, \mathcal{A}_D = \{k_1, k_2\}]$ is not strong Nash equilibrium because the deviation of both players from \mathcal{A}_D to \mathcal{A}_F increases payoff distribution.
4. $\psi = [\mathcal{A}_F = \{k_1, k_2\}, \mathcal{A}_D = \emptyset]$ is strong Nash equilibrium because the departure of one or both players from \mathcal{A}_F to \mathcal{A}_D decreases at least one player's payoff.

The unique strong Nash equilibrium is the strategy profile (F, F) which corresponds to coalition set of $\psi = [\mathcal{A}_F = \{k_1, k_2\}, \mathcal{A}_D = \emptyset]$, since no deviation can better off the payoff distribution vector $\mathbf{u} = [1 - c, 1 - c]$. In fact, at the strong Nash equilibrium, both players choose the “cooperative” and “altruistic” strategy of F in spite of the energy transmission cost.

In network problems, Zhong *et al.* show that using strong Nash equilibria context makes possible a collusion-resistant routing in non-cooperative wireless ad hoc networks [62]. Altman *et al.* in [63] examine a dynamic random access game with orthogonal power constraints in which the probability of transmission of a terminal in each slot depends on the amount of energy left prior to that slot. They show the existence of a strong Nash equilibrium point.

Conventional Nash equilibrium is concerned with the possibilities of only one step deviation by any player. The notion of strong Nash equilibrium requires an agreement not be subject to an improving (one step) deviation by any coalition of players given that all others coalitions be inert. This notion is stronger than Nash equilibrium, but it is not resistant to further deviation by sub-coalitions (the subsets of a coalition). Recognizing this problem, Bernheim *et al.* [60] introduced the notion of *coalition-proof Nash equilibrium*, which requires only that an agreement be immune to improving deviations which are *self-enforcing*. The definition of a self-enforcing deviation is recursive.

Definition 15 *For a singleton coalition, a deviation is self-enforcing if it maximizes the player's payoff. For a coalition of more than one player, a deviation is self-enforcing if: i) it is profitable for all its members, and ii) if there is no further self-enforcing and improving deviation available to a proper sub-coalition of players.*

Generally, a deviation by a coalition is self-enforcing if no sub-coalition has an incentive to initiate a new deviation. In the forwarder's dilemma game, the Nash equilibria is upset by a deviation of the coalition of both players k_1 and k_2 . At the pure strategy Nash equilibrium where each player choose strategy D , they each obtain a payoff of 0. By jointly deviating (both choosing F instead) k_1 and k_2 each earn a payoff $1 - c$. This deviation is not self-enforcing even though the movement to the pure strategy $\psi = [\mathcal{A}_F = \{k_1, k_2\}, \mathcal{A}_D = \emptyset]$ is profitable for both players. At strong Nash pure strategy (F, F) , the player k_1 tempts to move to strategy (D, F) to get more payoff, and player k_2 to that (F, D) . Thus, the strong Nash equilibrium is not immune against self-enforceability.

This notion of self-enforceability provides a useful means of distinguishing coalitional deviations that are viable from those that are not resistant to further deviations. With the concept of self-enforceability, our notion of coalition-proofness is easily formulated.

Definition 16 *In a one player game, a strategy is a coalition-proof Nash equilibrium if it maximizes the player utility. In a game with more than one player, a combination strategy is coalition-proof Nash equilibrium, if no sub-coalition has a self-enforcing deviation that makes all its members better off.*

This solution concept requires that there is no sub-coalition that can make a mutually beneficial deviation (keeping the strategies of non-members fixed) in a way that the deviation itself is stable according to the same criterion. In the forwarder's dilemma game, the strong Nash equilibrium profile (F, F) is not equivalent to coalition-proof Nash equilibrium. This is due to the fact that, the deviation of $\{k_1\} \subset \mathcal{A}_F = \{k_1, k_2\}$ to the strategy (D, F) increases payoff of k_1 . In this game there does not exist any coalition-proof Nash equilibrium, due to the fact that all pure strategies have at least one self-enforcing deviation.

Bernheim *et al.* [60] note that for 2-person games the set of coalition-proof equilibria coincides with the set of Nash equilibria that are not Pareto dominated by any other Nash equilibrium. However in n -person games ($K \geq 3$) the equilibrium concepts are independent. At coalition-proof Nash equilibrium, the deviations are restricted to be stable themselves against further deviations by sub-coalitions. Moldovanu in [64] discusses the situations of a 3-player game, wherein coalition-proof Nash equilibrium is equivalent to the core set. The conditions under which the set of coalition-proof

Nash equilibria coincides with the set of strong Nash equilibria, are formulated by H. Konishi *et al.* in [65].

