Cluster formation by allelomimesis in real-world complex adaptive systems

Dranreb Earl Juanico, Christopher Monterola, and Caesar Saloma*
National Institute of Physics, University of the Philippines, Diliman, Quezon City, Philippines 1101
(Received 27 August 2004; published 11 April 2005)

Animal and human clusters are complex adaptive systems and many organize in cluster sizes s that obey the frequency distribution $D(s) \propto s^{-\tau}$. The exponent τ describes the relative abundance of the cluster sizes in a given system. Data analyses reveal that real-world clusters exhibit a broad spectrum of τ values, 0.7 (tuna fish schools) $\leqslant \tau \leqslant 4.61$ (T4 bacteriophage gene family sizes). Allelomimesis is proposed as an underlying mechanism for adaptation that explains the observed broad τ spectrum. Allelomimesis is the tendency of an individual to imitate the actions of others and two cluster systems have different τ values when their component agents display unequal degrees of allelomimetic tendencies. Cluster formation by allelomimesis is shown to be of three general types: namely, blind copying, information-use copying, and noncopying. Allelomimetic adaptation also reveals that the most stable cluster size is formed by three strongly allelomimetic individuals. Our finding is consistent with available field data taken from killer whales and marmots.

DOI: 10.1103/PhysRevE.71.041905 PACS number(s): 87.23.Cc, 89.75.−k, 82.30.Nr, 89.75.Fb

I. INTRODUCTION

Huge amounts of data have been collected and analyzed by researchers in various fields of the natural and social sciences concerning cluster formation in animals (fish schools, buffalo herds, etc.), humans (e.g., cities, slums, companies, etc.), and inanimate objects (gene families, colloids, galaxies, etc.). In the real world different cluster types often exist and share a common habitat and an accurate understanding of cluster formation in diverse animate and inanimate systems is valuable in wildlife preservation, environmental management, urban planning, economics, genetics, and even politics.

Animal and human clusters are complex systems with adaptive agents and many of them exist in cluster sizes s that obey the frequency distribution $D(s) \propto s^{-\tau}$. The exponent τ determines the relative abundance of the cluster sizes—a small τ (≈ 0) implies equal abundance of large and small clusters in a given system. Scale-free cluster distributions have been observed with gene families, colloids, fish schools, slum areas, city populations, business firms, and galaxies to name a few. Table I lists 46 different real-world cluster systems with their measured τ values ranging from $\tau = 0.7$ (tuna fish schools) [1,2] to $\tau = 2.95$ (galaxy clusters) [3].

To our knowledge, no single self-aggregation model could generate size-frequency distributions that cover the entire range of τ values in Table I. Available models are effective only at describing particular systems such as cities [4,5], firms [6,7], or gene families [8] because they utilize many interaction details that limit the range of their applicability [9].

Here, we propose that allelomimesis is an underlying mechanism for adaptation that can accurately explain the broad τ spectrum that is observed with real-world cluster systems. Allelomimesis is the act of copying one's kindred neighbors [10,11]. We show that the differences in τ values are caused by variations in allelomimetic behavior (described by a single parameter α where $0 \leqslant \alpha \leqslant 1$) of the agent phenotypes from one cluster system to another. We derive a nonlinear relation between τ and α that rationalizes the τ distribution which is observed in animal, human, and inanimate cluster systems. Cluster formation by allelomimesis predicts that the most stable cluster size is formed by three strongly allelomimetic ($\alpha \approx 1$) individuals ($s = 3$) which is consistent with previous observations in groupings of marmots [12] and killer whales [13].

Allelomimesis is the common mechanism underlying previous self-aggregation models. In animal aggregation models, it is manifested as biosocial attraction [2] or as conspecific copying [14]. For example, herding, which has been observed in panicking mice escaping from a two-door enclosure, is an interesting manifestation of nearest-neighbor copying [15]. In the percolation model of urban growth, allelomimesis is implicit in the concept of correlation [16]. The tendency of employees to associate with those belonging to the same income bracket in Axtell's model of firms [6,7] can also be construed as another expression of allelomimesis.

Copying is normal among social groups and it is an evolutionary mechanism in human societies [17]. Among strongly allelomimetic individuals, it is natural to expect that full cooperation is achieved quickly without the threat of punishment [18]. In gene families, cluster formation is explained as an intricate birth-death process [8]. In olivine crystal sizes, it is driven by complicated tectonic processes [19].

