Recommendation Extension to Iterated Prisoner's Dilemma Game

–Draft– v2019-10-06/20:11:53

Zeynep B. Cinar and Haluk O. Bingol

***Abstract*—We present an extended version of the Iterated Prisoner's Dilemma game in which each player receives recommendations about other players to decide whether to play with unknown opponents. An agent can receive more than one recommendations about the same opponent. They have to evaluate the recommendations according to their disposition such as optimist, pessimist, or realist. Since agents have limited memory, they have to use different forgetting mechanisms such as forget cooperators or defectors first. Although one expects that agents getting recommendations would perform better, our results show that this is not always the case. We observe that realist performs the best and optimist the worse.**

***Index Terms*—Multi-agent systems, Prisoner Dilemma, Recommendation Networks, Trust Behaviors**

I. Introduction

We live in a complex and crowded world. It is so crowded that it is impossible to "know" everybody. We know only a very small fraction of the population. We operate within this network of known people. As we interact with them, we classify them based on our firsthand experience. Next time we need to interact again, we use this information.

Since boldly interaction with an unknown person may not be a good idea, proceeding with first-hand experience is not the only way that we use to make a judgement about somebody. If we do not know a person, we ask for recommendations from people that we already know. We do that all the time in real life.

Z. B. Cinar and H. O. Bingol are with the Department of Computer Engineering, Bogazici University, Istanbul, 34342 Turkey e-mail: XXXX@boun.edu.tr.

Manuscript received XX XX 20XX; revised XX XX, 20XX.

It is common sense that having recommendation is better. In this work, we investigate whether having recommendation is always a better alternative in the context of iterated prisoner's dilemma game. Our findings indicate that it is not always the case.

II. Background

A. Optimism and Pessimism

There are different orientations in humans to suggest that what kind of things they could do in a given situation [1]. These orientations have an impact on trust. Marsh presents dispositions in terms of how an agent estimates trust [2]. Optimism, pessimism and realism are the notions of dispositions of trust. Each different disposition results in different trust estimations from an agent. An *optimist* expects the best in people and is always hopeful about the result of situations. A *pessimist*, unlike the optimist, sees the worst in people, always looks at the situations through doubting eye. While optimism and pessimism are extreme cases, there are other cases such as "realism".

B. Prisoner's Dilemma

Prisoner's Dilemma is a two-person game, in which players can choose between two different strategies: "Cooperate" or "Defect", without knowing their opponent's choice [3]. As summarized in Table I, if a person cooperates while the other defects, cooperator gets the sucker payoff S and defector gets the temptation payoff T. On the other hand, if both players choose to cooperate, then they both get the reward payoff R. Lastly, in the case of mutual defection, both players get the punishment payoff P. In Prisoner's Dilemma

TABLE I: Payoff matrix. $T = 5$, $R = 3$, $P = 1$, and $S = 0$.

		Player X	
		Cooperate	Defect
Player Y	Cooperate	(R, R)	(S, T)
	Defect	(T, S)	(P, P)

game, the payoffs should satisfy both $S < P < R < T$ and $S + T < 2R$.

For a single Prisoner's Dilemma game, it is more advantageous to defect. But, when the game repeats, things change. The iterated Prisoner's Dilemma is an extension of the standard Prisoner's Dilemma game. *Iterated prisoner's dilemma* differs from the original concept of a prisoner's dilemma because players can memorize the past interactions of their opponent and can change their strategy accordingly [3]. Players can learn about the behaviors of their opponent and have the opportunity to penalize the other for previous decisions.

C. Cetin and Bingol's Model

In the model proposed by Cetin and Bingol, agents play Iterated Prisoner's Dilemma game in which players can accept or refuse to play with their partner [4], [5]. There are N agents in the population. Agents are not pure cooperators or pure defectors. Agents have an internal parameter ρ. Agent i cooperates with probability ρ_i, which is called *cooperation probability*. There are two types of agents, one group, called *cooperators*, has a ρ value larger than 0.5. The other group, called *defectors*, has a ρ value, which is less than 0.5.

