Experimental Verification of the Allelomimesis Clustering Model

Dranreb Earl Juanico* and Caesar Saloma
National Institute of Physics, University of the Philippines,
Diliman, Quezon City 1101
E-mail: dojuanico@up.edu.ph

ABSTRACT

The allelomimesis clustering model is based on only two parameters α and p, which represent the probability of nearest-neighbor copying and the fraction of unresponsive agents, respectively. The model results into the formation of clusters of agents, the sizes of which obey a distribution that is determined by the values of α and p. Several experimental data are fitted by tuning the two parameters. In particular, the significance of the value of α that corresponds to an experimental data is discussed and justified according to ecological context. Recommendations for possible extensions of the model are also enumerated.

INTRODUCTION

The penultimate hallmark of complex systems is the principle of emergence—macroscopic regularity of the system arising from apparently irregular (disordered) microscopic interactions between the system's constituents. Among the most commonly observed of these emergent properties is *clustering*.

Clusters in nature exist in different sizes. An interesting and somewhat unexpected observation is the regularity of the statistical distribution of cluster sizes, which generally obeys what is known as a *power law*. If we denote by *s* the size of a cluster and by *f*(*s*) the frequency of occurrence of *s* (i.e., the number of times that a particular value of *s* is tallied), then one would witness that the plot of *f*(*s*) versus *s* in double logarithmic scale is a decreasing straight line. The slope of this line corresponds to the exponent of the power law. Let us refer to the value of this slope as τ.

Power-law behavior is a well-known result in the area of complex systems. This article presents a simple, yet general, mechanism that leads to this power-law behavior.

In Sec. 2, a phenomenological model based on allelomimetic behavior is discussed. Allelomimesis is the tendency of individuals to imitate its neighbors; hence, allelomimetic behavior is the best candidate as local interaction that could lead to the formation of clusters. The results of the model are discussed in Sec. 3. A comparison of these results with several cluster systems found in nature is also presented. A few recommendations for further extension of this study are pointed out in Sec. 4.

THE CLUSTERING MODEL

Behavior of a single agent

The probability that an agent "performs" a certain "action" is described as follows:

$$P(\sigma) = \begin{cases} 1, & \text{if } \sigma > \sigma_c \\ 0, & \text{if } \sigma \leq \sigma_c \end{cases}, \qquad (1)$$

wherein σ is the total stimulus received by an agent (both from its environment and its neighboring agents) and s_c is an arbitrary threshold stimulus level. Equation

---

*Corresponding author

(1) means that an agent actuates if its stimulus level exceeds the threshold. Such characteristic is akin to neuronal stimulation (Churchland & Sejnowski, 1992).

In reality, however, an agent may override the consequence of $\sigma > \sigma_c$ by a complementary "choice" of not performing an action. Let us assume that the probability the agent will not actuate if $\sigma > \sigma_c$ is $1 - \beta$. Hence, Eq. (1) may be rewritten in the following form:

$$P(\sigma, \beta) = \begin{cases} \beta, & \text{if } \sigma > \sigma_c \\ 0, & \text{if } \sigma \leq \sigma_c \end{cases} . \qquad (2)$$

Equation (2) incorporates the ability of an agent to decide.

Generally, β and σ_c may vary among agents. But to make the model as simple as possible, it is assumed that β is a mean value over a population of agents; hence, β is a constant with respect to a particular agent population. However, β is allowed to vary between different populations. On the other hand, two cases are considered in assigning the value of σ_c—it is either fixed ($\sigma_c = 4$) for all agents or it varies within a range ($2 \leq \sigma_c \leq 16$) among agents.

Lattice of agents

Agents are distributed in a d-dimensional lattice of length L consisting of L^d discrete cells. A cell may only accommodate a single agent. Thus, L^d also corresponds to the size of the agent population.

A fraction p of the agent population is designated to be *unresponsive*. Unresponsive agents are characterized by $\beta = 0$ such that $P(\sigma, \beta) = 0$ even if $\sigma > \sigma_c$. These agents may be thought of as impurities in the lattice. The value of p is varied between 0 and 1.