Allelomimetic interaction between agents can be described with few simple local rules [11]. A single measure α is sufficient to vary the exponent τ over a wide range of values. Agents search for neighbors (kindreds) and those that are strongly allelomimetic ($\alpha \approx 1$) are more likely to copy their neighbors, leading to the formation of relatively large clusters such as those observed in fish schools (see Table I). On the other hand, relatively large clusters are quite unlikely

*Corresponding author. FAX: +6329205474. Electronic address: csaloma@nip.upd.edu.ph

TABLE I. Measured τ values of real-world cluster systems with power-law $D(s)$ plots. Also shown are the corresponding values of the allelomimesis measure α.

Real-world cluster system	τ	α
Tuna near fish-aggregating device [1,2]	0.70	0.9895
Clupeid fish *Sardinella maderensis and S. aurea* [1,2]	0.95	0.9895
Alpine marmot *Marmota marmota* [12]	1.08±0.25	0.9895
African bufallo *Syncerus cafer* [2]	1.15	0.9894
Wasps *Ropalidia fasciata* [20]	1.19±0.008	0.9893
Tephritid flies [21]	1.37	0.9888
Free-swimming tuna with three species mixed [1,2]	1.49	0.9877
Spatial sizes of forest fires [22]	1.5	0.9877
Offshore-spotted dolphin *Stenella attenuata* [23]	1.79±0.05	0.9590
Three species of African baboons [24]	2.01±0.08	0.7841
West Indian manatee *Trichecus manatus* [25]	2.19±0.07	0.4009
Nomadic Serengeti cheetah *Acinonyx jubatus* [26]	2.49±0.35	0.0378
Nomadic Serengeti lion *Panthera leo* [26]	2.49±0.08	0.0379
Mathare valley squatter settlements, Kenya [27]	1.40±0.20	0.9888
Recife squatter settlements, Brazil [27]	1.60±0.20	0.9844
French manufacturing firms in 1962 [28]	1.84±0.08	0.9411
Japanese manufacturing firms in 1975 [28]	1.90±0.04	0.9067
Urban agglomerations, India [29]	1.93±0.002	0.87
Towns surrounding London in 1981 [4,16]	1.96	0.8514
Towns surrounding Berlin in 1981 [4,16]	1.98	0.8270
Swedish firms in 1993 [5]	1.98±0.08	0.8270
Urban areas in Great Britain in 1981 and 1991 [16]	2.03	0.7512
City populations in Brazil in 1991 and 1993 [29]	2.04±0.06	0.7329
Urban agglomerations in Russia in 1994 [29]	2.04±0.06	0.7285
Urban agglomerations in U.S.A. in 1994 [29]	2.04±0.07	0.7250
Urban agglomerations in France in 1982 and 1990 [29]	2.05±0.00	0.7049
U.S. firms in 1997 [7]	2.06±0.05	0.6974
City populations in Mexico in 1990 [29]	2.07±0.00	0.6814
City populations in China in 1990 [29]	2.11±0.00	0.5970
World's most populous cities in 2002 [29]	2.11±0.08	0.5959
British business and manufacturing firms in 1955 [30]	2.11	0.5881
City populations in Germany in 1994 [29]	2.15±0.20	0.5033
City populations in Japan in 1994 [29]	2.16±0.09	0.4692
Gene family sizes of *M. pneumoniae* [8]	2.69	0.0056
Gene family sizes of *S. cerevisiae* [8]	2.81	0.0018
Gene family sizes of *E. coli* [8]	2.84	0.0013
Gene family sizes of *Synechocystis sp.* [8]	3.17	0.0001
Gene family sizes of *H. influenzae* [8]	3.27	0.0000
Gene family sizes of *M. janaschii* [8]	3.62	0.0000
Gene family sizes of *Vaccinia virus* [8]	3.80	0.0000
Gene family sizes of *M. genitalum* [8]	4.02	0.0000
Gene family sizes of *T4 bacteriophage* [8]	4.61	0.0000
Olivine crystal sizes of GP45, granoblastic [19]	2.83±0.16	0.0015
Olivine crystal sizes of GP30, coarse [19]	3.03±0.18	0.0002
Olivine crystal sizes of LANZ3, porphyroclastic [19]	3.81±0.41	0.0000
Database of 8000 galaxy clusters [3]	2.95±0.36	0.0005

in gene families, colloids, and galaxies which are systems with (asocial) components that are incapable of copying each other ($\alpha=0$).