Decision to Play. Cetin adopts choice-and-refusal rule [6]: If an agent "knows" that the opponent is a defector, then it refuses to play. Otherwise, it plays. That is, two agents, say i and j, are randomly selected and offered to play Iterated Prisoner's Dilemma game. Both agents evaluate their opponent and decide whether to play or not. (i) If an agent does not know the opponent, then it has to play. (ii) If it "knows" the opponent as "cooperator", then it again plays. (iii) It i "knows" j as "defector", then i refuses to play.

Perception. In order agents to "know" each other, agents have some memory so that they can keep track of previous games with the same agent. Suppose agent i plays with agent j. Agent i keeps two numbers in its memory. Number c_j is the number of times that j cooperates and d_j is the number of times that j defects. Then the *perceived cooperation ratio* is defined as

$$t_{ij} = c_j/(c_j + d_j). \tag{1}$$

If the perceived cooperation ratio is larger than 0.5, i considers j as *cooperator*, otherwise as *defector*.

Memory. We assume that agents have the same memory capacity of $M \leq N$, called *memory size*. That is, each agent can keep track of at most M opponents. *Memory ratio*, defined as $\mu = M/N \in [0, 1]$, is the percentage of people that can be kept in ones memory.

Forgetting strategies. Every new opponent gets its own space in the memory. Eventually, the agent will run out of memory. After that point, to create memory space for a new opponent, the agent has to forget a known opponent. There are several forgetting mechanisms investigated in the model. (i) Players prefer to forget cooperators first, denoted by FC. (ii) Players prefer to forget defectors first, denoted by FD. (iii) Players prefer to forget randomly, denoted by FR [4].

III. Proposed Model

In Cetin and Bingol's model, if an agent does not know the opponent, it has to play [4], [5]. However, in real life, we use our social network to obtain information about a person that we do not know. We extend the model so that agents can get recommendations about the opponent that they do not know.

A. Perception Graph

Perceptions between agents can be represented as a directed graph as given in Fig. 1. In a *perception graph*, agents are represented by vertices. When i plays with j for the first time, two arcs will be created, namely, one from i to j, and one from j to i. The perceived cooperation ratio of j with respect to i, denoted by t_{ij}, is assigned to arc from i to j as a weight.

As the game proceeds, symmetry is broken as in the case of k and j. Suppose agents k and j played before. Hence arcs (k, j) and (j, k) were

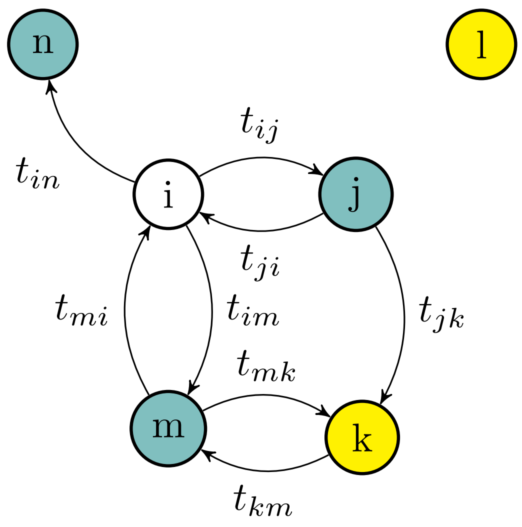


Fig. 1: Perception graph. Agents j, m and n are in the 1-neighborhood of i, while ℓ and k are not.

created. Later, when agent k chooses to forget j due to lack of memory space, arc (k, j) and its corresponding weight t_{kj} removed from the graph. Note that since j still keeps k in its memory, arc (j, k) and t_{jk} are intact.

B. Decision to play or not

Suppose i, in Fig. 1, is one of the selected agent to play. There are three cases for i's opponent. (i) i knows its opponent as in the case of j. Then all i has to do is check the perceived cooperation ratio. If $t_{ij} > 0.5$, it plays. However, if there is no data about the opponent in its memory, as in the cases of k or ℓ, agent i plays with it in Cetin and Bingol's model [5]. In our model, agent i asks for "recommendations" from its neighbors. There are two possible cases: (ii) Case of ℓ. If nobody in its neighborhood knows the opponent, i plays with ℓ. (iii) Case of k. Any neighbor, such as j, that knows k, provides its own perception t_{jk} about k. If there were only one such agent j, decision of i would be relatively easy: it will play if $t_{jk} > 0.5$. But usually, there are many such agents, e.g. j and m. Then i has to evaluate conflicting recommendations received from them.