Dynamics of action propagation

The lattice is constantly bombarded by external stimuli (e.g., environmental conditions or public information in the form of advertisements, and the like). The consequence of this is that at each time step of the numerical experiment, $\sigma \rightarrow \sigma + 1$ for all agents (i.e., stimulus levels are raised by a unit).

An agent is then randomly chosen, say an agent at cell i,j in a $d = 2$ lattice, to behave according to Eq. (1). Hence, if $\sigma(i,j)$ exceeds $\sigma_c(i,j)$, then this agent outputs a particular action **A**. Otherwise, nothing happens and another agent is randomly chosen. Allow us to distinguish this chosen agent as a *harbinger*. The harbinger is the initiator of an action. Once a harbinger is selected, the bombardment of external stimuli is momentarily paused to allow us to focus on the consequence of the harbinger's action to the entire lattice. Furthermore, we assume that there can only be one harbinger at a time, but any agent is a potential harbinger at any given time.

By the performance of **A** the harbinger's stimulus level decreases: $\sigma(i,j) \rightarrow \sigma(i,j) - 2d$, as though releasing tension. Meanwhile, the stimulus level of each of the harbinger's $2d$ nearest neighbors is increased by a unit due to their observance of **A**. These neighbors behave according to Eq. (2) to decide whether or not to mimic the harbinger and actuate **A**. For a responsive neighbor, sufficient stimulation ($> \sigma_c$) resulting from the observance of **A** makes it actuate **A** with a probability β. Let us assume that β is of a particular value α. In contrast, any amount of stimulation brought about by the observance of **A** will have no effect on an unresponsive neighbor.

The harbinger's neighbors in turn pass around the information to their corresponding neighbors by actuating **A**. This propagation continues up to the last agent that performs **A** without influencing its corresponding nearest neighbors.

The parameter α is defined as the *allelomimesis index*. Its value is tuned between 0 and 1. On one hand, $\alpha = 0$ is equivalent to setting $p = 1$, i.e., all agents in the population are unresponsive to their neighbors; hence, noncopying or nonallelomimetic. On the other hand, $\alpha = 1$ implies a highly allelomimetic population of agents, wherein imitation of neighbors is a big factor that promotes clustering.

The dynamics of action propagation can be summarized as two fundamental processes: (i) the selection of a harbinger that initiates an action and (ii) propagation of action through nearest-neighbor connections.

Cluster and cluster size

Process (ii) is repeated until the action initiated by the harbinger ceases to propagate. All agents that actuate or have actuated **A** are considered to belong to a cluster and the total number of these agents corresponds to the size of the cluster s. Subsequently, the bombardment of external stimuli is resumed for the proceeding time step. Process (i) results in the initiation of another action and process (ii) propagates this action through the lattice, hence, establishing the formation of another cluster. By repeating processes (i) and (ii) over several time steps, one generates different values of s. This allows one to deduce the statistical distribution $f(s)$.

RESULTS AND DISCUSSION

Numerical simulations

Figure 1 illustrates the morphology of clusters at different settings of the parameters α and p. These clusters apparently exhibit fractal structures resembling those that were produced by models of urban growth through *diffusion-limited aggregation* (Makse et al., 1995).