Of particular interest are human cluster systems such as slums of informal settlers, cities, and business firms. Table I reveals that such systems occupy a narrow band of the τ spectrum ($1.4 \leqslant \tau \leqslant 2.16$). More than other animals, humans are capable of overcoming the communal pull of blind copying (herding) and to make (unpopular) decisions that take into account the merit of available information. We determine the possible classes of allelomimetic interactions by establishing a quantitative relation between α and τ.

Our presentation proceeds as follows: In Sec. II we describe briefly our agent-based model for clustering by allelomimesis while in Sec. III we compare the predictions of the model with those observed in real-world clusters such as those listed in Table I. We end our presentation by discussing the results of our comparison.

II. METHODOLOGY

A. Agent-based model of allelomimetic interaction

A square $L \times L$ lattice with *free* boundaries [31] is utilized as a platform for our agent-based model of cluster formation. Initially at time step $q_0=0$, the lattice cells are empty and for every subsequent time step q, an agent is injected into a randomly selected vacant cell. An agent at cell location (x,y) is assigned a state $\Psi_m(x,y)$ that is taken from a set of possible states $\{\Psi_m\}=\{1,2,\ldots,M\}$. State $\Psi(x,y)$ represents a particular trait, preference, action, or any other social attribute that characterize an agent at any given time. For example, $\{\Psi_m\}$ could be different species of tuna swimming within the same field of observation [1] or the different types of behavior in groups of lions (e.g., hunting, sleeping, yawning, etc.) as described by Schaller [26]. An empty cell is assigned the value of $\Psi=0$.

An agent at (x,y) searches for other agents of similar state by occupying the next available cell of its Moore neighborhood which consists of the agent's eight nearest cells at locations $\{(x+j,y+k)\}$, where indices $j,k=-1,0,1$. The location $(x+0,y+0)=(x,y)$ represents the current (default) location of the agent. In deciding to occupy a neighboring cell $(x+j,y+k)$ in the next time step $q+1$, the agent evaluates the viability of its current position (x,y) with those of its neighboring cells using the discrete potential function Φ,

$$\Phi(x+j,y+k) = N_{Moore}\{1-\delta[\Psi(x+j,y+k)]\} - \delta[\Psi(x+j,y+k)] \sum_{u,v}^{N_{Moore}} \{\delta[\Psi(x,y) - \Psi(x+j+u,y+k+v)]\}, \quad (1)$$

where N_{Moore} is the number of sites in the Moore neighborhood of an agent and $\delta(w)$ is the Dirac delta function which is nonzero and equal to unity only when $w=0$. For a two-dimensional Moore neighborhood, $N_{Moore}=8$. The summation in Eq. (1) is taken from $u=-1$, $v=-1$ to $u=1$, $v=1$. The possible values for u and v are $-1,0,1$. The cell at $(x+j,y+k)$ is not ready for occupancy by the agent when $\Psi(x+j,y+k)\neq 0$ and the potential barrier $\Phi(x+j,y+k)$ is at its highest value of N_{Moore}. On the other hand, the cell at $(x+j,y+k)$ is vacant when $\Psi(x+j,y+k)=0$ and $-N_{Moore} \leqslant \Phi(x+j,y+k) \leqslant 0$.

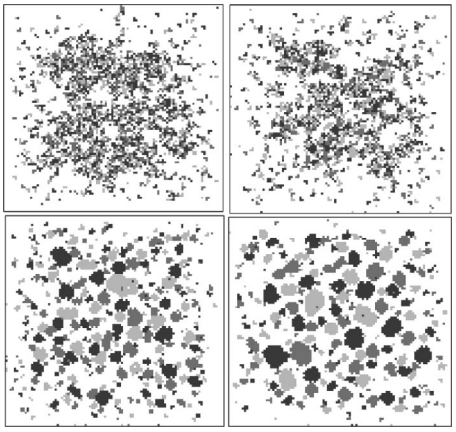

FIG. 1. Typical configurations of agents with $\alpha=0$ (first row, first column), 0.5 (first row, second column), 0.99 (second row, first column), and 0.995 (second row, second column) after $q_T=600\,000$ iterations ($L=100$). Different gray-levels represent three ($M=3$) possible phenotypes of an agent.