In the network engineering literature, F3legyh3azi *et al.* in [66] apply the concept of coalition-proof Nash equilibria to achieve a stable and fair channel allocation solution in a competitive multi-radio multi-channel wireless cognitive network. Gao *et al.* investigate multi-radio multi-channel allocation in multi-hop ad-hoc networks [67]. To better understand the concepts of self-enforceability and coalition-proof Nash equilibrium, let us introduce an intuitive *subcarrier allocation game in an OFDMA network*. Let us focus on three wireless transmitters $\mathcal{K} = \{k_1, k_2, k_3\}$ and an OFDMA base-station with two subcarriers $\mathcal{N} = \{1, 2\}$. Every subcarrier $n \in \mathcal{N}$ has a frequency spacing Δf . Each user $k \in \mathcal{K}$ experiences a Gaussian complex-valued channel gain $|H_{kn}|^2$ on the n th subcarrier to the base station. We assume that each subcarrier can be shared among more than one transmitter. The payoff of each player (transmitter) is defined as the achieved Shannon channel capacity. Each user $k \in \mathcal{K}$ is allowed to either spend a certain power \bar{p}_k on only one chosen subcarrier, or equally divide it among both subcarriers. In the pure strategy a_1 , player k transmits with the maximum power \bar{p}_k on subcarrier $n = 1$ and does not transmit any information on subcarrier $n = 2$. The strategy a_2 is contrary to a_1 , i.e. exclusively transmitting on subcarrier $n = 2$ with maximum power. Finally strategy a_3 equally divides its power on two subcarriers and exploits transmitting on both tones. The terminal k achieves a channel capacity:

$$C_k = \sum_{n \in \mathcal{N}} C_{kn} \quad (1.20)$$

where C_{kn} is the Shannon capacity achieved by user k on the n th subcarrier:

$$C_{kn} = \Delta f \cdot \log_2 \left(1 + \frac{|H_{kn}|^2 p_{kn}}{\sum_{k \neq i \in \mathcal{K}} |H_{in}|^2 p_{in} + \sigma_w^2} \right) \quad (1.21)$$

wherin p_{kn} represents the power allocated by terminal k over the n th subcarrier and where the interference term $\sum_{k \neq i \in \mathcal{K}} |H_{in}|^2 p_{in}$ is approximated with a Gaussian random variable of equal mean and variance. Choosing the strategy a_1 means selecting $p_{k1} = \bar{p}_k$ and $p_{k2} = 0$. For the strategy a_2 , $p_{k1} = 0$ and $p_{k2} = \bar{p}_k$, and for strategy a_3 , $p_{k1} = p_{k2} = \frac{\bar{p}_k}{2}$. The parameter σ_w^2 is the power of the additive white Gaussian noise (AWGN). Note that, in an OFDMA system, there is no interference between adjacent subcarriers. Hence, C_{kn} considers only intra-subcarrier noise, that occurs when the same subcarrier is shared by more terminals.

		$k_3(a_1)$			$k_3(a_2)$			$k_3(a_3)$		
		k_2			k_2			k_2		
		a_1	a_2	a_3	a_1	a_2	a_3	a_1	a_2	a_3
k_1	a_1	8, 5, 6	11, 11, 7	9, 11, 6	12, 7, 10	15, 10, 10	11, 10, 9	9, 5, 10	12, 9, 10	9, 10, 9
k_1	a_2	12, 8, 6	8, 7, 10	9, 11, 8	11, 11, 8	7, 6, 6	8, 11, 7	13, 9, 11	8, 6, 11	8, 10, 10
k_1	a_3	15, 6, 6	14, 8, 8	13, 10, 7	14, 8, 9	15, 6, 7	13, 9, 7	14, 6, 10	14, 7, 10	12, 9, 10

$\mathbf{u} = [u_{k_1}, u_{k_2}, u_{k_3}]$

Tab. 1.2: Subcarrier allocation in OFDMA network game in strategic form. The three strategies for three players k_1 , k_2 and k_3 are: Transmitting with the maximum power only on subcarrier number 1 (a_1), transmitting with the maximum power only on subcarrier number 2 (a_2) and equal division of the maximum power among both subcarriers (a_3). The player's payoff is the achieved channel capacity in kb/s.

Tab. 1.2 reports the simulation results obtained after 100 random realizations of a network with terminals distributed at a distance between 3m and 50m from the base-station. In the pure strategy matrix form of Tab. 1.2, player k_1 chooses the row, player k_2 chooses the column, and player k_3 chooses the matrix. Each payoff reports the (rounded) value of the achieved Shannon channel capacity in kb/s. We consider the following parameters for our simulations: the maximum power of each terminal k is $\bar{p}_k = 10$ mW; the power of the ambient AWGN noise on each subcarrier is $\sigma_w^2 = 100$ pW, and finally the carrier spacing is $\Delta f = \frac{10}{1024}$ MHz.⁴ The path coefficients $|H_{kn}|^2$, corresponding to the frequency response of the multipath wireless channel, are computed using the 24-tap ITU modified vehicular-B channel model adopted by the IEEE 802.16m standard [68].

It is easy to show that the (pure) Nash equilibrium strategies of Tab. 1.2 are (a_3, a_3, a_3) equivalent to $\psi = [\mathcal{A}_{a_1} = \emptyset, \mathcal{A}_{a_2} = \emptyset, \mathcal{A}_{a_3} = \mathcal{K}]$ and (a_1, a_2, a_2) to $\psi = [\mathcal{A}_{a_1} = \{k_1\}, \mathcal{A}_{a_2} = \{k_2, k_3\}, \mathcal{A}_{a_3} = \emptyset]$. The Nash equilibrium strategy (a_3, a_3, a_3) is neither coalition-proof nor strong. With deviation of the coalition \mathcal{A}_{a_3} to the strategy profile (a_2, a_1, a_3) all players profit more with payoff distribution [13, 9, 11]. This change is no longer valid since, there exists a self-enforceability for player k_1 to transit to the strategy profile (a_3, a_1, a_3) . This transition is not favorable for players

