Effect of defectors for cooperation: How strictly should defectors be eliminated from the newcomers?

(Extended Abstract)

Hitoshi Yamamoto
Faculty of Business Administration, Rissho University
3-2-16, Osaki, Shinagawa-ku
Tokyo, Japan
hitoshi@ris.ac.jp

Isamu Okada
Faculty of Business Administration, Soka University
1-236 Tangi-cho, Hachioji City
Tokyo Japan
okada@soka.ac.jp

Yuki Ogawa
Center for Service Research, AIST
Central2, 1-1-1 Umezono, Tsukuba-shi
Ibaraki, Japan
yuki.ogawa@aist.go.jp

ABSTRACT

Yamamoto et. al.[6] have discovered that cooperation can be robustly maintained in a metanorms game by introducing into the population a small number of agents that always act non-cooperatively. They call this a "social vaccine" effect. In this paper we focus on the implications of a social vaccine. We therefore consider a model where there is a constant flow of newcomers into the population. How strictly should non-cooperators be eliminated from the newcomers in such a model? In this paper, by assuming a case where cooperative participants and non-cooperative participants are trying to participate in a population where metanorms are functioning, we investigate how well cooperation within the population is maintained by a strict population management policy where only cooperative participants are allowed to participate, and a simple population management policy where non-cooperative participants are admitted to some extent.

Categories and Subject Descriptors

I.6.6 [SIMULATION AND MODELING]: Simulation Output Analysis

General Terms

Design

Keywords

Social vaccine, Meta-norms, Evolution of cooperation, Agent-based simulation, Public goods game

1. METANORMS GAME WITH NEWCOMERS

The metanorms game[1] is a well-known model for maintaining norms in a population. As an extension of the n-person prisoner's dilemma, this game provides an excellent model for studying how norms are maintained in a population without a centralized authority, such as problems involving cooperation on international affairs. For example, Heck[2] and Horne[3] performed a psychological experiment in which metanorms were shown to exist.

We consider a situation where there is an influx of new participants into a population. Assuming a case where cooperative participants and non-cooperative participants are attempting to join a population, we discuss what sort of control policy is effective for maintaining cooperation when the population operates a strict control policy of only admitting cooperative participants, and when it operates a simple control policy where non-cooperative participants are also admitted to some extent. It could be said that this is a highly abstracted model of the problem of whether the stability of a society is more effectively maintained by adopting an immigration policy of only admitting people who have a strong affinity with the country's policies, or by adopting a lenient policy and accepting some degree of risk.

We consider groups on a social network with a population size of 100. However, a metanorms model featuring mutual surveillance among all members of a group leads to an upper limit in the number of group members due to cognitive limits and that a system of mutual surveillance is an unrealistic, severe restriction. In light of these criticisms, extending the metanorms game to a partial group[5] and limiting the study to mutual surveillance in a small world network[4] have been proposed. In order to understand the basic properties of the model, the initial state of the network structure is assumed to be a non-oriented regular network. The average degree of the population is D.

The agents play a metanorms game on a network where they are all interconnected by links. An agent j that has been defected by agent i and has received a payoff of H is an agent with a link to agent i, and an agent capable of punishing agent i must also have a link to this agent. In the evolution process, agents that are capable of becoming the parent of each agent must also be linked agents.

At the stage where the first generation of the metanorms game has completed, the F agents with the lowest payoffs in the population are withdrawn. An equal number of agents are then admitted to the population. The strategies of these newly admitted agents are discussed below. Newcomer agents are linked with randomly selected exist-

Appears in: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012), Conitzer, Winikoff, Padgham, and van der Hoek (eds.), 4-8 June 2012, Valencia, Spain.

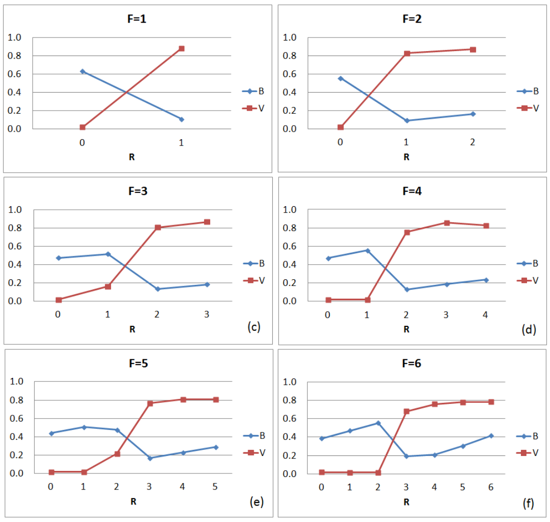

Figure 1: Effect of a social vaccine in newcomers

ing agents up to the average degree (D).