C. Evaluation of recommendations

We define *1-neighborhood* of i, denoted by $\Gamma(i)$, as all agents to which there is an arc from i. Note that $\Gamma(i)$ is composed of agents in i's memory. Hence if an agent j removed from i's memory, it will also be removed from $\Gamma(i)$.

The set of recommender agents of i about k is denoted by $R_i(k) = \{j \mid j \in \Gamma(i) \text{ and } k \in \Gamma(j)\}$. For example, in Fig. 1, $R_i(k) = \{j, m\}$. Note that, n is not in $R_i(k)$ since n does not know k. Every agent in $R_i(k)$ gives a recommendation to i.

Once the agent receives the recommendations, the evaluation process begins. Evaluation of the recommendations by agents varies according to their dispositions. We consider three types of dispositions:

(i) *Optimists.* An optimist agent i takes the maximum of the recommendations that it receives, that is,

$$t_{ik} = \max\{t_{jk} \mid j \in R_i(k)\}.$$

(ii) *Pessimists.* A pessimist agent i takes the minimum of the recommendations that it gets, namely,

$$t_{ik} = \min\{t_{jk} \mid j \in R_i(k)\}.$$

Optimist and pessimist agents are the opposites of each other. We consider a third type in between.

(iii) *Realists.* A realist agent i takes the average of the recommendations, that is,

$$t_{ik} = \frac{1}{|R_i(k)|} \sum_{j \in R_i(k)} t_{jk}.$$

In the literature, there are realist agents that use the mode or the median of the recommendations, too [7].

(iv) *Self-assured (SA).* To compare our agents with the previous work [4], [5], we also consider agents that do not ask for recommendations.

D. Perceived Cooperation Ratio

We modify perceived cooperation ratio given in Eq. 1. According to Eq. 1, if agent j plays with i once and cooperates, t_{ij} will be 1. Similarly, if j plays with i, say 10 times and cooperates in all of them, t_{ij} is still equal to 1. However, in real life, the trustworthiness of a person who was honest with us only once and that of a person who was honest with us many times are not the same. With this idea, we propose a different approach to calculations of perceived cooperation ratio. *Perceived cooperation ratio* of j with respect to i is defined as

$$t_{ij} = \frac{c_j + 1}{(c_j + 1) + (d_j + 1)}. \quad (2)$$

With our new formula, if j plays with i once and cooperates, t_{ij} becomes 0.66. On the other hand, if j plays with i 10 times and cooperates in all of them, t_{ij} becomes 0.91. However, according to the perceived cooperation ratio calculation in [5], t_{ij} was calculated as 1 for both cases. We feel that Eq. 2 provides more realistic perception evaluation. Note that both Eq. 1 and Eq. 2 produces the same value of $t_{ij} = 0.5$ when $c_j = d_j$. Therefore, optimist and pessimist agents are not affected from this new definition of perceived cooperation ratio. On the other hand, realist agents are affected.

E. Metrics

We want to compare the performances of the cooperators and the defector at the end of the game. To do that we define *average payoff* of all agents in a set A as

$$\overline{P_A} = \frac{1}{|A|} \sum_{i \in A} \text{payoff}(i).$$

where payoff(i) denotes the accumulated payoff by agent i at the game end.

We have two types of agents. The set C of agents with cooperation probability $\rho > 0.5$ are considered cooperators. The rest of the agents are called defectors and denoted by D. Now, we can define *payoff ratio* as

$$\phi_C = \overline{P_C} \;/\; \overline{P_{C \cup D}}.$$

Note that we are interested in cases of $\phi_C > 1$, where cooperators are doing better than the average, that is, they perform better than defectors.