⁴This is the carrier spacing of each subcarrier at a base station with 10 MHz bandwidth and 1024 subcarriers.

k_2 and k_3 . The player k_2 is tempted to transit to the Nash equilibrium point to earn a higher payoff. Whereas, the Nash equilibrium strategy profile (a_1, a_2, a_2) with payoff vector $[15, 10, 10]$ is a strong and coalition-proof Nash equilibrium. This is due to the fact that, in $\psi = [\mathcal{A}_{a_1} = \{k_1\}, \mathcal{A}_{a_2} = \{k_2, k_3\}, \mathcal{A}_{a_3} = \emptyset]$ there is no deviation and self-enforceability that can improve the payoff distribution. As can be seen, all players prefer to stay at the coalition-proof Nash equilibrium rather than the pure Nash equilibrium strategy (a_3, a_3, a_3) . Note that, a strong or coalition-proof Nash equilibrium does not necessarily coincide with a Nash equilibrium strategy profile, and the result of Tab. 1.2 is an exception.

In general, the existence of a pure cooperative or non-cooperative Nash equilibrium for subcarrier allocation game in OFDMA network is not guaranteed. Given different parameters approaches to quite different channel capacities and this may results a matrix form without any type of Nash equilibrium. There even might exist a Nash equilibrium which is Pareto-dominated by another strategy profile. This shows that in OFDMA networks, an appropriate resource allocation technique is needed [26].

1.7 Coordinated equilibrium

The most common solution concept in (non-cooperative) game theory, Nash equilibrium, assumes that players take mixed actions independently of each other. Cooperative games allow players to coordinate each other to find out possible equilibria and (joint) optimizations that the players can perform on their own. Unlike evolutionary games [10, Ch. 3], in coordinated games the interaction between players is implemented once, among all players by a central authority, to increase their throughput. The notion of *correlated equilibrium* was introduced by R. Aumann [69]. Correlated equilibria are defined in a context where there is an intermediary who sends random (private or public) signals to the players. An intermediary needs not have any intelligence or knowledge of the game. These signals allow players to coordinate their actions, and, in particular, to perform joint randomization over strategies. “Correlated strategies are familiar from cooperative game theory, but their applications in non-cooperative games are less understood”, says R. Auman [69]. This is because, the players of a coordination game, are not totally isolated and without a communication between them, achieving to coordinated strategy profile is not possible.

		k_2	
		W	A
k_1	W	0, 0	0, $1 - c$
	A	$1 - c, 0$	$-c, -c$

$\mathbf{u} = [u_{k_1}, u_{k_2}]$

Tab. 1.3: *The multiple access game in strategic form. The two moves for each player are: access (A) or wait (W).*

Let us start with an intuitive example. Consider the multiple access game [14, Table III] described in Tab. 1.3. The players k_1 and k_2 wish to send some packets to their receivers sharing a common resource, i.e., the wireless medium. They are in the sight of each other and accordingly, they interfere if transmitting at the same time. The users have two possible pure strategies: access (A) and wait (W). In this game two identical transmitters must simultaneously decide whether to access to channel or wait. The transmission of each packet has an energy cost of $0 < c \ll 1$. Each player earns a payoff 1 if it succeeds to transmit its packet without collision with the other. Waiting does not bring neither cost nor reward for the player. Each player's utility is its reward minus the cost. This game has three Nash equilibria: (A, W), (W, A) and a mixed strategy Nash equilibrium where each player transmits with the probability $1 - c$ [14, Sec. 2.3, 2.4]. The utilities of Nash equilibria strategies are: $(1 - c, 0)$, $(0, 1 - c)$ and $(0, 0)$, respectively. It is clear that the mixed strategy is not resistant to an improving deviation. In the following, we give the possibility of preplay communication to achieve a stable Nash equilibria.

In the game with “cheap conversation”, each player simultaneously and publicly announces whether it decides to access or wait. Following the announcements each player makes its choice. Suppose the players agree to participate to the game binding the following agreement: each player announces A with probability $\frac{3}{4}$. If the profile of announcements is either (A, W) or (W, A), then each player plays its own announcement. Otherwise, each player plays A with probability $\frac{1}{2}$. Note that no further

		k_2	
		W	A
k_1	W	$\frac{10+32c}{64}$	$\frac{11-16c}{32}$
	A	$\frac{11-16c}{32}$	$\frac{10+32c}{64}$

$$\mathbf{u} = [u_{k_1} = u_{k_2}]$$

Tab. 1.4: *The strategic form matrix of the multiple access game with preplay agreement.*

communication is possible. The use of joint deviation requires the unanimity of all members of the deviating coalition. A player agrees to be a part of a joint deviation if given its own information the deviation is profitable. Thus, if a joint deviation is used, it is common knowledge that each deviator believes that deviation is profitable.

This tradeoff results in an expected payoff for each player of $\frac{11-16c}{32} > 0$, while in the mixed Nash equilibrium of the original game each player has an expected payoff of 0. In this coordinated Nash equilibrium of the game, the players effectively play the *correlated strategy* [69, 70] (of the original game) given in Tab. 1.4, in order to face a higher utility in strategy profiles (A, W) and (W, A) . It is important to note that, this joint probability distribution is not the product of its marginal distributions and therefore cannot be achieved from a mixed strategy profile of the game without correlation among players.

