Instablity of a punishment strategy in correlated equilibria

Takuho Mitsunaga*, Yoshifumi Manabe†, Tatsuaki Okamoto‡
*Dept. of Social Informatics, Kyoto university
Yoshidahonmachi, Kyoto city, Kyoto, Japan
Email: mitsunaga@ai.soc.i.kyoto-u.ac.jp
†NTT Communication Science Laboratories
3-1 Morinosato-Wakamiya, Atsugi city, Kanagawa, Japan
‡NTT Information Sharing Platform Laboratories
3-9-11 Midoricho, Musashino city, Tokyo, Japan

***Abstract*—In game theory, to achieve correlated equilibria without a trusted mediator, the idea of replacing the mediator with protocol execution by players is suggested. Before players take actions in a game, players communicate with each other by following a protocol. In that model, the concept of a punishment strategy is defined for cases in which a player (or some players) aborts the protocol. In this paper, we present an example of game in which a punishment strategy does not work and suggest an improved definition of a punishment strategy.**

***Index Terms*—Game theory, Nash equilibrium, Correlated equibrium, Punishment strategy.**

I. Introduction

For years, in the field of cryptography, researchers have been concerned with applying game theory to cryptography. This is because cryptography and game theory pertain to the study of interactions among mutually distrusting players. Cryptographic protocols are designed under the assumption that some players are honest and faithfully follow the protocol, while some players are malicious and behave arbitrarily. However in game theory, all players are considered to be rational and behave in order to maximize their profits. In traditional cryptography theory, if a player is corrupted, he is considered to be dishonest and may even take an unreasonable action that the other players can not expect. However in game theory, almost in the same way as in the real world, it is assumed that each player selects his action from the viewpoint of the profit he can achieve even if he is not honest.

One of the most important ideas in game theory is equilibrium which is the best way for all players to follow actions. Two kinds of equilibrium were proposed. First, Nash equilibrium (named after John Forbes Nash, who proposed it) is a solution concept for a game involving two or more players, in which each player is assumed to know the equilibrium strategies of the other players and no player has anything to gain by only changing his own strategy ([1]). The other is a correlated equilibrium, which was proposed by Robert Aumann [2], and is a solution concept that is more general than the well known Nash equilibrium. The idea is that each player chooses his action according to his observation of the value of the single public signal. This signal is supposed to be sent by a trusted third party called a mediator. The mediator chooses the set of moves according to the right joint distribution and privately informs each player of what his designated move is. Then the next question is ”can we remove the mediator by using some protocols?”. In the case of a two-player game, it is well known that in the standard cryptographic models the answer is positive, provided that the two players can interact (see [3]). This positive result can be carried over to the game theory model as well. Especially, we consider an extended game, in which the players first exchange some messages (this part is called ”cheap talk” in game theory), and then choose their actions and execute them simultaneously as in the original game. In [4], Dodis et al. suggested the concept of a punishment strategy, which is a kind of rule for players not to abort the protocols in the cheap talk phase. If a player aborts during the cheap talk phase, the other players take actions that cause the utility level of the aborting player to decrease. So all players have incentives not to abort during the cheap talk phase or to deviate from the actions in the original game. This topic is similar to that of strong equilibrium in the terms of an equilibrium for deviation of multiple players. However, a strong equilibrium is an equilibrium in a game for every subset of players where they can not increase their utilities by deviating from an equilibrium. On the other hand, our topic is to achieve the correlated equilibrium by using communications among players called cheap talk and punishment strategy when players do not follow the protocol in the cheap talk phase. However, definitions of a punishment strategy so far have focused only on the utilities of the punished players. Thus punishment might also decrease the utilities of punishing players. So under the assumption that malicious players select their actions rationally in terms of their utilities, there is a case in which the punishment strategy does not work, i.e, when the punishment strategy is not better than any other strategies. In this paper, we show an example of a game in which the punishment strategy does not work and suggest an improved

		P_2		
		$\acute{A}$	$\acute{B}$	$\acute{C}$
P_1	A	(11,6)	(7,8)	(8,10)
	B	(8,6)	(10,10)	(6,7)
	C	(8,12)	(4,3)	(10,9)

Fig. 1. Two player game.

definition of the punishment strategy.

