Non-Cooperative Game of Multi-Agent Countermeasure Systems Based on Cognitive Hierarchy Theory

Xiaonan Dong¹, Xuehuan Li², Zongcheng Liu³ and Qiuni Li³

¹ Power Company, State Grid Corporation of China. Zoucheng, Shandong, China
² The People’s Hospital. Zoucheng, Shandong, China
³ Aeronautics Engineering College, Air Force Engineering University. Xi’an, Shaanxi, China
Email: liu434853780@163.com

Abstract. In order to research the non-cooperative and opposed problem in the real world, we establish the n-players n-strategies game model for multi-agent countermeasure network system with multi-tasking nature. Cognitive hierarchy (CH) theory is introduced to predict the game evolution, as game player is not the same reasoning in reality. A distributing virtual policy learning algorithm based on heterogeneous cognitive is proposed, which has the advantage of self-learning, self-organizing and self-optimizing. Simulation results indicate the proposed algorithm could make dynamic attack-defend mission planning effectively and give good predictions. Compared the experiment dates under different number of nodes and different parameter τ value, we find that the best τ value is between 1 and 2 and assuming appropriate τ value will optimize the whole performance of the countermeasure system.

1. Introduction
Game theory has developed into a fundamental calculus for economic and social theory, with advantages on analyzing actions and strategies on different targets or tendencies. John Nash has present the concepts of Nash Equilibrium and proofed the existence of the Nash point with the fixed point theorem (1950-1951) [1]. For the moment, game theory is developing continuous prosperity. Many experts have been awarded Nobel Prize for the contribution to game theory, such as Wekery and Mortis (1994), Robert Aumann and Thomas Schelling (2005), Leonid Hurwicz, EricMaskin and Roger Myerson (2007), AlvinRoth and Lloyd Shapley (2012).

Since the concept of the Nash equilibrium appeared, many methods and applications of the equilibrium have been developed. Simplicial methods of approximating fixed points of continuous mappings were devised in [2] and substantially developed in subsequent literatures. Mathematician Davidenko [3] developed a differential equation approach to the homotopy formulation. Based on Kohlberg and Mertens’ structure theorem, literatures [4] and [5] combined the global Newton method with the homotopy method to obtain an algorithm for computing Nash equilibrium. Other programming approaches have been implemented with different methods: a modified Newton approach [6, 7], a penalty interior point algorithm [8], distributed supervision algorithm [9] [10], a variational inequalities algorithm [11], a nonlinear complementarity approach [12], a quadratic optimization method or polynomial algebra algorithm [13] and so on. Nash equilibrium has been employed to analyze hostile situations such as the arms race and war [14], and to predict how conflict may be mitigated by repeated interaction [15]. It has also been used to study to the extent to which people with different preferences can cooperate, and whether they will take risks to achieve a cooperative outcome [16]. The electricity market models proposed in [17] form finite n-player games.
in normal form. In the financial sector, Nash equilibrium theory has been applied to the studies of currency crises and bank runs [18]. Some other applications include the organization of auctions, regulatory legislation, such as environmental regulations, penalty kicks in soccer [19].

Decision mechanism considering bounded rationality participators is also the focus of game theory research. Level-k and Cognitive Hierarchy models are two important theories which have been developed to interpret the heterogeneity rationality players in a game process. In recent years, the CH Model in [20] is used extensively to analysis and prediction, as applied to research on risk analysis in defend-attack model [21], cooperate effectively in the tacit communication game [22], prediction of the email game assumptions on higher order beliefs [23], behavior and auction prices with different information in English auctions with resale [24], learning and convergence on the beauty contest [25], and so on.

In reality, game is not always cooperative and coexist, but is non-cooperative and opposed, with jamming, attacking, defending and anti-attacking. The paper researches this kind of game problem. The major contributions we make can be concluded into two aspects: (1) A distributing virtual policy learning algorithm based on heterogeneous cognitive is proposed to predict the game evolution, this learning method is suitable for the circumstance while the nodes is dynamic change, and the heterogeneous cognitive hierarchy is near to the reality game. (2) After analyzing the experiment results, the parameter \( \tau \) value in Poisson distribution and reasoning capacity allocation is relative in the confrontation, that is, assuming appropriate \( \tau \) value will optimize the whole performance of the countermeasure system.