As can be seen, the proposed correlated deviation from the mixed strategy equilibrium makes both players better off. Note that the players are allowed to bind an agreement only on the space of feasible outcomes. In the correlated multiple access game the outcome is feasible since the correlated results are in the range between the smallest and highest possible payoff. In fact, the set of correlated equilibria contains those equilibria from which no coalition has a self-enforcing deviation making all members better off.

Let us describe a more complicated correlated equilibrium. We study the near-far effect game established by G. Bacci *et al.* in [71, Fig. 6]. The basic idea of near-

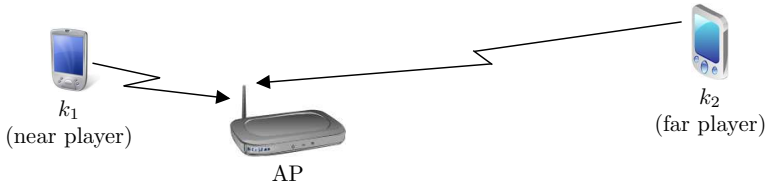


Fig. 1.3: *The network scenario in the near-far effect game.*

far effect game scheme is depicted in Fig. 1.3. Two wireless terminals k_1 and k_2 , are placed close to and far from a certain access point (AP), respectively, in a code division multiple access (CDMA) network with high SINR regime. The strategy of each player is either to transmit with the maximum power \bar{p} , or with a weakened level $\eta\bar{p}$, where $0 < \eta < 1$. Due to the interference at the AP, the throughput (the amount of delivered information) of each player depends on the strategies chosen by both players. Transmitting with a higher power increases the BER, and this results decreasing the throughput. Each player is rewarded u if it successfully delivers its packet and a reduced δu , if it delivers a corrupted version of the packet, where $0 < \eta < \delta < 1$. If the near player k_1 decides to transmit with the power \bar{p} , the farther player k_2 will not be able to deliver any information to the AP.

This results in no benefit for k_2 and causes a power consumption cost equal to $-\eta c$ if k_2 chooses strategy $\eta\bar{p}$ and $-c$ otherwise, where $c \ll u$. Obviously, transmitting with power \bar{p} for k_1 , results in a complete information delivery. This concerns a payoff equal to reward minus power consumption cost, i.e. $u - c$, irrespective of the k_2 strategy. The packets of player k_2 are successfully delivered if it chooses the maximum power \bar{p} , and player k_1 that reduced $\eta\bar{p}$. On the other hand, if both players decide to transmit with reduced power $\eta\bar{p}$, the near player takes the payoff $\delta u - \eta c > 0$, whilst the farther player k_2 will not successfully deliver any packet and suffers only a power cost $-\eta c$.

The payoff matrix of the near-far effect game is depicted in Tab. 1.5. As can be seen, the unique pure strategy of this game is represented by the strategy $(\bar{p}, \eta\bar{p})$ with benefits $u - c$ and $-\eta c$ for k_1 and k_2 , respectively. This means that, at the Nash equilibrium point, the farther player is not able to send any information. On the other hand, the Pareto optimal⁵ solution of the game are the strategies $(\bar{p}, \eta\bar{p})$ and

⁵Pareto optimal is named after Vilfredo Pareto (1848-1923). He was an Italian engineer and

		k_2	
		$\eta\bar{p}$	\bar{p}
k_1	$\eta\bar{p}$	$\delta u - \eta c, -\eta c$	$\delta u - \eta c, \delta u - c$
	\bar{p}	$u - c, -\eta c$	$u - c, -c$

$$\mathbf{u} = [u_{k_1}, u_{k_2}]$$

Tab. 1.5: *Payoff matrix for the near-far effect game with power control and variable throughput.*

$(\eta\bar{p}, \bar{p})$. This is an unsatisfactory outcome for the far player k_2 , while the near player k_1 takes the highest possible payoff. Now, let us find the mixed strategy of the game. We denote α_1 the probability with which the near player k_1 decides to transmit with the maximum power \bar{p} and α_2 the same probability for the far player k_2 . The payoffs of the players k_1 and k_2 are represented by:

$$u_{k_1} = \alpha_1 \left((1 - \delta) u - (1 - \eta) c \right) + (\delta u - \eta c) \tag{1.22a}$$

$$u_{k_2} = \alpha_2 \left((1 - \alpha_1) \delta u - (1 - \eta) c \right) - \eta c \tag{1.22b}$$

Both players want to maximize their own payoff. As can be seen, u_{k_1} takes its maximum value $u - c$ with $\alpha_1 = 1$. On the other hand, with $\alpha_1 = 1$, the far player k_2 earns a negative payoff whatever $\alpha_2 \in [0, 1]$. Instead, with $\alpha_1 = 0$ the near player k_1 gains $\delta u - \eta c$, and the player k_2 setting up $\alpha_2 = 1$ achieves the payoff $\delta u - c$. Thus, the best values for α_1 and α_2 are 0 and 1, respectively. The conclusion is that the mixed strategy is equivalent to the pure strategy $(\eta\bar{p}, \bar{p})$ with payoff $\mathbf{u} = [\delta u - \eta c, \delta u - c]$. In this game there is no (totally) mixed strategy and that is equal to the one of the pure Pareto optimal points.