II. Preliminaries

A. Game theory

In game theory we assume players take actions and have their own utility functions that are determined by a set of all players actions. An n-player game, Γ is denoted as $\Gamma = (\{A_i\}_{i=1}^n, \{u_i\}_{i=1}^n)$.

A_i is a set of actions of each player $i(P_i$ from now on). Player P_i selects an action $a_i \in A_i$. u_i is a utility function of P_i. $N = \{P_1, P_2, \ldots P_n\}$ is the set of all players. The game is played by having every player takes action $a_i \in A_i$ simultaneously. The payoff to P_i is given by $u_i(\boldsymbol{a})$, where $\boldsymbol{a}$ is the tuple of the action of each player, ($\boldsymbol{a} = (a_i, \ldots, a_n)$). P_i prefers outcome $\boldsymbol{a}$ to outcome $\acute{\boldsymbol{a}}$ iff $u_i(\boldsymbol{a}) \geq u_i(\acute{\boldsymbol{a}})$. We say P_i strictly prefers outcome $\boldsymbol{\alpha}$ to outcome $\acute{\boldsymbol{\alpha}}$ if $u_i(\boldsymbol{\alpha}) > u_i(\acute{\boldsymbol{\alpha}})$ and P_i weakly prefers $\boldsymbol{\alpha}$ to $\acute{\boldsymbol{\alpha}}$ if $u_i(\boldsymbol{\alpha}) \geq u_i(\acute{\boldsymbol{\alpha}})$. We assume that information of all possible actions of the players, A=$A_1 \times \ldots \times A_n$ and utility functions $u = u_1 \times \ldots \times u_n$ are common knowledge among the players.

We show an example of a two-player game in Fig. 1. It can be represented in a matrix form by mapping actions A_1 to rows and A_2 to columns.

The entry in the cell at row $a_1 \in A_i$ and column $a_2 \in A_2$ contains tuple (u_1,u_2) indicating the payoffs to P_1 and P_2, respectively, given the outcome $\boldsymbol{a}$ = (a_1,a_2). The example in Fig.1 represents a game where $A_1 = \{A, B, C\}$, $A_2 = \{\acute{A}, \acute{B}, \acute{C}\}$, and e.g., $u_1(A, \acute{A}) = 11$ and $u_2(A, \acute{A}) = 6$.

B. Nash equilibrium

If players play a game and P_1 knows the actions the other players will take, P_1 will select an action $a_1 \in A_1$ that maximizes $u_1(\boldsymbol{a})$. If a_1 is the best way a_1 is called the best response to the actions of the other players for P_1. If for every player action a_i is the best response to the other actions, we call the tuple of actions ($\boldsymbol{a} = (a_1, ..., a_n) \in A$) a Nash equilibrium. Formally, we define $\boldsymbol{a_{-i}} = (a_1, .., a_{i-1}, a_{i+1}, .., a_n)$ and let $(\acute{a_i}, \boldsymbol{a_{-i}})$ denote $(a_1, ..., a_{i-1}, \acute{a_i}, a_{i+1}, ..., a_n)$.

In a Nash equilibrium each player can not receive additional profit by deviating from his strategy. In the example in Fig. 1, P_1 may think that P_2 selects $\acute{A}$ to receive maximum payoff 12 ((a_1,a_2) = (C,$\acute{A}$)), so P_1 may select strategy A to receive maximum payoff 11 under the assumption that P_2 will take $\acute{A}$. However, if P_2 thinks that P_1 will take this strategy, $\acute{C}$ becomes a better strategy for P_2. In this case, if players try to maximize their payoffs, their strategies and the prediction of strategies that the other player will take are changing except for the point that is the best response for each player. In the example of Fig. 1., we can see, $u_1(A, \acute{B}) \leq u_1(B, \acute{B}) \geq u_1(C, \acute{B})$ and $u_2(B, \acute{A}) \leq u_2(B, \acute{B}) \geq u_2(B, \acute{C})$.

So B is the best response to actions of P_2 and $\acute{B}$ is the best response to actions of P_1. In this case, the set of actions $(B, \acute{B})$ fulfills the condition of the Nash equilibrium $u_i(\acute{a_i}, \boldsymbol{a_{-i}}) \leq u_i(\boldsymbol{a})$ for all i.