Let us first deal with the effect of varying α by setting $p = 0$. Figure 2 plots the power-law cluster size distribution (CSD) in double logarithmic scale for different values of α. Notice how the lines steepen with increasing value of α, indicating that the scaling exponent τ is negatively correlated with the parameter
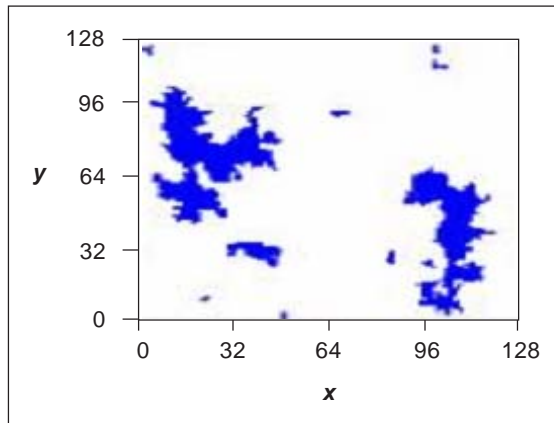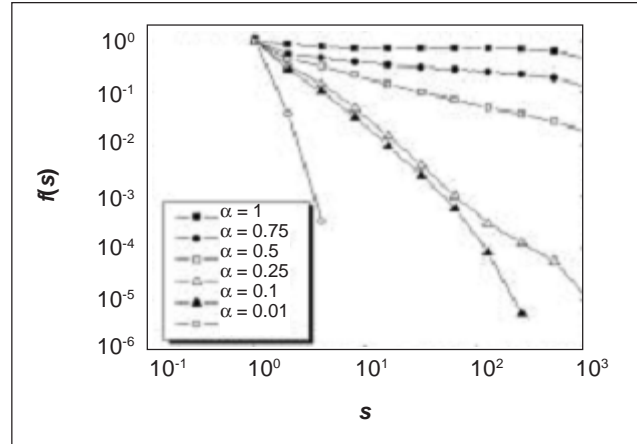

Fig. 2. CSD for different values of α at p=0.
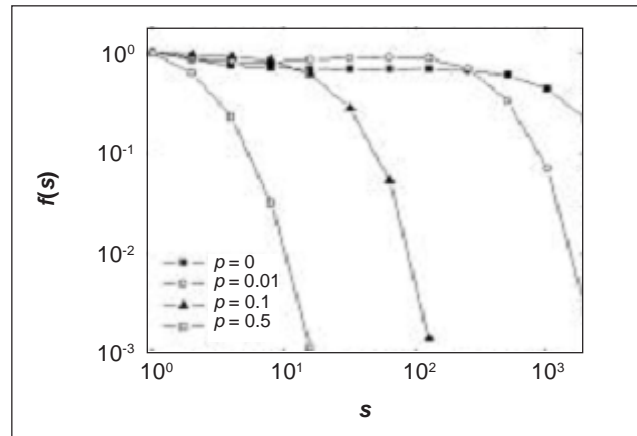

Fig. 3. CSD at $\alpha = 0$ and different values of p.

α. It is expected that $\tau \to \infty$ as $\alpha \to 0$, consistent with a dirac-delta CSD centered at s=1 for α=0 (i.e., no clusters are formed).

To show the effect of the parameter p on the CSD, we fix α to a value of 1. Figure 3 exhibits a distortion of the CSD at large values of s. The degree of such distortion intensifies with increasing p.

Considering that thresholds σ_c may vary among agents, we compare the set of CSDs (with different α and $p = 0$) for the case wherein $2 \le \sigma_c \le 16$ with the set for which $\sigma_c = 4$. Figure 4 plots the CSDs as data points in the former case and as broken lines in the latter case. There is no observable difference and this implies that the exact value of σ_c of an agent does not affect the CSD. Hence, the CSD is robust to variations of σ_c within an agent population.
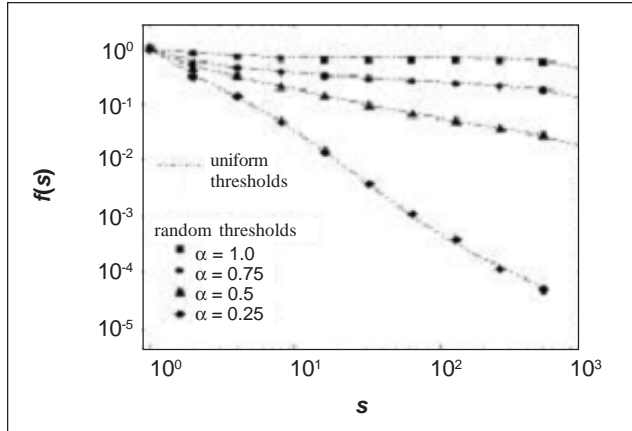

Fig. 1. Clusters that result for α=1 and p=0.