The near player earns the highest possible payoff at the Nash equilibrium, hence, it does not leave this strategy profile. The highest possible payoff for the far player is on the contrary $\delta u - c$. We show that an appropriate agreement among players can

economist and he made several important contributions to economics.

		k_2	
		$\eta\bar{p}$	\bar{p}
k_1	$\eta\bar{p}$	0, 0	0, 0
	\bar{p}	$\kappa, \kappa \frac{(\delta u - c)}{(u - c)}$	$1 - \kappa, (1 - \kappa) \frac{(\delta u - c)}{(u - c)}$

$\mathbf{u} = [u_{k_1}, u_{k_2}]$ normalized to 1

Tab. 1.6: The strategic form matrix of the near-far effect game with preplay agreement, and with $\kappa = \frac{\delta u}{(1 - \eta)c}$.

satisfy both of them at correlated equilibrium. Players k_1 and k_2 can guarantee an expected payoff of $\mathbf{u} = [u - c, \delta u - c]$ by playing the correlated strategy profile:

$$\frac{\delta u}{(1 - \eta)c} \cdot (\bar{p}, \eta\bar{p}) + \left(1 - \frac{\delta u}{(1 - \eta)c}\right) \cdot (\bar{p}, \bar{p}) \quad (1.23)$$

This is a plausible end, since both players earn their own highest possible payoff. The correlated strategy (1.23) is derived from the fact that, picking any real number κ in the expression $\kappa \cdot (\bar{p}, \eta\bar{p}) + (1 - \kappa) \cdot (\bar{p}, \bar{p})$ is indifferent for the near player k_1 , since it gets its own highest possible payoff, $u - c$ as well. To satisfy the far player k_2 , it is enough to solve the following equation for u_{k_2} :

$$\kappa \cdot (\bar{p}, \eta\bar{p}) + (1 - \kappa) \cdot (\bar{p}, \bar{p}) = [u - c, \delta u - c] \quad (1.24)$$

Supposing $\kappa = \frac{\delta u}{(1 - \eta)c} < 1$, the correlated strategy (1.23) means that the near player always transmits at its highest power level \bar{p} , and the far player transmits at that reduced $\eta\bar{p}$ with probability $\frac{\delta u}{(1 - \eta)c}$, and the maximum power \bar{p} otherwise. Actually, the near and far players effectively play the matrix form game of Tab. 1.6.

Bonneau *et al.* in [72] show that the coordination among mobile users can significantly increase the performance of access to a common channel in ALOHA setting. A coordination mechanism is also considered by Bonneau *et al.* in [73] to achieve the optimal power allocation in a wireless network wherein each terminal knows only

its own channel state. The concept of correlated equilibrium is also introduced in a multi-user interference channel context in [74]. Different types of coordination is deeply discussed and widely used in [70].

1.8 Dynamic learning

Until now, we have realized that the Nash equilibrium suffers from the lack of far-sighted stability, i.e., the relative results can be unsatisfactory and because of this any player can have incentive to improve its outcome by moving to another strategy. The existence of the strong and coalition-proof Nash equilibrium is not guaranteed and even if so, when the number of pure strategies is large, finding such solutions is very complicated. The challenge of finding a profitable accord among players is persistent in coordinated equilibria solution. In this section, the main question we seek an answer to is: *How can the players be led to a stable joint pure strategy gaining an acceptable payoff?* This question is important, even if multiple equilibrium points with the same payoff have been identified, since each player may autonomously decide to stay in a different strategy.

Dynamic learning [75] has been widely used in order to get rid of the anarchy derived from the conflicts between selfish decisions. Learning is a joint adaptive process for agents to converge and to get the best final response. The agents either have a common interest like a team work, or each agent has its own greedy goal. Generally, there are three learning process types: *individual learning*, *joint-action learning* and *stochastic learning*. In individual learning process, the independent agents cannot observe one another's actions; i.e. for each players the opponents are passive agents. Instead, during joint-action learning process, the notion of the "optimality" is improved by adding the observation of other concurrent learners to accomplish a stable optimal solution. The stochastic learning framework, having Markovian property and a stochastic inter-state transition rule, enables each player to observe the opponents' actions history.

In the network engineering literature, Schaar *et al.* in [76] introduce a stochastic learning process among autonomous wireless agents for the optimization of dynamic spectrum access given QoS of multimedia applications. A reconfigurable multi-hop wireless network is studied by Shiang *et al.* [77] wherein a decentralized stochastic learning process optimizes the transmission decisions of nodes aimed at supporting

mission-critical applications. In [78], Lin *et al.* propose a reinforcement learning among agents of a multi-hop wireless network based on Markov decision process. Each terminal autonomously adjust transmission power in order to maximize the network utility, in a dynamic delay-sensitive environment.

