

# Chaos Behavior in Group Formation

Kodai Shinkawa and Isamu Shioya

**Abstract**—This paper considers a basic autonomous multi-agent model of group formation, whose agents carry feature vectors and meet each other at random in a free moving space. The agents combine to form groups, when their feature vectors are matched or their compatibilities are higher than a certain threshold. Similarly, if the compatibilities of the feature vectors between two groups are higher than the threshold not only between two agents but also between groups, the two groups are united into one. On the other hand, forming groups reduces a satisfaction of the groups by getting groups larger. This paper experimentally shows that, for a given threshold, there is an optimal threshold that maximizes the amount of satisfactions among groups, while the compatibility keeps the threshold, and the optimum threshold is related to the size of feature vectors, but the optimum threshold is independent to the number of agents.

**Index Terms**—Multi-agents, group formation, optimum threshold on satisfaction.

## I. INTRODUCTION

This paper discusses an autonomous multi-agent model for group formation, whose agents meet each other at random in a free moving space and combine to form groups if their feature vectors are matched or are compatible. Note that each agent carries a feature vector, and assumes that when two agents meet each other in the space and their compatibilities are higher than a certain threshold, they form groups and are unified into one. That is, if the degrees of conformity between two groups are higher than a threshold not only for each agent but also between groups, the groups are united into one group. In this paper, we first define the compatibilities based on inner product between two feature vectors of agents, and extend this definition to the compatibilities between agent groups, that is, to a set of agents. On the other hand, forming groups reduces the satisfactions of the groups by getting the size of groups larger. Providing a threshold, so-called "ideal height", allows to form groups. When the threshold value is higher, the ideal height among groups is also higher. Conversely, the ideal height is lower when the threshold is lower. Forming groups is to become the satisfaction of groups lower. Then, it is difficult to keep groups if the ideal is too either high or low. This paper shows that there is an optimal threshold value that maximizes the amount of the satisfaction values among the groups. By appropriately setting the threshold value, we show that the amount of the goodness of fit among the group becomes highest. It indicates that there is an optimal threshold. This work is closely related to the work Smith and Neil [1] on two kind

of groups formation. Another related work is presented in [2] on the community of DBLP. Our model is quite the opposite to Schelling segregation model [3], and negative elements attract gatherings so that the satisfaction increases. On the other hand, in our model, while similar items make groups, their satisfaction decreases.

In this paper, we show our related works in the following section. In Section III, we propose our model and present an example in the following section. In Section IV, we show our experimental results, our related works, and our discussions in the following. The final section presents our conclusions in this paper.

## II. RELATED WORKS

The closely related work is [1], which discusses two kinds of formation for looking another kind of partner in the same as a free space. The partners keep finite time before disassociating, the compatibility of two agents is defined as well, and the goal is to maximize the duration of two pair. The work [2] discusses the actual world on the community of DBLP or around. By using the techniques of decision trees, time series and diffusion, they find several aspects to join the community so that to join the community is not only related the number of friends, but also how those friends connected to other friends. Also, they are related to their topics of interest. But, the community is complicated for obtaining explicit results how a lot of factors are interrelated. The work [3] is not only a simple model to produce a segregation of groups as well as our work, but the results are also simple. These features are sometimes complicated as [4], [5]. However, our real world behaves collective dynamics [6] and has scaling feature [7] as we would try in the next.

The opposite one to ours is Anderson [8]. The complexities of systems quite come to be complicated. But, increasing the number of agents is independent to highest satisfaction. In other words, the internal of systems is complicated, but the outcomes of the systems are different. Note that, in our work, there are interactions among agents or their groups, but the results show that the satisfaction is independent to the number of agents. Also, our work is closely related to clustering [9]. A clustering is to collect similar objects into clusters for grasping the overviews of objects, and the processes are to produce recognized clusters. The difference between clustering and group formation is discussed in [1].

## III. PROPOSED MODEL

We suppose that agents  $A_1, A_2, \dots, A_N$  have a fixed  $S$ -dimensional feature vector  $x_i$ , and are moving at random in a free moving space. Then, the agents form groups consisting

Manuscript received November 21, 2018; revised February 17, 2019.

Kodai Shinkawa and Isamu Shioya are with the Faculty of Science and Engineering, Hosei University, Kajino-cho 3-7-2, Koganei-shi, Tokyo, Japan (e-mail: kodai.shinkawa.5z@stu.hosei.ac.jp; shioyai@hosei.ac.jp).

some agents if their agents meet at random in the space, and their feature vectors match or the similarity value is higher than a certain threshold, i.e. we consider a model in which agents with high adaptability between two agents form a group.