A wide variety of policies can be considered with regard to how newcomers are controlled, but in this paper we assume a simple model to observe the effects of the social vaccine. Here, newcomers are assumed to be either good or bad. In this context, a bad newcomer refers to a social vaccine. The population control policy is expressed as a level of rigor ranging from a strict monitoring policy where bad newcomers are never admitted, to a lenient policy where a blind eye is turned to these admissions to some extent.

Specifically, the strategy of good newcomers is taken to be $(B, V) = (0, 0)$, and the strategy of bad newcomers is taken to be $(B, V) = (1, 0)$ where B and V stand for boldness and vengefulness, respectively, as well as Axelrod[1]. The degree of rigor is expressed as the number R out of F newcomers for which a blind eye is turned to the admission of bad individuals ($R \leq F$).

For the payoff parameters of this section, we used the same values as in the Axelrod's experiment[1], i.e., $T = 3, H = -1, E = -2$ and $P = -9$. Also, the strategies (B, V) of first-generation agents are each given by uniform random numbers.

2. RESULTS

We analyze the effects of control rigor on the numbers and influx of newcomer agents, with the average degree D fixed at 20 the same as the population size of Axelrod's basic model (Fig. 1). Each of the graphs in Fig. 1 shows a plot of R on the horizontal axis and the average values of B and V of a population at the end of the simulation on the horizontal axis, for values of F ranging from 1 through 6.

When the number of newcomers is $F = 1$ (Fig. 1(a)), in the state where $R = 0$ - i.e., where bad individuals are completely prevented from entering the population and only good people can enter - it can be seen that cooperation is not achieved. However, when $R = 1$ - i.e., when newcomer agents adopt a defection strategy - cooperation is achieved at a high level. Similarly, when $F = 2$ (Fig. 1(b)), cooperation is not achieved when $R = 0$ but is achieved when bad agents are admitted. A similar trend was observed for $F \geq 3$, but with a gradual increase in the threshold value of R for which the social vaccine functions effectively. For example, when the number of newcomer agents is $F = 4$ (Fig. 1(d)), a value of $R = 1$ indicates the state where one bad individual (social vaccine) enters the population, while the other new entrants are all good. In this case, cooperation is not achieved. However, cooperation is achieved when $R = 2$. Cooperation is also maintained for $R = 3$ and $R = 4$, albeit not to as great an extent as for $R = 2$.

A characteristic feature of these experimental results is that the value of V for the population (i.e., its vindictiveness with regard to defection) differs widely in the vicinity of the threshold value at which cooperation is achieved, and is maintained at a high level in environments where cooperation is achieved. Specifically, the admission of a certain level of social vaccine into a population prevents the value of V for the group from decreasing, and as a result realizes a society that is robust against defection. This phenomenon resembles the immune function of resistance to a pathogen whereby inoculation with a weakened pathogen leads to the creation of antibodies to the pathogen. We call this a "social vaccine" effect.

3. CONCLUSION

We assumed a state where there is a constant influx of new participants into the population in a metanorms game, and we analyzed what sort of admissions policy the population should apply to newcomers in order to ensure that cooperation is robustly maintained. For simplicity, we expressed the admissions policy as the degree to which a blind eye is turned to the admission of uncooperative agents when fully cooperative agents and fully non-cooperative agents are both trying to enter the population.

In simulation experiments, we found that cooperation collapses in populations where entrants are subject to constant rigorous monitoring so that only cooperative agents are allowed to enter, but conversely cooperation is maintained at a high level when the entry of non-cooperative agents is overlooked to some extent.

4. REFERENCES

[1] R. Axelrod. An evolutionary approach to norms. American Political Science Review, 80(4):1095–1111, 1986.

[2] D. Heckathorn. Collective sanctions and compliance norms: A formal theory of group-mediated social control. American Sociological Review, 55(3):366–384, 1990.

[3] C. Horne and A. Cutlip. Sanctioning costs and norm enforcement: An experimental test. Rationality and Society, 14(3):285–307, 2002.

[4] D. Newth. Altruistic punishment social structure and the enforcement of social norms. In R. Khosla et al. (Eds.): KES 2005, LNAI 3683, pages 806–812, 2005.

[5] M. Prietula and D. Conway. The evolution of metanorms: quis custodiet ipsos custodes? Computational and Mathematical Organization Theory, 15(3):147–168, 2009.

[6] H. Yamamoto and I. Okada. Vulnerability and evolution of cooperation in the metanorms game. In Proceedings of 3rd World Congress on Social Simulation, 2010.