Here, we study a well-known individual reinforcement learning task, namely the so-called *Q-Learning* [79]. We assume a set of players \mathcal{K} , and each player k has a finite set of individual actions \mathbf{A}_k . Each agent k individually chooses a pure joint action (strategy) to be performed $\mathbf{a}_k = (a_1, \dots, a_K) \in \mathbf{A}_1 \times \dots \times \mathbf{A}_K$ from the available joint strategy space. Q-learning enables the individual learners to achieve optimal coordination from repeated trials. Q-learning introduces a certain value Q as the immediate reward obtained after having moved to the new strategy. Each player individually updates a Q value for each of its actions. In each time step and after having selected the new joint action \mathbf{a}_k , the values of Q_k^t is individually updated. In particular, the value of $Q_k^{t+1}(\mathbf{a}_k)$ estimates the utility of performing the joint strategy \mathbf{a}_k for user k . In the seminal paper of Watkins *et al.* [79], the Q value is updated by the following recursion:

$$Q_k^{t+1}(\mathbf{a}_k) \leftarrow (1 - f_k^{t+1}) \cdot Q_k^t(\mathbf{a}_k) + f_k^{t+1} \cdot (u_k(\mathbf{a}_k) + \delta_k \cdot Q_k^t(\mathbf{a}_k)) \quad (1.25)$$

where $\delta_k \in (0, 1)$ is a discount factor and $u_k(\mathbf{a}_k)$ is a reward of the joint action \mathbf{a}_k for the respective player, and f_k is a function of t which is related to “learning rate”. Watkins *et al.* showed that given bounded rewards, learning rate $0 \leq f_k^t < 1$, and

$$\sum_{t=1}^{\infty} f_k^t = \infty, \quad \text{and} \quad \sum_{t=1}^{\infty} (f_k^t)^2 < \infty \quad \forall k \in \mathcal{K} \quad (1.26)$$

all Q_k values updating (1.25) converge a common joint pure strategy with probability one. The reward u_k is defined by a learning policy and it is not necessarily equal to the payoff defined by the game. The learning policy is greedy with respect to the Q value, i.e. the particular action \mathbf{a}_k will be selected in long-run if it makes Q value better off. Q-learning is guaranteed to converge to an optimal and stable joint strategy regardless of the action selection policy. Q-learning is not applicable where the strategy space is continuous or the number of strategies is not finite. Claus *et al.* [80] establish a simplified version of the Q recursion (1.25) which updates the Q value by the following recursion:

$$Q_k^{t+1}(\mathbf{a}_k) \leftarrow Q_k^t(\mathbf{a}_k) + \delta_k \cdot (u_k(\mathbf{a}_k) - Q_k^t(\mathbf{a}_k)) \quad (1.27)$$

For the sake of simplicity, we apply the Q recursion (1.27). In a multi learners scenario, a major challenge of Q -learning is strategy selection. When the number of strategies and players are large, the number of time step to achieve an optimal joint action exponentially increases. It is fairly clear that the best manner is to start with “exploration” of different strategies and then focus on “exploitation” of the strategies with the best value of Q . Kaelbling *et al.* in [81] recall *Boltzmann function* as an efficient strategy selection to strike a balance between exploration and exploitation. Boltzmann functions define a probability distribution among different joint actions. At each time step $t + 1$, every player will individually select the joint strategy \mathbf{a}_k with the probability $p(\mathbf{a}_k)$:

$$p(\mathbf{a}_k) = \frac{e^{E_k(\mathbf{a}_k) / T}}{\sum_{\forall \mathbf{a}_i \in \prod_{\forall k \in \mathcal{K}} \mathbf{A}_k} e^{E_k(\mathbf{a}_i) / T}} \quad (1.28)$$

The $E_k(\mathbf{a}_k) = (\delta_k)^t \cdot u_k(\mathbf{a}_k)$ is the discounted reward for taking action \mathbf{a}_k by the user k in time step t . The T is a function which provides a randomness component to control exploration and exploitation of the actions. Practically, the *temperature function* T is a decreasing function over time to decrease the exploration and increase exploitation. High values of T yields a small $p(\mathbf{a}_k)$ value and this encourages exploration, whereas a low T makes $Q(\mathbf{a}_k)$ more important, and this encourages exploitation. At time $t = 0$, each player randomly chooses a strategy and assign a random number to its own Q value. At time step t , after having been updated function T , each concurrent agents’ experience consists of a sequence of stages [80]:

- 1) Computing $p(\mathbf{a}_k)$ for all $\mathbf{a}_k \in \prod_{\forall k \in \mathcal{K}} \mathbf{A}_k$,
- 2) Generating a random number ξ_k^t uniformly distributed in $[0, 1]$, and then choosing the best joint strategy \mathbf{a}_k , i.e. the highest $p(\mathbf{a}_k)$ such that $\xi_k^t \geq p(\mathbf{a}_k)$. If $\xi_k^t < p(\mathbf{a}_k)$ for all $\mathbf{a}_k \in \prod_{\forall k \in \mathcal{K}} \mathbf{A}_k$, then the learner randomly picks a strategy,
- 3) Updating the Q_k^t value according to (1.27). If Q_k^t grows, then the learner moves to selected joint strategy \mathbf{a}_k , otherwise it stays in the current joint action and do not update Q .

Despite the individual best strategy selection of the learners, this process reach a common stable joint strategy such that all players stay there forever, i.e. no player deviates from the (common) achieved joint strategy.