The set of cells $\{(x+j+u,y+k+v)\}$ represents the Moore neighborhood of the agent's neighboring cell $(x+j,y+k)$. The term $\delta[\Psi(x,y)-\Psi(x+j+u,y+k+v)]$ is a comparison between $\Psi(x,y)$ with those in the Moore neighborhood of the vacant cell at $(x+j,y+k)$. It is equal to unity when $\Psi(x,y)=\Psi(x+j+u,y+k+v)$. The summation in Eq. (1) yields the global minimum $\Phi=-N_{Moore}$ if all the agents occupying the Moore neighborhood of the vacant cell have the same state as the highlighted agent at (x,y). Otherwise, Φ will have a value greater than $-N_{Moore}$ but not exceeding 0.

Equation (1) is applied to every cell in the Moore neighborhood. An agent occupies the cell that yields the least value for Φ, expressed in the following minimum condition:

$$\Phi_{min} = \min_{j,k} \Phi(x+j,y+k), \quad (2)$$

wherein j,k run through all cells in the Moore neighborhood of (x,y) including itself. If more than one cell satisfies Eq. (2), then the agent randomly selects among these cells. Successive application of the above-mentioned mechanism for a sufficiently long period of time gives rise to the formation of clusters such as those shown in Fig. 1. A cluster is defined as a contiguous group of agents of the same state Ψ that are connected via neighboring cells.

The tendency to copy its neighbors can vary from one agent phenotype to another. We use a single parameter α to describe different levels of allelomimetic behavior where 0 (nonallelomimetic) $\leqslant \alpha \leqslant 1$ (blind copying). The state of a completely allelomimetic agent ($\alpha=1$) always depends on the condition of its Moore neighborhood while that of a nonallelomimetic agent is oblivious of the states of its Moore neighborhood. For computational simplicity, we assume that all the agents in a square lattice have the same α value.

Agent segregation and clustering are attributed directly to local behavior of the individual agents instead of a globally defined probability of segregation and clustering [1].

Let ζ be a uniformly random variable between 0 and 1. If $\zeta > \alpha$, then an agent randomly selects a state from the set $\{\Psi_m\}$. On the other hand, if $\zeta \leqslant \alpha$, then the agent's state is set by its Moore neighborhood such that the state with the highest occurrence within the neighborhood is the one copied by the agent. If more than one state satisfies this condition, then the agent selects randomly from among these states.

To avoid completely filling the lattice through the persistent addition of agents, we also include a constant probability of death for each agent. The death probability is held sufficiently low (1 in 10 000 for every time step) to maintain a relatively high lattice population density ρ as time progresses. When an agent dies, its cell location is vacated in the next time step. The agent population density ρ in the lattice consistently increases with q and agents are more likely to stay in their respective positions after a sufficiently long period of time when the number of available vacant cells has diminished.

Each simulation is run for a time-step duration of $q_T = 600\,000$ which is sufficient to allow the agent population to reach a steady-state density $\rho(q_T) \approx 0.5$. To increase the probability of conspecific agents meeting and coalescing within a reasonably short period of computational time, we choose $M=3$—i.e., $\{\Psi_m\}=\{1,2,3\}$. The choice of $M=3$ is consistent with the dimensional reduction hypothesis of Bonabeau *et al.* [2] which states that clustering is more likely to occur at low effective dimensions. We also mentioned that interesting clustering behavior has been observed in real-world systems wherein three different species of individuals are mixed within the same territory [1,2,24].

The essential aspects of scaling found in the $D(s)$ distributions that were generated with $M=3$ are also maintained at $M>3$. However, large cluster sizes ($s>100$) take significantly longer simulation times to form with $M>3$. Our model is based on the coalescence of kindreds and having a large number of different phenotypes (i.e., $M>3$) reduces the likelihood of an individual agent finding a kindred within its neighborhood at any particular time. This reduction coupled with agent death probability decreases the rate of cluster formation, thereby increasing the amount of time that is needed to form larger clusters.