It could be the case that even if agents collect recommendations, they misjudge the opponent. The number of misjudgments of cooperators as defectors is denoted by η_{cd}. Similarly, η_{cd} is the number of misjudgments of defectors as cooperators. Then, the *accuracy ratio* of evaluations defined as the ratio of the total number of correct judgments to the total number of recommendations received, namely,

$$\delta = 1 - \frac{\eta_{cd} + \eta_{dc}}{\eta_r}.$$

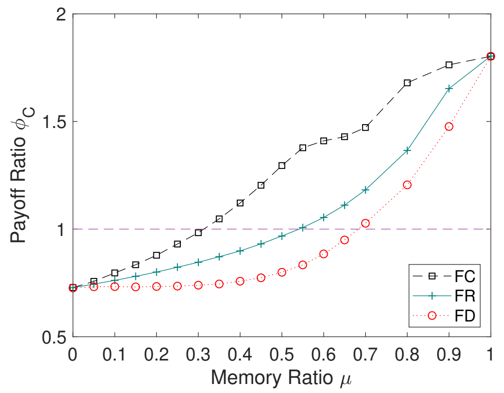


Fig. 2: Comparison of forget mechanisms for self-assured agents.

IV. Experiments

We run experiments for various memory ratios and report our findings as a function of μ.

We consider homogeneous agents, i.e. one type of cooperators with $\rho = 0.9$ playing against one type of defectors with $\rho = 0.1$.

In an experiment, both cooperators and defectors use the same forget strategy such as forget cooperators first.

Defectors are always *self-assured*, i.e. they do not ask for a recommendation. In contrast to defectors, cooperators do ask for a recommendation. Therefore, an experiment actually consists of four different simulations where cooperators are either (i) optimistic or (ii) pessimistic or (iii) realist or (iv) self-assured playing against self-assured defectors. We will be using plots such as Fig. 3(a) to compare performances of different dispositions.

In each experiment, there are $N = 100$ agents, where 50 defectors play against to 50 cooperators, i.e. 50 % cooperators. We terminate the experiments after $\tau\binom{N}{2}$ pairs invited to play as in the case of Cetin and Bingol's model [5], where $\tau = 30$. We report the average of 10 realizations.

A. Self-assured agents

First of all, we consider agents that do not get recommendations as in the case of Cetin and

Bingol model [4]. Fig. 2 agrees with the finding of Ref [4] that forgetting cooperators first (FC) is a better strategy compared to random forgetting (FR) or forgetting defectors first (FD). Even for this strategy, in order for cooperators to go above average payoff, i.e. $\phi_C = 1$, considerable memory ratio, $\mu = 0.3$, is required. Strategies forget randomly and forget defector first call for more than 50 % memory ratio for $\phi_C > 1$.

B. Forget Cooperators First

We investigate forgetting cooperators first (FC) strategy since it is better compared to FR and FD strategies according to Fig. 2 and Ref [4]. Cooperators of a disposition play against self-assured defectors. Both cooperators and defectors use forget cooperators first strategy, That is, if there is no memory space left, they forget a randomly selected cooperator first. If there is no cooperator left in the memory, then randomly select a defector to forget.

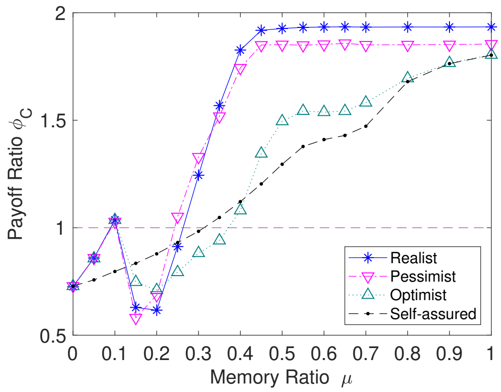
Performance. As expected, Fig. 3(a) shows that self-assured agents increase their performance as memory ratio increases. Note that the self-assured curve in Fig. 3(a) and the FC curve in Fig. 2 are the same.

Interestingly, in Fig. 3(a), we observe unexpected fluctuations in the performances of the optimists, pessimists and realists around $\mu = 0.2$. Instead of increase, they decrease. Note that, although agents receive recommendations, they perform worse than self-assured agents in this region.