The theory of learning in games studies how and which equilibria might arise as a consequence of a long-run non-equilibrium process of learning. A natural question is: *Can learning algorithms find a Nash equilibrium?* The reason for asking this question is in the hope of being able to achieve Nash equilibria, as a plausible concept, via a reasonable learning algorithm in particular when there are a large number of players and strategies. At the first look, the stability of the above addressed dynamic learning approach is described as to converge to a *pure* joint strategy and it is clear that the existence of a *pure* Nash equilibrium is not guaranteed. The fact is, in general, a dynamic learning algorithm is not able to guarantee to achieve a non-cooperative or cooperative Nash equilibrium. In the literature, there are some efforts to present a dynamic learning algorithm that achieves a Nash equilibrium in dynamic and repeated games under particular constraints [82–85].

We present now some results about Q-learning in a CDMA network. In what follows, the experimental work is presented highlighting how the agents learn to increase their individual rewards by revealing their actions. As above mentioned, the strategy selection can significantly influence the number of time steps to converge. Choosing an appropriate temperature function is a heuristic search. In our experiment, we define $T = q \cdot e^{-mt}$ as our temperature function wherein m controls the rate of exponential decay and $q > 1$ encourages the exploration of different strategies in the initial time steps.

We illustrate the behavior of mobile terminals as Q-learners in a CDMA network. Our example is a *power control problem in a CDMA network* applying Q-learning and Boltzmann function. Assume a CDMA network with K mobile terminals denoted by set \mathcal{K} . The players wish to transmit data to a certain AP. The strategies of every player is a set of discrete power levels denoted by $\mathbf{A} = \mathbf{A}_k = [\Delta p, 2\Delta p, \dots, M\Delta p]$ where Δp is our power step and $M > 1$ is an integer number. Each user has M actions to choose from, and accordingly the matrix game is made by $\times_{k \in \mathcal{K}} \mathbf{A}$ which consists of M^K joint strategies. The Shannon capacity between player k and the AP is

$$C_k = \log_2 \left(1 + \frac{N_s \cdot |H_k|^2 p_k}{\sum_{\mathcal{K} \ni i \neq k} |H_i|^2 p_i + \sigma_w^2} \right) \quad (1.29)$$

with N_s , $|H_k|^2$ and σ_w^2 denoting (the common) spreading factor for all players, user's k path gain, and the AWGN power, respectively, and where the $p_k \in \mathbf{A}$ denotes the transmit power of user k .

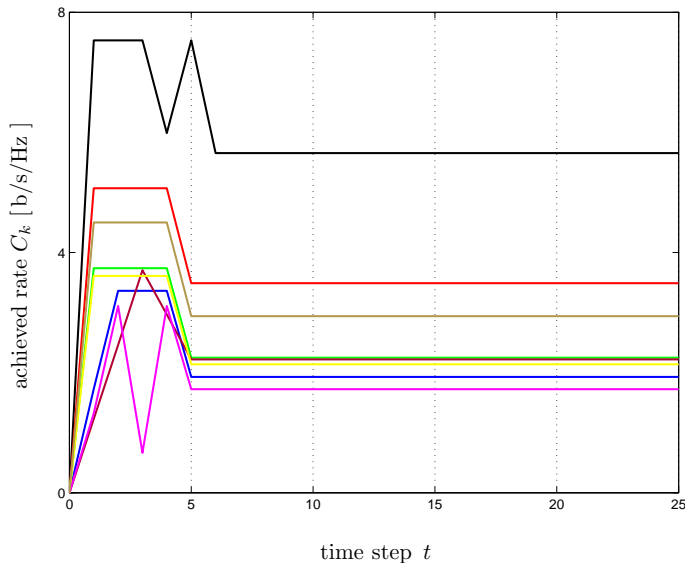


Fig. 1.4: Achieved rates as functions of the iteration step.

We introduce an individual work in which each player must individually choose the joint strategy at which achieves the best Shannon channel capacity. We simulate a learning process with $K = 8$ players such that each player k must choose the best p_k between $M = 5$ strategies. The power step is assumed to be $\Delta p = 100$ mW, the power of AWGN $\sigma_w^2 = 1$ nW and the spreading factor $N_s = 64$. The players are uniformly located at a distance between 3 and 50m from the AP. The matrix form of this game is composed of 390625 joint strategies and there may exist different power combinations (joint strategies) which achieve the same Shannon channel capacities. Q-learning leads the players to that joint strategy $(p_1, \dots, p_8) \in \mathbf{A}^8$ in which all players are satisfied of the proper achieved Shannon channel capacities. In the Q function of (1.27) for all players the discount factor parameter is fixed to $\delta_k = 0.09$ and the payoff function u_k is defined as:

$$u_k(\mathbf{a}_k) = C_k \quad \forall k \in \mathcal{K} \quad (1.30)$$

Our experiments with different parameters show that good values of the temperature function parameters are $m = 0.001$ and $q = 50$, and we start with $Q_k^{t=0} = 0$. It is obvious to say that existing a strategy in which all $u_k(\mathbf{a}_k)$ are maximal value is not

always guaranteed. Since, there is a huge conflict of interest between the players to choose different strategies.

Fig. 1.4 reports the behavior of the (reward) achievable rate C_k of $K = 8$ terminals as a function of the time step t in our scenario. The figure exhibits the convergence of all learners to a stable joint strategy after 6 time steps. Numerical results of 500 random realization of a network show the convergence of all players to a stable joint strategy after (in average) 6 steps of the iterative Q-learning algorithm wherein each joint action is probabilistically chosen according to the distribution of Boltzmann function. Furthermore, it is experimentally observed that the sum of the achieved Shannon channel capacities is (in average) 22.4 b/s/Hz and that is 94% of the maximum possible of $\sum_{k \in \mathcal{K}} C_k$.