The rest of our paper is organized as below: In the following section, we present the multi-agent countermeasure systems model. In Section 3 and 4, we introduce cognitive hierarchy (CH) model and illustrate the distributing virtual policy learning algorithm. In Section 5, we analyze the method’s effectiveness and the plausible reasoning capacity allocation while a conclusion is drawn in the final section.

2. The Multi-Agent Countermeasure System Model of the Game

We consider two multi-agent countermeasure network systems which participate in the dynamic game and consist of \( n \) nodes and \( m \) nodes, respectively. These two systems are denoted as \( \mathcal{N} \) and \( \mathcal{M} \), which is the attacking side fighters and the IADS defending side, respectively. The countermeasure purpose of this game is summarized as follows: the attacking individuals will always try to suppress the defending side, such as weakening the scope of counteraction or damage capacity, striking the defending individuals, making them lose the confrontation capacity thoroughly, and ensuring own safety at the same time. In order to protect their own safety and minimize the loss, the defending individuals will try to anti-attack, jam or detect the attacking side.

Let the nodes’ attributes: the detecting scope \( \mathcal{A} \), the offensive scope \( \mathcal{B} \), the damage capacity \( \mathcal{C} \), the communication ability \( \mathcal{D} \), the speed \( \mathcal{V} \), the value of the entity node \( \mathcal{P} \), the detecting probability \( \mathcal{P}_{ij} \), and the attacking probability \( \mathcal{P}_{ij} \). Suppose that the information of the nodes is known to the opposite.

Assume that the information of the nodes of both sides in this game is completely known by the opposition, which means the numbers and properties of all the nodes are known by the opposition.

Let \( \mathcal{G} = \{O, E, F\} \) be the strategy action space of the attacking side and \( \mathcal{G}' = \{O', E', F'\} \) be the strategy action space of the defending side. \( E, F \) and \( O \) represent three kinds of strategic actions: jamming, attacking and no actions, respectively. Every multi-agent node chooses one kind of strategic action \( g_i \in \mathcal{G} \) or \( g'_i \in \mathcal{G}' \) in each round of the game, so as to increase threats to the opposition and keep the value of the actor, as well as maximize the expected payoffs of the actor. In addition, the probability of strategy satisfies the constraints: \( g_{G,O} + g_{G,E} + g_{G,F} = 1 \) and \( g'_{G',O} + g'_{G',E} + g'_{G',F} = 1 \).

The mixed strategy vectors are described as follows. Consider that node \( i \in \mathcal{N} \) jams the opposition node with the probability \( \pi_{i,g_i} \), let \( \Pi_i = \{\pi_{i,g_i} \mid g_i \in \mathcal{G}\} \) be the mixed strategy vector of all the possible strategies \( \Pi_i = \{\pi_{i,g_i} \in R : \sum_{g_i \in \mathcal{G}} \pi_{i,g_i} = 1\} \), and denote \( \Pi_{-i} = \{\pi'_{i,i'} \in \mathcal{N} \backslash \{i\}\} \). Similarly, consider node
where $q_1$, $q_2$, $q_3$, $q_4$ ($q_1 + q_2 + q_3 + q_4 = 1$) are the weight coefficients.

Then the expected payoffs are calculated by Equation (3) and Equation (4), respectively. Where $E_{x, \Phi}$ is the expected action on the probability distribution $\{\Pi, \Phi\}$.

$$V_i(g_i, \pi_i, \Phi) = E_{x, \Phi}[V_i(g_i, \pi_i, \Phi)] = \sum V_i(g_i, \pi_i, \Phi) \cdot \Pi \quad (3)$$

$$U_j(g_j, \phi_j, \Pi) = E_{x, \Phi}[U_j(g_j, \phi_j, \Pi)] = \sum U_j(g_j, \phi_j, \Pi) \cdot \Phi \quad (4)$$

3. Level-K Reasoning Based no Cognitive Hierarchy Theory

The traditional virtual game is based on the same reasoning. It will be lack of close to truefulness. Moreover, in a statics game, if the actual strategy is not stable, using historical strategy to instead of actual strategy will result in unreasonable outcomes possibly. In the most real game, the players are not the same reasoning capability, but rather than heterogeneous. Therefore, in the multi-agent countermeasure systems, we introduced a one-parameter cognitive hierarchy (CH) model [20] to investigate behavior in one-shot games, and initial conditions in repeated games. Literatures [29-30] show that this model fits empirical data, and explains why equilibrium theory predicts behavior well in the game.