The use of $M=1$ merely leads to an uninteresting result of filling up the entire lattice while $M=2$ leads to a binary mixture of two types of agents competing for space within the lattice. Either one phenotype fills the lattice while the other is extinguished or the two phenotypes coexist and are segregated by a boundary. Both $M=1$ and $M=2$ do not produce a wide range of cluster sizes which is required to show a statistical distribution.

Figure 1 presents typical configurations of a cluster system ($M=3$) at different α values. A relatively high local density of agents is found in the lattice interior due to the "free" boundary condition which eliminates all agents that wander beyond the lattice boundaries. We measure the cluster size s (in cell units) using the Hoshen-Kopelman algorithm [32]. A histogram (bin size=1) of the average cluster sizes is calculated for every α value using ten trials. The cluster-size distribution is fitted by a power-law function over a finite range of sizes, $s_{min} \leqslant s \leqslant s_{max}$, where s_{min} and s_{max} are the minimum and maximum cluster sizes, respectively. We determine the τ value from the best fit curve and use it to characterize the cluster-size distribution. Table II presents the relation between the allelomimesis measure α with exponent τ.

TABLE II. Some α values and resulting cluster-size distribution fitted by $F(s)=As^{-\tau}\exp(-s/s_c)$ with exponent $-\tau$ taken over the cluster-size range $s_{min} \leqslant s \leqslant s_{max}$.

α	τ	s_{min}	s_{max}
0.0	2.451±0.092	3	21
0.1	2.318±0.096	3	22
0.2	2.298±0.080	3	22
0.3	2.233±0.077	3	20
0.4	2.127±0.085	3	20
0.5	2.110±0.039	3	22
0.6	2.200±0.047	3	24
0.7	2.060±0.072	3	24
0.8	2.040±0.033	3	27
0.9	1.858±0.030	3	25
0.91	1.887±0.048	3	25
0.95	1.789±0.047	3	27
0.99	1.343±0.018	3	34
0.995	1.199±0.015	3	43
0.9996	1.045±0.016	3	42

We also determine the average length of time $\langle Q \rangle$ that an agent remains in a particular state and correlate it with the cluster size to which the agent belongs. $\langle Q(s) \rangle$ is the amount of time that an agent stays in a cluster of size s. It can serve as a rough measure of cluster stability.

B. Data from real-world cluster systems

The predictions of our model are compared with measurements taken from different real-world cluster systems (see Table I). Group dynamics is confined within a two-dimensional plane which is applicable to real-world clusters formed by human beings and terrestrial animals. Our model remains valid even to fishes which have been found not to utilize fully the three-dimensional character of oceanic space [2].

Cluster-size distributions are normally taken from direct-count observations and presented as absolute frequency $D(s)$ plots when the number of data points is sufficiently large or as cumulative-frequency plots when the data set is sparse. To within a predefined accuracy, both methods yield the same τ value of the $D(s)$ plot. Cluster-size data from different kinds of animals are presented commonly as direct-count values and data reliability depends heavily on the accuracy of spotting the correct number of members in an animal group.

Data on free-swimming tuna, buffalo, and sardinella are taken from Bonabeau's study [1,2] while those of Serengeti lions and cheetahs were sourced from Schaller [26]. Infor-

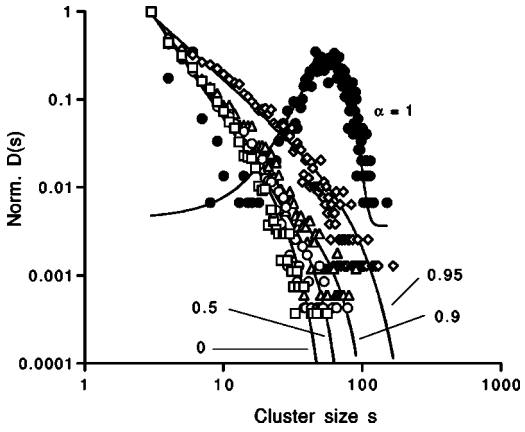

FIG. 2. Normalized cluster-size distributions. $D(s)$ plots for different α values generated from our model ($L=100, M=3$). Solid lines represent curve fits derived from a scaling function of the form: $F(s)=As^{-\tau}\exp(-s/s_c)$.