Fig. 4. Comparison between imposing uniform threshold ($\sigma_c = 4$) and random threshold ($2 \leq \sigma_c \leq 16$) for different values of α at $p=0$.

Comparison with data for real systems

We fit our model to different CSDs taken from experimental observations of actual clusters. The goodness-of-fit is measured in terms of the mean square error (MSE) between the data and the curve generated from the model. Figures 5 and 6 present data on four animal systems and four distinct human cluster systems, respectively.

Remarkably, α is high for animal systems (except for Serengeti lions), implying that allelomimesis is strongly expressed in animals. According to Wagner and Danchin (2003), "conspecific copying" (or allelomimesis) is a ubiquitous mechanism behind the formation of aggregates such as leks and colonies. Bonabeau and Dagorn (1995) showed that "biosocial attraction" (another form of allelomimesis) promotes schooling in fishes. Parrish and Edelstein-Keshet (1999) further proposed that allelomimesis is a generic mechanism that maintains the cluster as a cohesive unit. Indeed, ecological evidence for a high value of α in the animal kingdom is compelling. The low α (=0.1) for Serengeti lions is explained by the fact that the lions that were observed by Schaller (1972) were *nomadic*. This means that they are likely to wander alone or in small groups, hence, a small value of α.

The value of α for human cluster systems is low as compared with animal systems, implying that allelomimesis is only moderately expressed in human beings. This can be justified by considering that humans
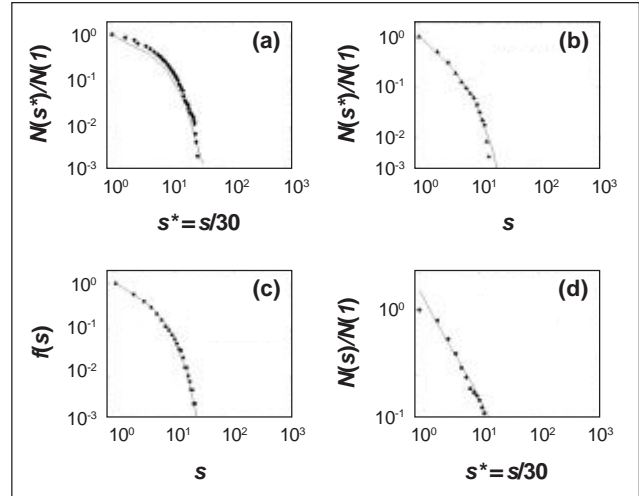

Fig. 5. Fitting the model to four different animal systems. (a) Spotted dolphin, *Stenella attenuata* ($\alpha = 0.75$, $p = 0.3$, MSE = 0.00381); (b) West Indian manatee, *Trichecus manatus* ($\alpha = 1$, $p = 0.45$, MSE = 0.00034); (c) Wasp, *Ropalidia fasciata* ($\alpha = 0.75$, $p = 0.35$, MSE = 0.00059); (d) Serengeti lion, *Panthera leo* ($\alpha = 0.1$, $p = 0$, MSE = 0.08661).
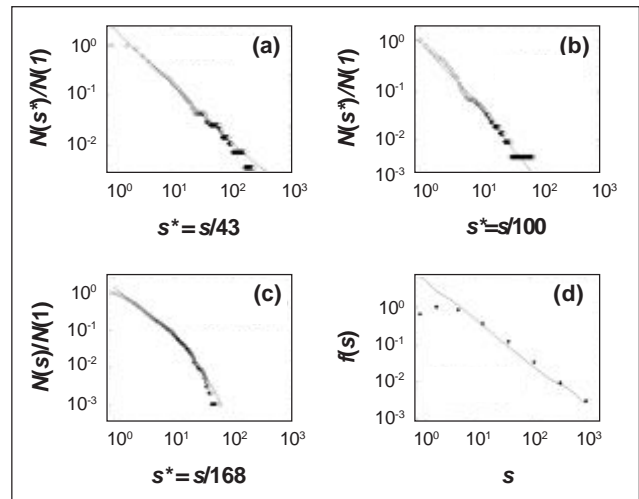