1.9 Summary

Cooperation can be seen as the action of obtaining some advantage by giving, sharing or allowing something. In this contribution, we aimed at mapping different coalitional game approaches into communications and networking systems. A very important boundary condition for cooperation is that each participating entity is gaining more by cooperation than they would by operating alone. It is not important that all entities contribute the same effort, gain the same amount, or even have the same gain to cost ratio, but the effect of cooperation should bring advantage or gain to each cooperating entity. One different form of cooperation is altruism, a strategy wherein one of the players may sacrifice and does not gain from the cooperation to support others. In networking, for instance, one terminal sacrifices battery power and bandwidth to act as a relay for others terminals and to increase the throughput of the whole system. In some communication systems, network protocols themselves can be seen as an implicit cooperation to achieve better performance, e.g. ALOHA system. In some communication systems, network entities establish a cooperation with each other to achieve better performance, e.g. relay communications.

Cooperative game theory is a branch of game theory which aims at studying the cooperations among individual and rational participants. Unlike non-cooperative game approaches, cooperative game concepts are centralized and they need a central authority for exchange of information and policy making process. The most challenging part of a cooperative game theoretic framework is the choice of characteristic

function, since it interprets the agents perceptions of gain and satisfaction.

The main fundamental question in coalitional game theory is the question how to allocate the total generated gain by the collective of all players over the different players in the game. The distribution of payoff is described as a binding contract between the players and various criteria have been developed. The problem of the gain distribution is approached with the aid of solution concepts in coalitional game theory like core, Shapley value, kernel, and nucleolus. Core solution is the most classic solution whose result is stable against deviation of coalitions. The core solution is useful where the negotiation process is centralized and no subset of players can selfishly and privately negotiate with each other. The core set can be empty. Shapley value is the unique single-valued solution which explores the fairness in every possible prospective coalition forming. The kernel solution be understood as the set of all efficient allocations for which no pair of players want to exchange payoff. The nucleolus selects the unique imputation that successively (lexicographically) minimizes the maximal excesses. This defining property makes the nucleolus appealing as a fair single-valued solution. The kernel of a game always contains the nucleolus. The process of computing the kernel and nucleolus of arbitrary transferable utility games is hard.

The most fundamental solution concept for non-cooperative game is that of Nash equilibrium. In a Nash equilibrium no agent is motivated to deviate from its strategy given that the others do not deviate. If every player individually agrees on a certain profile of strategies without binding an agreement, then these strategies constitutes a Nash equilibrium. Nash equilibrium does not account for the possibility that groups of agents (coalitions) can change their strategies in a coordinated manner. A strategy profile is in strong Nash equilibrium if no subgroup of agents is motivated to change their strategies given that others do not change. Often the strong Nash equilibrium is too strong a solution concept, since in many games no such equilibrium exists. Coalition-proof has been suggested as a partial remedy to this problem. This solution concept requires that there is no sub-group that can make a mutually beneficial deviation (keeping the strategies of nonmembers fixed) in a way that the deviation itself is stable according to the same criterion. These solution concepts which allow coalitions to make agreements simultaneously typically suffer from incompatibility of agreements, which can give rise to empty solution sets in games of networking interest. Mixed (vs. pure) strong and coalition-proof Nash equilibrium have not been

introduced.

In a game wherein there are a huge number of agents and strategies, finding a pure cooperative/non-cooperative Nash equilibrium is hard and maybe even impossible. A learning process leads participants to a common joint action with an acceptable payoff. During a learning process agents act as independent learners, i.e. they only get information about their own action choice and payoff. As such, they neglect the presence of the other agents. The learning process happens at regular time steps and is basically a signal for the agents to start an exploration phase. During each exploration phase, some agents exclude their current best action so as to give the team the opportunity to look for a possibly better joint action. This technique of reducing the action space by exclusions was only recently introduced for finding periodical policies in games of conflicting interests. There are two problems in the process of learning optimal cooperative pursuit strategy for multiple agents. One is the probability of circulation among the actions chosen by agents which make the learning, process not converging, the other is there are many conflicts among the actions chosen by agents which make the learned pursuit strategy not optimal. Q-learning with the Boltzmann action-selection strategy guarantees the convergence of multi agents to a common and optimal joint strategy after a few time step.

1.10 Discussion

This chapter has provided a unified reference for network engineers investigating the applicability of coalitional game theory to practical problems. Different approaches such as core solution, Shapley value, kernel and nucleolus, were shown to provide a strong foundation for finding possible and stable resource/cost sharing arrangements. The results confirm the apparent analogy between the definition of Nash equilibrium in non-cooperative and coalitional game theory: both strong and coalition-proof Nash equilibria reflect on unprofitability of coalition deviations rather than an individual player deviation. In a network wherein informational exchange is possible, either through a central controller or among players themselves, the concept of coordinated equilibrium arises. The results of intuitive examples show a significantly improvement in coordinated equilibrium when compared with non-cooperative schemes. When the number of agents or strategies is large, the ability of jointly reach a consensus through environmental learning guarantees convergence to the best joint action.