Cognitive hierarchy (CH) theory is a model of hierarchies of beliefs. Formally, the CH model has two components: (1) decision rules for players doing $k$ levels of thinking, and (2) an frequency distribution of levels $f(k)$. The CH model consists of iterative decision rules for players doing $k$ levels of thinking, that is a step-by-step reasoning procedure and best-responds. And the frequency distribution $f(k)$ of level $k$ players assumed to be Poisson distribution, from level 0 to level $k$ - 1. Strategic categories are defined in The CH model as follows inductively: level 0 players randomize; and level $k$ thinkers’ best-respond, assuming that other players are distributed over level 0 through level $k$ - 1. Each player assumes that he understands the game better than the other players and his strategy is the most sophisticated. That is to say, the same level players have the same beliefs, make the same choices, and have the same expected payoffs, while they are in the same state of the game; the higher level reasoning players can conclude the lower level reasoning players’ strategies and actions exactly. In addition, they have a memory of the game history. The lower level reasoning players would not realize the existence of the higher level reasoning players.
This paper proposed an iterative approach to study levels of reasoning, as well as beliefs and actions in a repeated normal-form game. Use a level-\( k \) model of limited strategic reasoning and allowing for other-regarding preferences. On the base of empirical research from Camerer, Ho and Chong, The estimates of \( \tau \) generally fell between 1 to 2 and fit actual choices reasonably well, and are very likely to predict as accurately as Nash equilibrium. In our model’s study, we main focus on \( \tau \in \{0.2,0.3,0.4,...,2.8,2.9,3\} \), and analyze the dynamic performance in the process decision-making and the game evolution, under the heterogeneous cognitive circumstance. As aforementioned, we infer the following equations: \( N_i = f(k) = \frac{\tau^k e^{-\tau}}{k!} \); \( M_k = f(k) = \frac{\tau^k e^{-\tau}}{k!} \), and \( \sum N_i = \sum f(k) = N \), \( \sum M_k = \sum f(k) = M \).

Combining with the illustrated multi-agent countermeasure systems model in the previous section, we make supposition: (1) In order to save resources, when the high level player estimate the attacking target node of the opposite will be destroyed, it won’t choose the node for target. (2) In order to ensure its own safety, the high level player is prior to choose one of the low level players as target node, if the low level players are predicted to attack this high level player.

## 4. Virtual Policy Learning Algorithm

In order to gain the MSNE, we proposed a distributing virtual policy learning algorithm with the heterogeneous cognitive mechanism. Suppose the attacking side and the defending side will choose their own strategy with a certain probability. In each learning time \( t \), the attacking node \( i \) and the defending node \( j \) choose a pure strategy \( g_i^t \in G \) or \( g_j^t \in G^* \) respectively as the optimal response based on other players’ mixed strategies. Let \( \Pi^{-1}_i = \{\pi_i^{-1}, i' \in N \setminus \{i\} \} \) denote the mixed strategies of nodes \( i' \in N \setminus \{i\} \) at the time \( t-1 \), and \( \Phi^{-1}_j = \{\phi_j^{-1}, j \in M \} \) denote the mixed strategies of nodes \( j \in M \) at the time \( t-1 \). Similarly, let \( \Phi^{-1}_j = \{\phi_j^{-1}, j' \in N \setminus \{j\}\} \) and \( \Pi^{-1}_i = \{\pi_i^{-1}, i \in N \} \). As a consequence, \( g_i^t \) and \( g_j^t \) can be expressed as

\[
g_i^t = \arg \max_{g_i \in G} E_{\pi_i^{-1}, \Phi^{-1}_j}[V(g_i, \pi_i^{-1}, \Phi^{-1}_j)]
\]

\[
g_j^t = \arg \max_{g_j \in G} E_{\pi_i^{-1}, \Phi^{-1}_j}[U_j(g_j, \phi_j^{-1}, \Pi^{-1}_i)]
\]