C. Correlated equilibrium

The concept of a correlated equilibrium is suggested in [2]. This equilibrium may give a better payoff than the Nash equilibrium for every player P_i. A correlated equilibrium can be described by means of a joint distribution over the strategy sets.

Let $\Gamma = (\{A_i\}_{i=1}^n, \{u_i\}_{i=1}^n)$ be an n-player game. Then, $\alpha \in A_1 \times \ldots \times A_n$ denotes the set of n-tuple strategies of Γ. We assume the existence of external party M called the mediator and define a mediated version of Γ that relies on M.

The game is now played in two stages: first, the mediator M chooses a tuple of actions, $\boldsymbol{a} = (a_1, \ldots, a_n) \in A$, according to some known distribution D, and then hands the recommendation a_i to player P_i. Secondly the players play Γ as before by choosing any action in their respective action sets. Players are supposed to follow the recommendation of M, and it is the best response for each player to realize a correlated equilibrium. To define formally this notion, let $u_i(\acute{a_i}, \boldsymbol{a_{-i}}|a_i)$ denote the expected utility of P_i, given that he plays action $\acute{a_i}$ after having received recommendation a_i and all other players take their recommended actions $\boldsymbol{a_{-i}}$.

Definition 1: Let $\Gamma = (A_i, u_i)$. Distribution $D \in \Delta(A)$ is a correlated equilibrium if for all $\boldsymbol{a} = (a_1, \ldots, a_n)$ in the support of D, all i, and all $\acute{a_i} \in A_i$, it holds that $u_i(\acute{a_i}, \boldsymbol{a_{-i}}|a_i) \leq u_i(\boldsymbol{a}|a_i)$.

D. Realizing correlated equilibrium with cheap talk

Consider some n-player game Γ = (A_i,u_i) in normal form, along with a correlated equilibrium D. We then define the extensive form game Γ_{CT} in which all players first communicate in a cheap talk phase before the original game Γ. Following the game-theoretic convention, all players must take some actions in Γ, i.e., we do not allow player P_i to abort in Γ unless this is an action in A_i. On the other hand, following the cryptographic convention we allow players to abort during the cheap talk phase. In case players abort during the cheap talk phase, we must consider a new idea for each player to move properly.

III. Punishment strategy

A punishment strategy was suggested as a kind of rule to prevent players from aborting protocols in the cheap talk phase. If a player aborts, the other players take actions that cause the ulitity level of aborting player to decrease. So there in no incentive for any players to abort in the cheap talk phase and to deviate from an action in the original game. The initial result of employing the punishment strategy was shown in [4], which examines the case of a two-player game. The basic idea is described hereafter. Let D be a correlated equilibrium in a two-player game Γ in Γ_{CT}, the two players run a protocol Π

to calculate $(a_1, a_2) \leftarrow D$, where player P_i receives a_i as an output. This protocol, Π, is secure-with-abort (cf.[6]), which informally means that privacy and correctness hold. On the other hand, fairness does not hold in particular, we assume it is possible for P_1 to receive its output even though P_2 does not. After running Π, each player takes the action it received as the output in Π. If P_2 does not receive an output from Π then it plays the minimax profile against P_1. The minimax profile against P_i is an action $a_{-i} \in A_{-i}$ that minimizes $max_{a_i \in A_i}$ $u_i(a_i, a_{-i})$. Kats generalized this punishment strategy from two players to n-players in [7]. Assume that some players select actions following the recommended actions from the outputs of Π, while some collude with each other (which is called coalition C) and deviate from the recommendation. C prefers σ to $\acute{\sigma}$ only if every player in C weakly prefers σ to $\acute{\sigma}$ and some player in C strictly prefers σ to $\acute{\sigma}$.

Definition 2: Let Γ be an n-player game with correlated equilibrium D. A strategy vector ρ is a t-punishment strategy with respect to D if for all $C \subseteq$ N with $\mid C \mid \leq$ t, and all $\acute{\sigma}_C$, it holds that for all $i \in C$, $u_i(\acute{\sigma}_C, \rho_{-C}) \leq u_i(D)$.