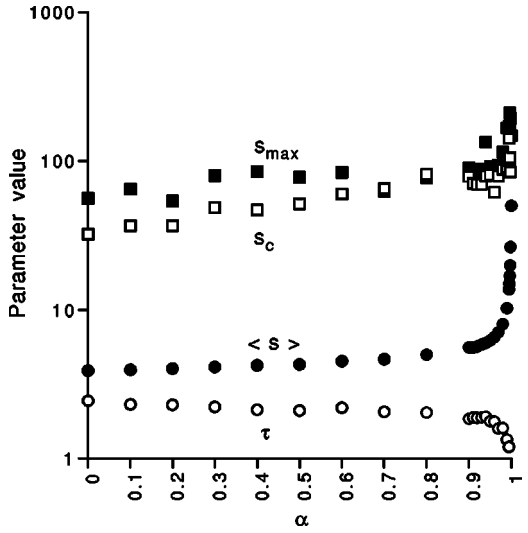

FIG. 3. Parameter dependence. The scaling exponent τ (open circles), mean cluster size $\langle s\rangle = M_1/M_0$ (solid circles), cutoff size s_c (open squares), and maximum cluster size s_{max} (solid squares) versus the allelomimesis measure α.

mation on baboons is accrued from several separate studies by Altmann and Altmann [24] on three different baboon species. Marmot data were quoted from Grimm [12]. Cluster formation data in dolphins and manatee were obtained from direct counting and their reliability was limited by visibility.

Information about city populations and slum areas was taken from Brinkhoff [29] and Sobreira and Gomes [27], respectively. The available data sets were plotted as a cumulative frequency distribution, $C(\geqslant s)=\Sigma_{s'}N(s')$, where $N(s')$ are the number of clusters of size s'. The summation is taken from $s'=s$ to the largest available cluster size s_{max}. The exponent τ is calculated using the property of power-law distributions which states that if τ' is the exponent of $C(\geqslant s)$, then it follows that $\tau=\tau'+1$ [33]. Data about the employee sizes of U.S. firms were taken from Axtell [7] while those from Swedish, Japanese, and British firms were obtained from Johansson [5] and Simon and Bonini [30], respectively.

We also studied cluster-size data of inanimate systems ($\alpha\approx 0$) such as gene families in various kinds of bacteria [8], olivine crystal sizes in xenoliths of the lithospheric mantle [19], and galaxy clusters [3].

III. EXPERIMENTAL RESULTS

Figure 2 shows plots of normalized cluster-size frequency distributions $D(s)=N(s)/N(s_{min})$, with $s_{min}\leqslant s\leqslant s_{max}$, for different α values ($0\leqslant\alpha\leqslant 0.996$). To determine its corresponding τ value, we fit a given $D(s)$ with the function $F(s)=As^{-\tau}\exp(s/s_c)$, where s_c is the cutoff size and A is a normalization constant. The s_{min}, s_{max}, and s_c values are related according to $s_{min}=3<s_c<s_{max}$. Sizes $s=1$ and $s=2$ are excluded from the curve fitting procedure to minimize deviations from the power-law distribution caused by boundary effects and finite agent population. Note that the $D(s)$ plot obtained with $\alpha=1$ is Gaussian like with a characteristic size at $s\approx 56$.

The fitting function $F(s)$ is similar to the one utilized by Bonabeau *et al.* [2]. The cutoff size s_c accounts for a number of characteristics that are observed in the formation of large clusters. One is cluster breakup or fragmentation which reduces the frequency of large clusters in a system. Another is the effect of boundaries, which becomes significant when agents are competing over a finite amount of space. The death rate and migration dynamics are factors that favor cluster breakups. The $M=3$ cluster phenotypes drive the competition for space in a finite-size lattice. The above-mentioned factors do not favor the formation of very large clusters and cause the tail of $D(s)$ to decay at a rate that is faster than a power law.

Figure 3 plots the dependence of τ with α (open circles) which shows that τ is independent of α for $\alpha<0.9$. However, as $\alpha\rightarrow 1$, τ rapidly decreases to zero, indicating a nonlinear relation between τ and α. Also shown is the dependence of the mean cluster size $\langle s\rangle=M_1/M_0$, with α (solid circles) where M_0 and M_1 are the zeroth- and first-order moments of the size distribution $N(s)$, respectively.