In addition, set three dimensional vectors for nodes \( N \) or nodes \( M \) \( h_i^t = [h_i^0, h_i^1, h_i^2] \) and \( h_j^t = [h_j^0, h_j^1, h_j^2] \) to storage the decision-making vector at a certain time. when the multi-agent node choose no action, jamming and attacking, the value is \([1,0,0], [0,1,0], [0,0,1]\) respectively.

The strategies update according to Equation (7) and Equation (8).

\[
\pi_i^t = \pi_i^{-1} + \frac{1}{t}(h_i^t - \pi_i^{-1})
\]

\[
\phi_j^t = \phi_j^{-1} + \frac{1}{t}(h_j^t - \phi_j^{-1})
\]

It is the deviation linear combination of the last mixed strategy and the cumulative mixed strategy in every step’s update. The learning process will continue iterate until its convergence precision reaches a factor small enough. So that the MSNE is found out, and both sides take actions with stable probabilities.

The distributing virtual policy learning algorithm based on heterogeneous cognitive:

1: Initialization: set \( t=1 \), initialize \( \Pi^0 \) and \( \Phi^0 \), \( h_i^0 = [h_i^0, h_i^0, h_i^0] = [1,1,1] \) and \( h_j^0 = [h_j^0, h_j^0, h_j^0] = [1,1,1] \);
2: Repeat
3: For all $k_i \in k$, do
4:    repeat
5:        for all $i \in N_k$, do
6:            Decision-making $g_i$ according to equation (5), update $h_i$;
7:        end for
8:    end for
9:    for all $j \in M_k$, do
10:    Decision-making $g_{ij}$ according to equation (6), update $h_{ij}$;
11:    Update $\pi_{ij}$ according to equation (7);
12: end for
13: End for
14: Delete the destroyed nodes, update the node’s information
15: $t = t + 1$;
16: Until $\|\Pi - \Pi^{t-1}\| + \|\Phi - \Phi^{t-1}\| < \varepsilon$

Form the above, we can find that the complexity of the method is $\sum (N_k + M_k) * k$ and the payoff matrix space of the method is $6t*n*m$ dimensions. Since at a certain time slot $t$, the payoff matrix of every node is $n*3$ or $m*3$ dimensions, and the payoff matrix space of all one side’s nodes is $n*m*3$ dimensions. So, at all time slots, the payoff matrix space of all nodes is $6t*n*m$ (that is $2*3*n*m*t$) dimensions.

5. Experiment and Results
In this simulation, both nodes $N$ and nodes $M$ of the two multi-agent countermeasure systems are distributed randomly in the space of 1000km*1000km. The initial strategy probability vectors are assumed to be [0.33,0.33,0.34] and [0.5,0.5]. Set weight coefficient $q_1 = 0.35$, $q_2 = 0.4$, $q_3 = 0.1$, $q_4 = 0.15$ for nodes $N$ and $q_1 = 0.4$, $q_2 = 0.35$, $q_3 = 0.1$, $q_4 = 0.15$ for nodes $M$. The convergence precision is assumed to be 0.001, and threshold value $C_{\text{threshold}}$ is set as 0.5. $d_{nj}$ is set as all 1’s matrix, that is the initial communication is fully connected, the weight of connection $w_{nj}$ is assumed to randomly generate between 0 and 1. The other parameters used in the simulation are assumed to randomly generate as Table 1:

| parameters | A (km) detecting domain | B (km) offensive domain | C damage capacity | P_c/(10^4CNY) value | V (km/h) speed | $w_{nj}$ weight of connection | P_d detecting probability | P_a attacking probability |
|------------|--------------------------|-------------------------|------------------|---------------------|----------------|-----------------------------|--------------------------|--------------------------|
| range      | (0, 100)                 | (5, 30)                 | (10, 100)        | (10,000)            | (100,1000)    | (0, 1)                       | (0.5, 0.99)              | (0.5, 0.99)              |

We observe the game’s attacking and defending evolution and observe how the average expected payoffs and the accumulated sum of expected payoffs change under different number of nodes and different parameters in the Poisson distribution.