We introduce another definition of punishment strategy as described in [5]. In [5], Dolev et al. considered a case with k-immune, which means that the strategy is tolerant to at most k Byzantine failure players. Byzantine fault tolerant means that there is nothing that players in a set T of size at most k can do to give the rest of players a worse payoff, even if the players in T can communicate with each other. For simplicity of discussion, this paper assumes that k=0, that is, there is no Byzantine failure players. They also consider type t_i which is an input given to each player at the beginning. This paper does not consider type t_i, that is, there is a single type for every player. The example in this paper can be easily extended to cases where there are multiple types for players.

Definition 3: If Γ is an underlying game with a mediator M, a strategy profile ρ in Γ is a t-punishment strategy with respect to a strategy profile σ in Γ if for all subsets $C \subseteq N$ with $\mid C \mid \leq$ t, all strategies ϕ in Γ with a cheap talk CT(C) among players in C, and all players $i \in C$, $u_i(\Gamma, \sigma) > u_i(\Gamma + CT(C), \phi_C, \rho_{-C})$.
A remarkable difference between Definition 2 and Definition 3 is the allowing of equal utilities. In regard to this, Definition 3 requires a stronger condition. Intuitively, for any set C, even if all players in C collude and communicate with each other during the cheap talk, no player in C can obtain a better payoff than the correlated equilibrium if the rest of the players select the punishment strategy. In [7], Katz showed that if there is a punishment strategy, in a five-players game with two malicious players, Nash equilibrium can be implemented. In [8], In an n-player game with t malicious players, if $n > 2t+k$ (k is the number of Byzantine failure players) and there is a punishment strategy, the Nash equilibrium can be implemented.

IV. Cheating players' actions against punishment strategy

This section describes an example in which the punishment strategy does not prevent the players in C from aborting during the cheap talk phase. We consider a five-player game with two malicious players. This satisfies the conditions in both [7] and [5] mentioned above. However, a table that shows a five-player game is very complicated to explain, so to simplify the example, we assume a dummy player as defined below.

Definition 4: Let δ_{-d} be the set of actions of the players other than the dummy player. A dummy player, P_d is a player who satisfies the following conditions,

1.His actions do not affect the other players' utilities.

$\forall \sigma_d,\ \acute{\sigma_d} \in A_d\ \forall \boldsymbol{\sigma}_{-d} \in \delta_{-d},\ u_{-d}(\sigma_d, \boldsymbol{\sigma}_{-d}) = u_{-d}(\acute{\sigma_d}, \boldsymbol{\sigma}_{-d})$.

2.His utility is not affected by the other players' actions except for a punishment strategy.

His utility is as defined below,

When a punishment strategy ρ_{-d} for P_d is taken, $\forall \sigma_d, \in A_d,\ \forall \boldsymbol{\sigma}_{-d} \in \delta_{-d}$, $u_d(\sigma_d, \boldsymbol{\sigma}_{-d}) > u_d(\sigma_d, \boldsymbol{\rho}_{-d})$.

otherwise,

$\forall \sigma_{-d},\ \acute{\sigma_{-d}} \in \delta_{-d} - \{\rho_{-d}\},\ \forall \sigma_d \in A_d$, $u_d(\sigma_d, \boldsymbol{\sigma}_{-d}) = u_d(\sigma_d, \acute{\boldsymbol{\sigma}_{-d}})$