Suppose two agents  $A_i$  and  $A_j$  ( $i \neq j$ ), and they have two feature vectors  $\mathbf{x}_i$  and  $\mathbf{x}_j$ , respectively. For the simplicity, suppose each element of their vectors is either 1 or -1 throughout of this paper. Then we define a compatibility  $C'_{ij}$  between  $A_i$  and  $A_j$  as the following:

$$C'_{ij} = (\mathbf{x}_i \cdot \mathbf{x}_j + S)/2,$$

where  $\cdot$  denotes an inner product between two vectors, and  $S$  is the size of feature vectors. There, we have  $C'_{ij} = C'_{ji}$  and  $0 \leq C'_{ij} \leq S$ ,  $i, j=1, \dots, N$ . Also,  $C'_{ij}$  follow a binomial distribution Bin ( $S, 1/2$ ). The higher similarity between two  $\mathbf{x}_i$  and  $\mathbf{x}_j$  becomes the higher compatibility  $C'_{ij}$ , and vice versa.

Initially, we suppose all agents belong to one of the groups  $G_1, \dots, G_N$ , and we have

$$G_i \cap G_j = \emptyset, i \neq j, i, j=1, \dots, N, \text{ and}$$

$$\bigcup_{k=1}^M G_k = \{1, \dots, N\}.$$

When the number of agents and groups is  $N$  and  $M$ , respectively, the variance  $v$  of the group size is defined as:

$$v = \frac{1}{M} \sum_{k=1}^M (|G_k| - \frac{N}{M})^2.$$

Further, we extend the degree of conformity between agents to the degree of conformity between groups

$$C_{ij} = \frac{\sum_{k \in G_i, l \in G_j, k < l} C'_{k,l}}{|G_i||G_j|}$$

$C_{ij}$  denotes a compatibility between groups  $G_i$  and  $G_j$ , and is the amount of the compatibilities of all the agents belonging to the groups.

Similar to  $C'_{ij}$ , we have  $C_{ij} = C_{ji}$  and  $0 \leq C_{ij} \leq S$ . When two groups  $G_i$  and  $G_j$  meet in the spaces and their compatibility  $C_{ij}$  is higher than a threshold  $\tau$  ( $0 \leq \tau \leq S$ ), the two groups is united into one group, and a new group is generated. We note that, when  $C_{ij} = \tau$ , two groups  $G_i$  and  $G_j$  is unified into one with the probability 1/2.

A satisfaction  $S_i$  of a group  $G_i$  is defined as following:

$$S_i = \sum_{k, l \in G_i, k < l} \mathbf{x}_k \cdot \mathbf{x}_l.$$

We assume that the satisfaction of  $G_i$  is zero, if  $G_i$  is a singleton.  $S_i$  is the amount of inner products of two agents on each  $i$  group, and the amount of the satisfactions for every groups is defined as:

$$S_{\text{sum}} = \sum_{i=1}^M S_i$$

By varying the threshold  $\tau$ , we can find the optimal  $\tau_{\text{op}}$  to maximize  $S_{\text{sum}}$ .

**Example:** We suppose the five agents  $A_1, A_2, A_3, A_4$  and

$A_5$ , the dimension of their vectors is  $S = 3$ , and their vectors are defined as

$$\mathbf{x}_1 = [1, -1, 1],$$

$$\mathbf{x}_2 = [-1, 1, 1],$$

$$\mathbf{x}_3 = [1, -1, 1],$$

$$\mathbf{x}_4 = [1, 1, 1],$$

$$\mathbf{x}_5 = [1, 1, -1].$$

Their compatibilities can be calculated, for an example,  $C'_{2,5}$  is as:

$$C'_{2,5} = (\mathbf{x}_2 \cdot \mathbf{x}_5 + 3)/2 = 1.$$

At the initial agent placement, each agent belongs to a group having only its own group as follows:

$$G_1 = \{1\}, G_2 = \{2\}, G_3 = \{3\}, G_4 = \{4\}, G_5 = \{5\}.$$

We set the threshold  $\tau = 2$ . Each member of  $G_1$  and  $G_3$  meet at random in a free space and merge into one by  $C_{1,3} = C'_{1,3} = 3 > \tau$ , and the following groups are formed:

$$G_1 = \{1, 3\}, G_2 = \{2\}, G_3 = \{4\}, G_4 = \{5\}.$$

Also, if each member of  $G_2$  and  $G_3$  meet and  $C_{2,3} = C'_{2,4} = 2 = \tau$ ,  $G_2$  and  $G_3$  form a new group with probability 1/2. Then, the current groups are:

$$G_1 = \{1, 3\}, G_2 = \{2, 4\}, G_3 = \{5\},$$

The compatibilities between groups are as:

$$C_{1,2} = \frac{C'_{1,2} + C'_{1,4} + C'_{3,2} + C'_{3,4}}{2 \times 2} = 1.5 < \tau,$$

$$C_{1,3} = \frac{C'_{1,5} + C'_{3,5}}{2 \times 1} = 1.0 < \tau, \text{ and}$$

$$C_{2,3} = \frac{C'_{2,5} + C'_{4,5}}{2 \times 1} = 1.5 < \tau.$$

Forming groups arrive inactive, anymore, and the compatibilities of their groups are:

$$S_1 = \mathbf{x}_1 \cdot \mathbf{x}_3 + \mathbf{x}_3 \cdot \mathbf{x}_1 = 6.0,$$

$$S_2 = \mathbf{x}_2 \cdot \mathbf{x}_4 + \mathbf{x}_4 \cdot \mathbf{x}_2 = 2.0, \text{ and}$$

$$S_3 = 0.$$

The variance of group size is:

$$v = \frac{1}{3} \sum_{k=1}^3 (|G_k| - \frac{5}{3})^2 \cong 0.22$$

The amount  $S_{\text{sum}}$  of the compatibility of the agents is:  $S_{\text{sum}} = S_1 + S_2 + S_3 = 8.0$

The final group composition may differ depending on the order of the groups to meet. In this research, the groups randomly meet in a free space, and the amount  $S_{\text{sum}}$  does not depend on the number of groups.

## IV. EXPERIMENTS

We simulated the formation of groups by agents and examined the thresholds that maximize  $S_{\text{sum}}$ . The following are our algorithm:

- 1) Make thirty agents, i.e.,  $N=30$ . Each element of all the feature vectors is either 1 or -1 with even probability.
- 2) The groups meet in the free space randomly and are unified into one if the compatibility of  $G_i$  and  $G_j$  is greater than a given threshold  $\tau$ . If the threshold is equal to  $\tau$ , they are unified into one with the probability  $1/2$ .
- 3) Repeat (2) until forming groups are inactive.
- 4) No longer group formation is impossible, and we calculate the amount of satisfiability of  $S_{\text{sum}}$ .

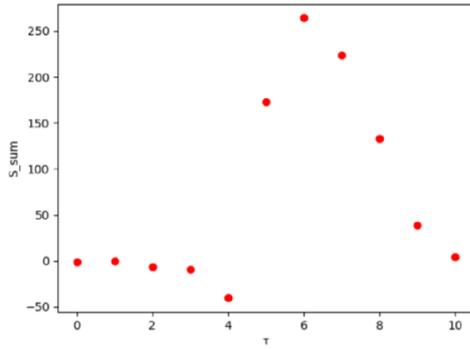


Fig. 1. The experiment of the case  $N=30, S=10$ .

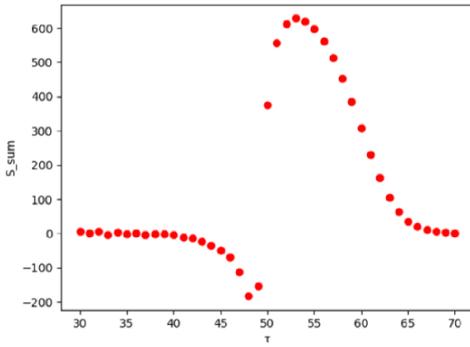


Fig. 2. The experiment of the case  $N=30, S=100$ .

For every  $\tau$ , we tried 1,000 times, and their average values  $S_{\text{sum}}$  of  $N=30$  and  $S=10$  are shown in Fig. 1, and are also shown in Fig. 2 for  $N=30$  and  $S=100$ . Their figures show that there is an optimum threshold  $\tau_{\text{op}}$  for each so that the compatibilities  $S_{\text{sum}}$  can be maximized experimentally, approximately. We experimentally say that the threshold  $\tau_{\text{op}}$  only depends on the size of feature vectors, i.e. the value is the sum of half size of feature vectors and the square root of the dimension  $S$  of feature vectors divided by 3. And, we can see that the  $\tau_{\text{op}}$  is independent to the number of agents.

$$\tau_{\text{op}} = \frac{S}{2} + \frac{\sqrt{S}}{3}$$

Similarly, we tried 1,000 times for each  $\tau$  and the average values of the variance on the group size is shown in Fig. 3. The figure shows that the variances  $v$  starts to increase near the median  $S/2$  of the feature vector size as Chaos behavior [3], [4]. The left shows the stable stagnated region for forming groups, and there whole consists of one group. And

it's also right, but it's just different to left. In the right, the number of groups is proportionally increasing to the threshold shown in Fig. 4, i.e., it's growth smoothly. However, in the middle region around  $\tau=50$ , the variances of the number of groups violently fluctuates.