We observe the game’s attacking and defending evolution and observe how the average expected payoffs and the accumulated sum of expected payoffs change under different number of nodes and different parameters in the Poisson distribution.
(a) $n=m=15$, $\tau_n = 1.5$, $\tau_m = 0.5$, an equilibrium is achieved
Figure 1. The average expected payoff and the accumulated sum of expected payoffs. The blue line describes nodes $N$, red line describes nodes $M$.

Figure 1 shows the average expected payoffs and the accumulated sum of expected payoffs change in the game process. There are two kinds final state: (a) the game reaches an equilibrium in the end. (b) all nodes $M$ is destroyed in the end. In Figure 1, the average expected payoff line emerges jumping phenomenon, when the nodes destroy the opposite’s nodes, the average expected payoff is going up; when the nodes are destroyed by the opposite’s nodes, the average expected payoff is going down. If jamming happens, the average expected payoff may fluctuate within a range. Both Figure 1 (a) and Figure 1 (b) the accumulated sum of expected payoffs line keep up continue. We observe that the game players would learn themselves, adjust the strategy to optimize the payoff in the dynamic attack-defend confrontation. The multi-agent countermeasure system can dynamically make mission re-planning, achieve target assignment, trajectory planning and strategy selection, has advantages of self-organizing and self-optimizing. In a word, the proposed method gives effective predictions for the attack-defend game evolution.
The parameter of Poisson distribution
The accumulated sum of expected payoffs

(a) \( n = m = 15 \)

The average expected payoffs

(b) \( n = m = 20 \)
The parameter of Poisson distribution

The average expected payoffs

The accumulated sum of expected payoffs

(c) $n=m=30$
The parameter of Poisson distribution

The average expected payoffs

The accumulated sum of expected payoffs

(d) $n=m=40$
Figure 2. The average expected payoff and the accumulated sum of expected payoffs under different parameter $\tau$, when an equilibrium is achieved or all nodes of one side are destroyed. The blue line describes nodes $N$, red line describes nodes $M$.

By changing the parameter $\tau$ of Poisson distribution while the other parameters is fixed, when the game reached equilibrium, The accumulated sum of expected payoffs presents a mean reversion phenomenon. It fluctuates around the mean value and has a return tendency. Figure 2 (a) shows when the number of nodes is small, the $\tau$ value has little influence in the game. Figure 2(e) shows fierce confrontation spring up among the two countermeasure systems, and all the nodes of vulnerable side are destroyed, the average expected payoff is near to 0, once the number of nodes increased to a certain extent. That is similar to Malthus effect. Comprehensively analysis on Figures 2 (a) to (e), we observe that the best $\tau$ value is between 1 and 2, As the empirical result of Camerer, Ho and Chong [20]. This model fits empirical data and the maximal accumulated sum of expected payoffs is obtained
at the best \( \tau \) value. More experiments are done. When the initial strategy probability vector is assumed to be random, the equilibrium status and the best \( \tau \) value will be changed, it is also between 1 and 2.