An example of five-player game is shown as follows. $N = \{P_1, P_2, P_3, P_4, P_5\}$, we assume P_5 is a dummy player, so his actions do not concern us here. The number of malicious players is 2 (t=2), and for $1 \leq i \leq 4$, P_i's action set is $A_i = \{a^i_1, a^i_2, a^i_3, a^i_4\}$. The utility u_i is shown in Fig 2. Fig. 2 consists of 4×4 sub-tables. The utilities when P_4 takes a^4_i and P_3 takes a^3_j are shown in the sub-table at the i-th row and j-th column. In each sub-table, the actions by P_1 are mapped to the rows and the actions by P_2 are mapped to the columns. Each entry is a tuple of utilities, (u_1, u_2, u_3, u_4). The correlated equilibria for this game are $(a^1_3, a^2_3, a^3_3, a^4_2)$ and $(a^1_1, a^2_1, a^3_2, a^4_3)$. In these cases, the utilities of the players are (5,5,5,5), as indicated by the bold outlined boxes. Let us consider the case when P_3 and P_4 abort during the cheap talk phase. After aborting the protocol, they declare that they will take actions a^3_1 and a^4_1 using the chap talk, the rest of players are supposed to select the punishment strategy (a^1_4,a^2_4). As a result, the set of actions is (a^1_4,a^2_4,a^3_1,a^4_1), and each player will receive a utilities (u_1,u_2,u_3,u_4) = (3,3,3,3). The utilities for P_3 and P_4 decrease from the correlated equilibria. So, these utilities satisfy the definition of a punishment strategy (for all $C \subseteq$ N and all $\acute{\sigma}_C$ it holds that for all $i \in C$, $u_i(\acute{\sigma}_C, \rho_{-C}) \leq u_i(D)$). However, the important point is that the utilities of P_1 and P_2 also decrease from other strategies. If they try increase their utilities, they must give up taking the punishment strategy for P_3 and P_4 and select different actions that could increase the utilities of P_3 and P_4. If the players are honest, they will select a punishment strategy even if they receive worse utilities than the other strategies. However, in game theory, all players are considered to be rational, so if there is a better set of actions for P_1 and P_2, it is natural for them to select a better action than the Nash equilibrium. P_1 and P_2 know the actions that P_3 and P_4 will take and their utilities when they select a punishment strategy. Thus, the aborting players think that they will not execute the punishment strategy. This is called an "empty threat" [4]

In this example, given that P_3 and P_4 take a_1^3 and a_1^4 (in this case, we use the table at the upper left), we repeat iterative elimination of strictly dominated strategies for P_1 and P_2. From the viewpoint of P_2, strategies a_3^2 and a_4^2 are dominated by a_2^2, so we can remove the possiblity that P_2 takes a_3^2 and a_4^2. On the other hand, from the viewpoint of P_1, strategies a_1^1, a_3^1 and a_4^1 are dominated by a_2^1 and a_1^2 are dominated by a_2^2. Then, (a_2^1,a_2^2) is found to be the dominant strategy for P_1 and P_2. Note that both P_1 and P_2 can independently calculate the strategy by itself. So all players try to receive the maximum profits under the assumption that all players are rational, and (a_2^1,a_2^2,a_1^3,a_1^4) is the equilibrium for the all players. (As a result, they will receive utilities (4,4,5,6)). In this case, even if P_3 and P_4 abort protocols during the cheap talk phase, the other players will not punish aborting players, rather they will help aborting players to receive more payoff to get more payoff than that received from the punishment strategy. This is indicated by as an arrow in Fig. 2. The players will select the set of actions (a_2^1,a_2^2,a_1^3,a_1^4), not the punishment strategy (a_4^1,a_4^2,a_1^3,a_1^4).

V. New definition of punishment strategy

The reason a punishment strategy does not work is that definitions in [7] and [5] do not care about utilities of punishing players. In [4], it was shown that for two-player games, a min-max strategy may be an "empty threat" without the proper setting. For multiple player games, the above example shows that a punishment strategy does not work. To avoid these cases, we suggest a new definition of a punishment strategy that considers punishing players' utilities.

Definition 5: Let Γ be an n-player game with correlated equilibrium D. A strategy vector ρ is a t-punishment strategy if for any strategy vector $\acute{\rho}$ with respect to D and for all $i \in C \subseteq$ N, $j \notin C$ with $\mid C \mid \leq$ t, all $\acute{\sigma_C}$ it holds that $u_i(\acute{\sigma}_C, \rho_{-C}) \leq u_i(D)$ and $u_j(\acute{\sigma_C}, \acute{\rho}_{-C}) \leq u_j(\acute{\sigma}_C, \rho_{-C})$,
where $\acute{\sigma_C}$ satisfies the condition $u_i(D) < u_i(\acute{\sigma}_C, \rho''_{-C})$ for some strategy vector ρ''_{-C}.
We add the idea of punishing player (punisher) utilities to the original definition of the punishment strategy to avoid the case in which the utilities of the punishers decrease when they punish the aborting players. We also add the condition for $\acute{\sigma_C}$ to avoid the the case where punishment strategy becomes a dominant strategy for punishers. A punishment strategy is the dominant strategy for punishers only when some players abort in the cheap talk phase.