Fig. 6. Fitting the model to four distinct human cluster systems. (a) Urban agglomerations of India, 1991 ($\alpha = 0.32$, $p = 0$, MSE = 0.00394); (b) major cities of Japan, 1994 ($\alpha = 28$, $p = 0$, MSE = 0.00021); (c) households/barrios of Metro Cebu, 2001 ($\alpha = 0.31$, $p = 0.25$, MSE = 0.00179); (d) firms/clusters of employees of US, 1997 ($\alpha = 0.32$, $p = 0$, MSE = 6.16504).

are generally more highly cognitive than animals, which consequently overrides their instinctive tendency to be allelomimetic. Furthermore, telecommunication technology (which only humans are capable of) diminishes the requirement of information transfer through nearest-neighbor connections such as

allelomimesis. Interestingly, α is not significantly different among distinct human cluster systems ($\alpha\sim0.3$). This result is quite expected because even though we consider clusters of cities, of households, or of employees to be distinct from one another, one fact remains common between them—these systems are all made up of human beings. It would be worthwhile to investigate the origin of such seemingly universal value of α from a psychological point of view.

RECOMMENDATIONS

The model due to its inherent simplicity has cut down on details as much as possible so that it can be considered generic, hence, applicable to a wide variety of systems. Here, we suggest some minor points of modification to allow a more realistic description.

The stimulation on an agent due to constant bombardment of external factors may not necessarily be equal to unity (i.e., $\sigma\rightarrow\sigma+1$). It can be expressed as $\sigma\rightarrow\sigma+\eta$, where η represents a positive Poisson number that appropriately describes the time fluctuation of the amount of external stimuli. Furthermore, subsequent stimulation of neighboring agents may not necessarily decrease the stimulus level of the harbinger by an amount that is equal to the number of its nearest neighbors. That is, one can write $\sigma\rightarrow\sigma-\varepsilon$, where ε is a positive number derived from a Gaussian or a binomial probability distribution. It follows that the ensuing stimulation of responsive neighbors that observe the actuation of the harbinger can be expressed as $\sigma_n\rightarrow\sigma_n+\delta_n$, where the subscript n represents the nearest neighbor and $\sum_n\delta_n = \varepsilon$. Note here that δ_n could either be positive or negative, implying that the stimulation is excitatory or inhibitory, respectively (Churchland & Sejnowski, 1992).

CONCLUSION

A simple model of cluster formation is proposed to explain the cluster size distribution observed for various cluster systems in nature. The model consists of two mutually independent parameters, namely, α and p. The value of α represents the probability that an agent mimics the action of its nearest neighbors, whereas p

is the fraction of unresponsive agents that characterizes the particular agent population. Resulting CSDs are highly dependent on α and p.

The model fits into experimental data corresponding to various cluster systems in nature. A high value of α generally characterizes animal systems, whereas $\alpha\sim0.3$ distinguishes human cluster systems.

REFERENCES

Churchland, P.S. & T. Sejnowski, 1992. The computational brain. MIT Press, Cambridge.

Makse, H., S. Havlin, & H.E. Stanley, 1995. Modeling urban growth patterns. *Nature.* **377**: 608.

Wagner, R.H. & E. Danchin, 2003. Conspecific copying: a general mechanism of social aggregation. *Anim. Behav.* **65**: 405.

Bonabeau, E. & L. Dagorn, 1995. Possible universality in the size distribution of fish schools. *Phys. Rev. E.* **51**: 5220.

Parrish, J.K. & L. Edelstein-Keshet, 1999. From individuals to aggregations: Complexity, epiphenomena, and evolutionary trade-offs of animal aggregation. *Science.* **284**: 99-101.

Schaller, G.B., 1972. The Serengeti lion: A study of predator-prey relations. University of Chicago Press, Chicago.