The sharp variation of τ and $\langle s\rangle$ as $\alpha\rightarrow 1$ implies a rapid increase in the degree of clustering among agents, hence a greater probability of the formation of large clusters. Figure 3 also plots the values of s_c (open squares) and s_{max} (solid squares) as a function of α. The plots reveal that both s_c and s_{max} increase rapidly and in a nonlinear fashion with α as $\alpha\rightarrow 1$. The plot behavior of τ, $\langle s\rangle$, s_c, and s_{max} versus α consistently indicates that strongly allelomimetic agents are capable of forming large, compact, and considerably stable associations. The stability of these clusters are further investigated by looking at the average amount of time $\langle Q(s)\rangle$ that an agent stays within a cluster of size s.

Figure 4 plots the nonlinear (α,τ) curve, which is described by the Fermi function:

$$\alpha(\tau)=\gamma\{1+\exp[\beta(\tau-\tau_c)]\}^{-1}, \tag{3}$$

with $\gamma\approx 1$, $\beta=0.104$, and $\tau_c=2.15$. We found that Eq. (3) strongly approximates the various (α,τ) pair of values that

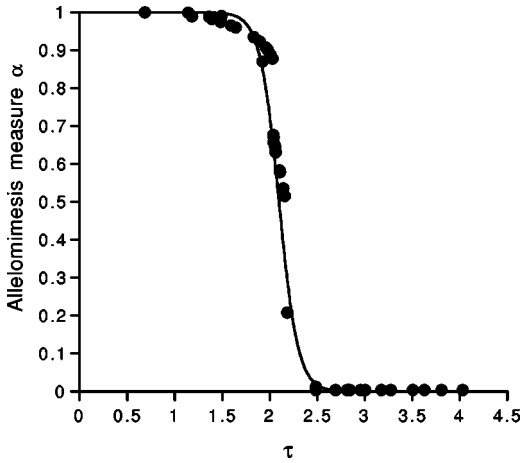

FIG. 4. Comparison of the Fermi fitting function (solid curve) with data (circles) from 32 real-world cluster systems. Measured τ values are plotted according to listing order of Table I.

were generated using our agent-based model. Also presented in Fig. 4 are the measured τ values (solid circles) from 32 real-world cluster systems in the order presented in Table I. The corresponding α value is interpolated from the known τ value of a real-world cluster system via the Fermi function $\alpha(\tau)$ (reduced $\chi^2=0.00529$, $R^2=0.9565$) in Eq. (3).

The behavior of $\alpha(\tau)$ indicates that allelomimetic interactions are of three general types: (1) *blind allelomimesis* ($\alpha \approx 1$) where agents are most likely to copy conspecifics, (2) *information-use allelomimesis* ($\alpha \propto \tau$) where agents are deliberate in their decisions to copy conspecifics, and (3) *non-allelomimetic* ($\alpha \approx 0$) where agents do not possess the social attribute to copy their neighbors.

Our findings are consistent with the experiment-based classifications proposed earlier by Wagner and Danchin [14]. Information-use copying may be considered an advance (evolutionary) trait found in humans. It plays a critical role in the expansion of business firms and cities [7,16]. Interestingly, our model indicates that the growth of slum areas is driven by blind copying and not by deliberate decisions based on available information.

Many animal clusters (e.g., fish schools, buffalos) benefit from blind copying ($0.98 < \alpha < 0.99$). Herding is more likely to emerge quickly in systems with strongly allelomimetic agents. On the other hand, baboons have a relatively low α at 0.78 because they live in heirarchical societies [24] where higher-ranked individuals are more likely to succeed in imposing their will on others below them—allelomimesis is biased towards dominance. The West Indian manatee have a relatively low α of 0.40 because they are solitary animals [25] and are unlikely to encounter other individuals of the same kind within their lifetime. Cheetahs and lions both have low α of 0.038 because they are nomadic and prefer to hunt alone or in small packs [26]. In these animals, the chances of being influenced by their neighbors are rather low.