Table 2. Stochastic assignment of related parameters

| Number of nodes \((n,m)\) | \(\tau\) value \((\tau_n, \tau_m)\) | The accumulated sum of expected payoffs | The total payoffs increased or decreased value and percentage | remarks |
|--------------------------|----------------------------------|----------------------------------------|------------------------------------------------|---------|
| (15,15)                  | (1.5,1.5)                        | (90.51,125.90)                         | \((\uparrow 16.99, \uparrow 4.2), (\uparrow 18.77\%, \uparrow 3.34\%)\) |         |
| (15,15)                  | (1.5,2.2)                        | (107.50,130.10)                        | \((9.6, [15.6], [10.61\%, \uparrow 12.39\%])\) |         |
| (15,15)                  | (1.5,0.5)                        | (80.91,110.30)                         | \((\downarrow 9.6, \downarrow 15.6), (\downarrow 10.61\%, \downarrow 12.39\%)\) |         |
| (30,30)                  | (1.8,1.8)                        | (157.90,173.60)                        | \((\downarrow 8.6, \downarrow 15.6), (\downarrow 9.6\%, \downarrow 12.39\%)\) |         |
| (30,30)                  | (0.2,1.8)                        | (109.80,179.70)                        | \((\downarrow 48.10, \uparrow 16.10), (\downarrow 30.46\%, \uparrow 3.51\%)\) | All nodes \(N\) destroyed |
| (30,30)                  | (3.1,8)                          | (123.10,206.90)                        | \((\downarrow 34.8, \uparrow 27.2), (\downarrow 22.04\%, \uparrow 19.18\%)\) |         |
| (50,50)                  | (1.8,1.8)                        | (55.75,74.26)                          | \((\downarrow 50.76, \downarrow 70.05), (\downarrow 91.04\%, \uparrow 3.51\%)\) |         |
| (50,50)                  | (1.8,0.5)                        | (4.99,4.21)                            | \((\downarrow 50.76, \downarrow 70.05), (\downarrow 91.04\%, \uparrow 3.51\%)\) |         |
| (50,50)                  | (1.8,2.2)                        | (29.27,36.97)                          | \((\downarrow 26.48, \uparrow 37.29), (\downarrow 50.76\%, \uparrow 50.22\%)\) |         |

Table 2 shows the total of expected payoffs increased or decreased situation under best \( \tau \) value and other \( \tau \) value. The increased or decreased data is compared with before \( \tau \) value changed. We can find that the best \( \tau \) value has obvious advantage on the accumulated sum of expected payoffs. When \( n=m=15 \), the advantage is not so clear. With number of nodes increasing, the superiority is obvious, the gap of total payoff is enlarge, even all nodes of one side are destroyed. This indicates the better \( \tau \) value, the more plausible reasoning capacity allocation. In other words, assuming appropriate \( \tau \) value will optimize the whole performance of the countermeasure system. However, the disadvantage is when the \( \tau \) values have little difference; the optimized performance may be not obvious.

6. Conclusion
In this paper, we establish the multi-agent countermeasure network system model to analyze non-cooperative mixed game, which is integrate with many characters of resources and has the property of target selection, detection, jamming, attacking and defending mission. Simulation results indicate that the proposed method is effective for researching the problem of attack-defend game evolution. Nodes can make dynamic mission planning coordinate. The introduced Cognitive hierarchy (CH) theory makes the prediction more reality. Compared the experiment data under different number of nodes and different \( \tau \) value in the Poisson distribution, we observe the best \( \tau \) value is between 1 and 2 and the better \( \tau \) value, the more plausible reasoning capacity allocation in the countermeasure system. In future work, the problem of specific game strategy will be researched base on Cognitive Hierarchy Theory.

7. Acknowledgments
The work described in this paper is supported in part by the National Natural Science Foundation of China under grant 61603411.