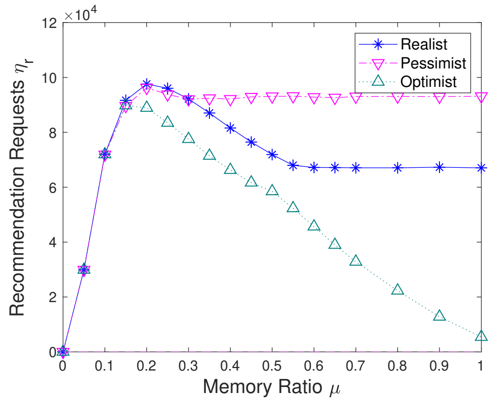
For values of $\mu > 0.2$, performances start to increase again but different dispositions have different trends. Pessimists and realists are quick to recover and pass $\phi_C = 1$ threshold around $\mu = 0.25$. They reach their peak values on around $\mu = 0.45$ and stay there. Optimists have a different path. They cross $\phi_C = 1$ threshold around $\mu = 0.35$. Their performance steadily increases and takes its peak value close to $\mu = 1$.

Number of requests. In order to understand the strange behavior in performance, we collect data on recommendations.

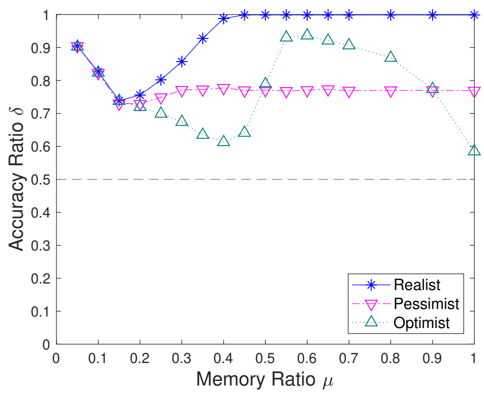
There are recommendation requests, which receive no response. Then, agent has to play. If it receives at least one response, then it acts accordingly. Fig. 3(b) plots η_r, the number of



(a) ϕ_C - Payoff ratios of agents.



(b) n_r - Number of recommendation they receive.



(c) δ - Accuracy ratios of evaluations of agents

Fig. 3: Detailed analysis of forgetting cooperators first. (a) describes payoff ratios of agents; (b) describes the number of recommendation they receive; (c)describes the accuracy ratios of evaluations of agents

recommendation requests that received at least one response. For all three dispositions, η_r increases as μ increases but it reaches its maximum around $\mu = 0.2$. Then different dispositions present different behaviors. Pessimists keeps the same level around $\phi_C = 9.5 \times 10^4$ for $\mu > 0.3$. For realists, the number of recommendation requests decreases till $\mu = 0.55$ and stays around $\phi_C = 6.5 \times 10^4$. For optimists, after its early peak value at $\mu = 0.15$, the number of recommendation requests decreases almost to none.

In Fig. 3(c), δ values of cooperators are shown. Initially, accuracy decrease as the memory of the agents increases for all dispositions. Around $\mu = 0.15$, dispositions start to deviate. For realists, accuracy increases and reaches its maximum value of perfect accuracy, i.e. $\delta = 1$, around $\mu = 0.4$. Accuracy of pessimists increases slightly and stays just below $\delta = 0.8$ starting from $\mu = 0.3$. Behavior of optimists is the most difficult one to explain. It keeps decreasing to just above $\delta = 0.6$ till $\mu = 0.4$, then it has a sharp increase that reaches to above $\delta = 0.9$ at $\mu = 0.55$, and then has a smooth decrease back to $\delta = 0.6$.

V. Discussions

Note that the way we set the experiment, cooperators are of one type of disposition only. For example, for a realist, all cooperators are also realist. That is, any recommendations that are received are coming from realists. And they are evaluated by the agent which is also a realist.

Optimist, by their nature, have an optimistic view of life. One positive recommendation is good enough for them to play. Once they play with an agent, they record this firsthand experience in their memory. Because of this, they do not need to ask recommendations for the same agent again. This explains the steady decrease of recommendation requests in Fig. 3(b) for $0.2 < \mu < 1$. For $0.2 < \mu < 0.4$ region in Fig. 3(c), its accuracy keeps decreasing, too. That is, it requests less recommendation and makes bad judgements. Yet it knows enough defectors so that it can keep its payoff ratio increasing.