Theorem 1: Let Γ be an n-player game with correlated equilibrium D and the punishment strategy as defined above. Correlated equilibria can be implemented even in the presence of at most t malicious players under the condition that $n > 2t$.

Proof When some players C abort during the cheap phase, the other players try to punish them using the punishment strategy. Since all punishing players' utilities for the punishment strategy are not worse than other strategies, they will select the punishment strategy. Malicious players know that the rest of the players will choose the punishment strategy whenever they abort the protocol, and they are not supposed to deviate from the protocol.

VI. Conclusion

We showed that there are cases when a punishment strategy does not work. We suggested a new definition of a punishment strategy to avoid these cases. As future directions of investigation, we are exploring cases in different settings and searching for better definitions for rational multi party protocols.

References

[1] E. Rasmusen, "*Games and Information : An Introduction to Game Theory*," Blackwell Publishing 1994
[2] R. Aumann. "Subjectivity and correlation in randomized strategies". J. Mathematical Economics, Vol. 1, No. 1, pp.67-96, 1974.
[3] O. Goldreich, S. Micali, and A. Wigderson. "How to play any mental game". In Proceedings of the 19th ACM Symposium on Theory of Computing, pages 218-229, 1987.
[4] Y. Dodis, S. Halevi and T. Rabin "Cryptographic Solution to a Game Theoretic Problem". Crypto 2000, LNCS Vol.1880, pp.112-130, 2000.
[5] I.Abraham. D. Dolev, and J. Y. Halpern. "Lower Bounds on Implementing Robust and Resilient Mediators". Theory of Cryptography Conference(TCC), LNCS Vol. 4948, pp.302-319, 2008, full version http://arxiv.org/abs/0704.3646v2.
[6] O. Goldreich. "Foundations of Cryptography, vol. 2: Basic Applicationss". Cambridge University Press, 2004.
[7] J. Katz. "Bridging Game Theory and Cryptography: Recent Results and Future Directions". Theory of Cryptography Conference (TCC), LNCS Vol.4948, pp.251-272, 2008.
[8] I. Abraham, D. Dolev, R. Gonen, and J. Y. Halpern. "Distributed computing meets game theory: Robust mechanisms for rational secret sharing and multiparty computation". Proc. 25th PODC, pp. 53-62, 2006.

P3: a^3_1

P4: a^4_1

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	3	3	3
	2	3	3	3
	2	2	3	3
	2	2	3	3
a^1_2	3	4	4	2
	4	4	3	3
	3	5	3	3
	4	6	3	3
a^1_3	3	3	3	3
	4	3	3	3
	3	3	3	3
	3	4	4	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P3: a^3_2

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	2	4	6	3
	3	4	4	3
	4	3	4	3
	4	2	4	3
a^1_2	4	6	3	2
	6	5	2	3
	4	4	5	3
	4	3	5	3
a^1_3	4	4	6	3
	3	2	4	3
	4	4	4	3
	5	3	3	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P3: a^3_3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	6	2	4	3
	4	4	5	3
	5	3	6	3
	4	3	4	3
a^1_2	4	6	4	2
	3	5	4	3
	4	4	2	3
	2	3	4	3
a^1_3	4	2	4	3
	6	4	4	3
	4	2	3	3
	5	3	4	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P3: a^3_4

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_2	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_3	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P4: a^4_2

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	4	4	6	3
	4	4	4	3
	3	3	4	3
	3	3	5	3
a^1_2	2	6	4	3
	6	5	5	3
	3	4	2	3
	4	3	4	3
a^1_3	3	2	4	3
	6	4	4	3
	3	2	3	3
	4	3	4	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	4	4	6	3
	6	3	4	3
	4	3	4	3
	3	4	4	3
a^1_2	4	5	3	3
	3	5	4	3
	6	4	4	3
	4	4	4	3
a^1_3	2	4	3	3
	3	5	6	3
	4	3	4	3
	2	4	4	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	6	4	4	3
	4	4	4	3
	5	3	4	3
	4	3	4	3
a^1_2	4	6	4	3
	3	5	4	3
	4	4	2	3
	3	3	4	3
a^1_3	4	2	5	3
	4	4	5	3
	4	2	5	3
	5	3	5	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_2	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_3	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P4: a^4_3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	4	2	6	3
	6	3	4	3
	4	3	4	3
	4	4	4	3
a^1_2	4	5	3	3
	4	2	4	3
	6	4	5	3
	4	4	4	3
a^1_3	2	4	4	3
	2	5	6	3
	4	3	4	3
	3	4	4	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	5	2	4	3
	5	4	4	3
	5	3	4	3
	5	3	6	3
a^1_2	4	6	4	3
	3	5	4	3
	2	4	6	3
	4	3	4	3
a^1_3	4	2	6	3
	4	2	4	3
	2	2	3	3
	5	3	4	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	2	6	3
	6	2	4	3
	3	3	5	3
	4	4	4	3
a^1_2	4	5	3	3
	3	5	4	3
	6	6	4	3
	2	4	4	3
a^1_3	2	4	3	3
	3	5	6	3
	4	3	4	3
	3	4	4	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_2	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_3	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P4: a^4_4