8. References
[1] Nash, J. F. 1996 Essays on game theory. John Nash.
[2] Scarf, H. 1967 the approximation of fixed points of a continuous mapping. Siam Journal on Applied Mathematics, 15(5), 1328-1343.
[3] Davidenko, D. F. 1953 on a new method of numerical solution of systems of nonlinear equations. Doklady Akademii Nauk Sssr, 88(4), 601-602.
[4] Govindan, S., & Wilson, R. 2003 A global newton method to compute nash equilibria. Journal of Economic Theory, 110(1), 65-86.
[5] Watson, L. T., Billups, S. C., & Morgan, A. P. 1987 Algorithm 652: hompack: a suite of codes for globally convergent homotopy algorithms. Acm Transactions on Mathematical Software, 13(3), 281-310.
[6] Govindan, S., & Wilson, R. 2003 A global newton method to compute nash equilibria. Journal of Economic Theory, 110(1), 65-86.
[7] Govindan, S., & Wilson, R. 2003 A global newton method to compute nash equilibria. Journal of Economic Theory, 110(1), 65-86.
[8] Hobbs, B. F., Metzler, C. B., & Pang, J. S. 2000 Strategic gaming analysis for electric power systems: an mpec approach. IEEE Trans Power Syst, 15(2), 638-645.
[9] Casavola, A., Garone, E., & Tedesco, F. 2014 A distributed multi-agent command governor strategy for the coordination of networked interconnected systems. Automatic Control IEEE Transactions on, 59(8), 2099-2112.
[10] Feng, X., Sun, G., Gan, X., & Yang, F. 2014 Cooperative spectrum sharing in cognitive radio networks: a distributed matching approach. IEEE Transactions on Communications, 62(8), 2651-2664.
[11] Campos, F. A., Villar, J., & Barquin, J. 2010 Solving cournot equilibriums with variational inequalities algorithms. Proc. Inst. Eng. Technol., Gen., Transm., Distri.,4(2), 268-280.
[12] Xian, W., Yuzeng, L., & Shaohua, Z. 2004 Oligopolistic equilibrium analysis for electricity markets: a nonlinear complementarity approach. IEEE Transactions on Power Systems, 19(3), 1348-1355.
[13] Datta, R. S. 2010 finding all nash equilibria of a finite game using polynomial algebra. Economic Theory, 42(1), 55-96.
[14] Schelling, T. C. 1980 the Strategy of Conflict. The strategy of conflict. Harvard University Press.
[15] Heap, S. H., & Varoufakis, Y. 1996 Game theory: a critical introduction. Journal De Economistes Et Des Etudes Humaines, 7(2-3), 453-462.
[16] Skyrms, B. 2004 The Stag Hunt and the Evolution of Social Structure. The stag hunt and the evolution of social structure. Cambridge University Press.
[17] Yang, Y., Zhang, Y., Li, F., & Chen, H. 2012 computing all nash equilibria of multiplayer games in electricity markets by solving polynomial equations. IEEE Transactions on Power Systems, 27(1), 81-91.
[18] Cooper, R. 1999 Coordination games. Cambridge University Press.
[19] Lee, K. H., & Baldick, R. 2007 tuning of discretization in bimatrix game approach to power system market analysis. IEEE Power Engineering Review, 22(11), 55-55.
[20] Camerer, C. F., Ho, T. H., & Chong, J. K. 2004 A cognitive hierarchy model of games. Quarterly Journal of Economics, 119(3), 861-898.
[21] Rothschild, C., Melay, L., & approach. Guikema, S. 2012 Adversarial risk analysis with incomplete information: a level-k Risk Analysis, 32(7), 1219–1231.
[22] Weerd, H. D., Verbrugge, R., & Verheij, B. 2015 Higher-order theory of mind in the tacit communication game. Biologically Inspired Cognitive Architectures,11, 10-21.
[23] Strzalecki, T. 2014 Depth of reasoning and higher order beliefs. Journal of Economic Behavior & Organization, 108(C), 108-122.
[24] Georganas, S. 2011 English auctions with resale: an experimental study. Games & Economic Behavior, 146(1), 359-371.
[25] Chen, S. H., Du, Y. R., & Yang, L. X. 2014 Cognitive capacity and cognitive hierarchy: a study based on beauty contest experiments. Journal of Economic Interaction & Coordination, 9(1), 69-105.
[26] Li, Q.N., Yang, R. N. 2017 Research on the Non-cooperative Game Strategy of Suppressing IADS for Multiple Fighters Cooperation. Journal of Xidian University, 44(6): 140-147.
[27] Li, Q.N., Yang, R. N., Liu, Z. C. 2017 Int. Conf. on Comm. and Inf. Syst. (Wuhan), China.
[28] Li, Q.N., Yang, R. N., Liu, Z. C. 2017 Int. Conf. on Inf. Tech (Singapore)
[29] Rothschild, C., Mclay, L., & approach. Guikema, S. 2012 Adversarial risk analysis with incomplete information: a level-k Risk Analysis, 32(7), 1219–1231.
[30] Weerd, H. D., Verbrugge, R., & Verheij, B. 2015 Higher-order theory of mind in the tacit communication game. Biologically Inspired Cognitive Architectures, 11, 10-21.