Pessimists have the opposite strategy in playing. One single negative recommendation is enough for pessimist not to play. If they do not play, then there will be no record of the opponent kept in their memory. If the same opponent is matched to play again, a pessimist has to ask recommendation once more. This explains the high number of queries of recommendation even at $\mu = 1$, where there is enough memory to keep all the population.

Realists are in the middle ground of optimists and pessimists. They play more than pessimists but less than optimists. In the region of $0.2 < \mu < 0.6$, similar to optimists, they play with new agents, but after that, they stop to play. From that point on, their actions are similar to that of pessimists, that is, reject to play and keep asking the same agent over and over again.

We further investigate the reasons for this surprising behavior in the region $0.15 < \mu < 0.25$. First, we ask whether they receive recommendations or not? By checking η_r values around $\mu = 0.2$, one can observe that the optimist, pessimist and realist agents received sufficient recommendations. Then, we need to investigate how accurately they evaluate those recommendations. One observes that accuracy declines as μ go from 0.1 to 0.2, which can explain the drop in performance ϕ_C in the same region.

In this region around $\mu = 0.20$, memory is quite small. One can hold at most 25 % of the population while 50 % of it is defector. Because of the forgetting strategy of FC, one keeps perceived defectors in the memory. Therefore any recommendation given would be towards no play. If a neighbor misperceives a cooperator as defector, this misjudgment prevents others to play with it. This escalation prevents others to play with this cooperator and gain points.

VI. Conclusion

We investigated Iterated Prisoner's Dilemma game where agents get recommendations if they do not know the opponent. Although we expect better performance as memory capacity increases, performances of all dispositions drops around 20 % memory ratio region. After that performances recover. Realists perform the best. Performance of the optimists is the worse. They barely outperform agents which do not get any recommendation at all.

In this work, we report, in detail, strong cooperators and defectors with $\rho = 0.9$ and $\rho = 0.1$. We also investigate mildly cooperators and defectors such as $\rho = 0.75$ and $\rho = 0.25$, and we obtained similar results.

This work can be expanded in various ways. In our current model, recommender agent gives its sincere perception as a recommendation. That is, even a defector agent provides its genuine opinion. This may not be the case in real life. We considered $N = 100$ with 50 % cooperators. One wants to try larger N values as well as different percentages. We consider homogenous types. One types of cooperators play against one types of defectors. In real life, there are always mixtures of all kinds. In such heterogeneous environments are difficult to investigate but definitely much more realistic.

References

[1] D. Susan and J. G. Holmes, "The dynamics of interpersonal trust: Resolving uncertainty in the face of risk," *Cooperation and Prosocial Behavior; Cambridge University Press: New York, NY, USA*, p. 190, 1991.

[2] S. Marsh, in *Optimism and Pessimism in Trust*. McGraw-Hill Publishing, 1994.

[3] R. Axelrod and W. D. Hamilton, "The evolution of cooperation," *Science*, vol. 211, no. 4489, pp. 1390–1396, 1981.

[4] U. Cetin and H. O. Bingol, "Iterated prisoners dilemma with limited attention," *Condensed Matter Physics*, vol. 17, no. 3, p. 33001, 2014.

[5] ——, "The dose of the threat makes the resistance for cooperation," *Advances in Complex Systems*, vol. 19, no. 08, p. 1650015, 2016.

[6] E. A. Stanley, D. Ashlock, and M. D. Smucker, "Iterated prisoner's dilemma with choice and refusal of partners: Evolutionary results," in *Proceedings of the Third European Conference on Advances in Artificial Life*. London, UK, UK: Springer-Verlag, 1995, pp. 490–502. [Online]. Available: http://dl.acm.org/citation.cfm?id=645300.648534

[7] S. P. Marsh, "Formalising trust as a computational concept," Department of Computing Science and Mathematics University of Stirling, Tech. Rep., 1994.