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_2	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_3	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_2	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_3	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_2	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_3	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_2	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_3	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

Fig. 2. An example of game that a punishment strategy does not work.

P3: a^3_1
P4: a^4_1

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	3	3	3
	2	2	3	3
	2	2	3	3
	2	2	3	3
a^1_2	2	3	4	3
	4	3	2	3
	3	5	3	3
	4	6	3	3
a^1_3	2	3	3	3
	4	3	2	3
	3	3	3	3
	3	4	4	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P3: a^3_2

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	2	4	6	3
	3	4	4	3
	4	3	4	3
	4	2	4	3
a^1_2	4	6	3	3
	6	5	2	3
	4	4	5	3
	4	3	5	3
a^1_3	4	4	6	3
	3	2	4	4
	4	4	4	3
	5	3	3	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P3: a^3_3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	6	2	4	3
	4	4	5	3
	5	3	6	3
	4	3	4	3
a^1_2	4	6	4	2
	3	3	4	4
	4	4	2	3
	2	3	4	3
a^1_3	4	2	4	3
	6	4	4	3
	4	2	3	3
	5	3	4	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P3: a^3_4

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_2	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_3	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P4: a^4_2

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	4	4	6	3
	4	4	4	3
	3	3	4	3
	3	3	5	3
a^1_2	2	6	4	3
	6	5	5	3
	3	4	2	3
	4	3	4	3
a^1_3	3	2	4	3
	6	4	4	3
	3	2	3	3
	4	3	4	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	4	4	6	3
	6	3	4	3
	4	3	4	3
	3	4	4	3
a^1_2	4	5	3	3
	3	5	4	3
	6	4	4	3
	4	4	4	3
a^1_3	2	4	3	3
	3	5	6	3
	4	3	4	3
	2	4	4	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	6	4	4	3
	4	4	4	3
	5	3	4	3
	4	3	4	3
a^1_2	4	6	4	3
	3	5	4	3
	4	4	2	3
	3	3	4	3
a^1_3	4	2	5	3
	4	4	5	3
	4	2	5	3
	5	3	5	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_2	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_3	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P4: a^4_3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	4	3	6	3
	6	3	4	3
	4	3	4	3
	4	4	4	3
a^1_2	4	5	3	3
	4	2	4	3
	6	4	5	3
	4	4	4	3
a^1_3	2	3	4	3
	2	5	6	3
	4	3	4	3
	3	4	4	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	5	2	4	3
	5	4	4	3
	5	3	4	3
	5	3	6	3
a^1_2	4	6	4	3
	3	5	4	3
	2	4	6	3
	4	3	4	3
a^1_3	4	2	6	3
	4	2	4	3
	2	2	3	3
	5	3	4	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	2	6	3
	6	2	4	3
	3	3	5	3
	4	4	4	3
a^1_2	4	5	3	3
	3	5	4	3
	6	6	4	3
	2	4	4	3
a^1_3	2	4	3	3
	3	5	6	3
	4	3	4	3
	3	4	4	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_2	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_3	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P4: a^4_4

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_2	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_3	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_2	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_3	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_2	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_3	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

P2/P1	a^2_1	a^2_2	a^2_3	a^2_4
a^1_1	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_2	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_3	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3
a^1_4	3	3	3	3
	3	3	3	3
	3	3	3	3
	3	3	3	3

Fig. 3. An example of game that satisfies the new definition of punishment strategy.