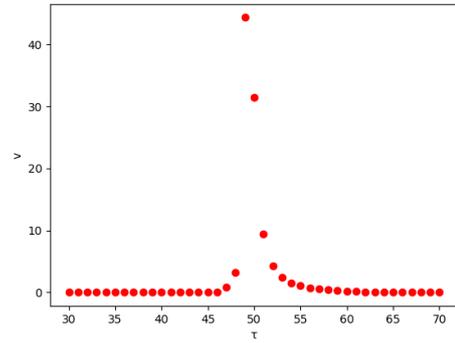


Fig. 3. The variances of the case  $N=30, S=100$ .

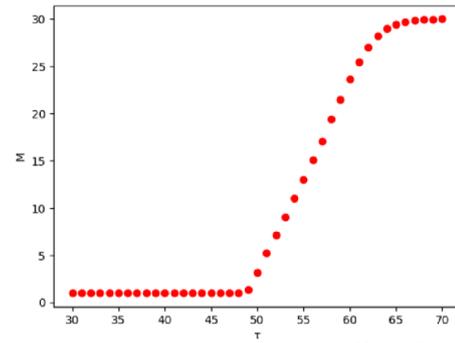


Fig. 4. The average number of groups of the case  $N=30, S=10$  for each  $\tau$ .

## V. DISCUSSIONS

The compatibilities  $C'_{ij}$  follow binomial distribution  $\text{Bin}(S, 1/2)$ , so the average  $S/2$  and the variance  $S/4$  follow along normal distribution. Therefore, if you set the threshold  $\tau$  to the value of the expression shown in the experiment, some agents  $A_i$  and  $A_j$  will merge with a probability of about 25%. In our research, the probabilities to meet two agents in the spaces are not depending on the sizes of the groups. Therefore, after unifying two groups into one, the probabilities to meet other groups in the spaces varies. Our next work is to obtain theoretical value of  $S_{\text{sum}}$  by considering a model in which the probability to meet is dependent on the size of the groups.

Why the satisfaction is negative in low fifty of Fig. 1 and Fig. 2? Because, some similar agents values are classified into different groups. For example, consider two groups A and B, and closer agents are classified into different. Then the threshold is loose, so the several inner products become negative. If the threshold is further relaxed, the satisfaction becomes 0, because it becomes one group.

Our experiments set the threshold by every step to 1. More accurate optimum threshold values are obtained by segmenting the step size finely.

We only discussed on the amount of satisfaction. The similar discussions are given, i.e., an average satisfiability on groups can discuss as well as this paper. Finally, we note that our results differ in the combination of groups that is created at each experiment as well as clustering.

## VI. CONCLUSIONS

This paper experimentally showed that, when autonomous agents are carrying feature vectors and they meet at random in moving free spaces forming groups, the optimum thresholds exist. For our future works, since our experiment was conducted on a simple network that the agent groups randomly meets in spaces, we will perform on more complex networks [6], [7]. Furthermore, we consider the model in which the probability of meeting depends on the size of the groups, and we will obtain the theoretical expression of  $S_{\text{sum}}$ .

## REFERENCES

- [1] D. M. D. Smith and N. F. Johnson, "Pair formation within multi-agent populations," *Physica A: Statistical Mechanics and Its Applications*, vol. 363, no. 1, pp. 151-158, 2006.
- [2] L. Backstrom, D. Huttenlocher, J. Kleinberg, and X. Y. Lan, "Group formation in large social networks: Membership, growth, and evolution, KDD," in *Proc. the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2006, pp. 44-54.
- [3] S. Strogatz, *Nonlinear Dynamics and Chaos*, Perseus Books, 2000.
- [4] Y. Kuznetsov, *Elements of Applied Bifurcation Theory*, Springer, 1998.
- [5] T. Schelling, "Dynamic models of segregation," *Journal of Mathematical Sociology*, vol. 1, pp. 143-186, 1971.

- [6] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, pp. 440-442, 1998.
- [7] L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509-512, 1999.
- [8] P. W. Anderson, "More is different," *Science*, vol. 177, no. 4047, pp. 393-396, 1972.
- [9] M. W. Berry and M. Castellanos, *Survey of Text Mining Clustering, Classification, and Retrieval*, 2nd ed. Springer, 2007.



**Kodai Shinkawa** is the undergraduate student of Faculty of Science and Engineering, Hosei University, Tokyo, Japan. He was born in Kanagawa, and his current interests are complex systems and multi-agent systems.



**Isamu Shioya** is a professor of Faculty of Science and Engineering, Hosei University in Tokyo, Japan. He was born in Niigata, and his interests are graph grammars, complex systems, moving multi-agents and traffic flow.