Figure 5(a) plots the model's predictions of the $\langle Q(s)\rangle$ values for $\alpha=0.96$ (open circles) and that of killer whales *Orcinus orca* (solid circles). For whales we interpret Q as the number of hours of observation in direct count studies. The

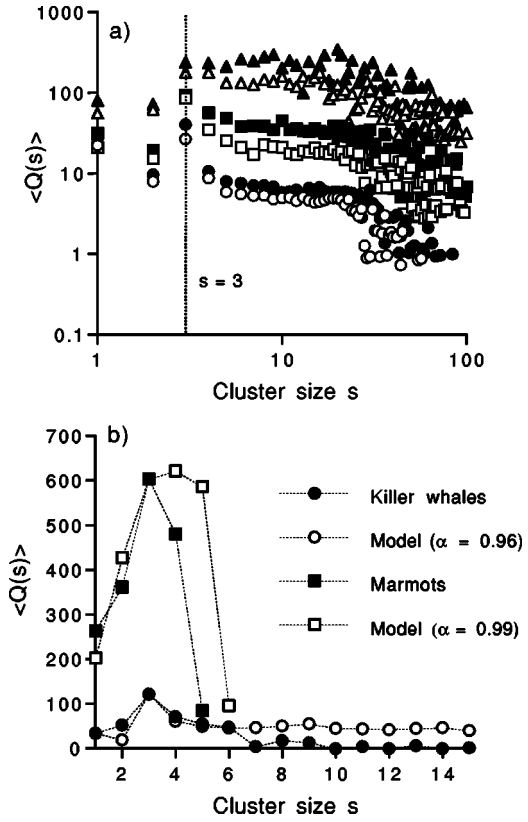

FIG. 5. Comparison of numerical predictions with real-world data: (a) $\langle Q(s)\rangle$ from an agent-based model at $\alpha=0$ (open circles), 0.5 (solid circles), 0.9 (open squares), 0.95 (solid squares), 0.99 (open triangles), and 0.995 (solid triangles). (b) $\langle Q(s)\rangle$ from killer whales (solid circles) and marmots (solid squares) and corresponding $\langle Q(s)\rangle$ from our model at $\alpha=0.96$ (open circles) and 0.99 (open squares), respectively.

cluster-size distribution data are unavailable for killer whales, so we assumed that they cluster by the same manner as dolphins for which group size data are available. Our assumption is justified because the *Orcinus orca* is a member of the dolphin family [34]. Dolphins exhibit a cluster-size distribution ($\tau=1.79\pm0.05$) that corresponds to $\alpha=0.9590$ (see Table I). The two plots show the same peak values and peak locations at $s=3$.

Yellow-bellied marmots (*Marmota flaviventris*) also exhibit optimum cluster stability at $s=3$ [35]. Also compared in Fig. 5(b) are the predicted $\langle Q(s)\rangle$ values for $\alpha=0.99$ (open squares) and that of marmots (solid squares). The plots show similar peak values and peak locations at $s=3$. We use the net reproduction rate to represent Q because it is directly related to the amount of time spent by individuals as a group. The predicted $\langle Q\rangle$ values for $\alpha=0.99$ were obtained using a large lattice ($L=500$) since marmots operate in relatively wide territories such as steppes, alpine meadows, pastures, or fields [36].

IV. DISCUSSION

We have shown that allelomimetic adaptation could explain the broad τ spectrum that is observed in a wide variety

of real-world cluster systems ranging from fish schools to shanties to gene family sizes and galaxies (see Table I). Cluster formation is governed by a few simple rules which makes the self-aggregation model effective in describing different kinds of cluster systems. A single parameter α $(0 \leqslant \alpha \leqslant 1)$ is needed to tune the exponent τ over a wide range of values.

A nonlinear relation exists between the degree of allelomimetic behavior (as measured by α) and τ which describes the relative abundance of the cluster size in the system. The $\alpha(\tau)$ curve is accurately approximated by a Fermi function [Eq. (3)]. Strong correlation is achieved between the predictions of the model and experimental evidence even with a simplified self-aggregation model that neglects possible differences in the allelomimetic behavior among agents. In the present investigation, all agents in a given system are assumed to have the same α and are confined to interact on a two-dimensional field.

The nonlinear character of the $\alpha(\tau)$ curve indicates that allelomimetic adaptation is of three general classes: namely, *blind copying*, *information-use copying*, and *nonallelomimetic*. Our findings are consistent with previous claims that were derived directly from experimental evidence [14]. Different cluster systems that benefit from blind copying are characterized by a small α range that is near unity. Such systems yield cluster-size frequency distributions with a wide range of possible τ values $(0 \leqslant \tau < 2)$, implying that the relative abundance of the cluster sizes in systems with strongly allelomimetic agents is sensitive to small variations in α. A similar sensitivity characteristic also occurs with nonallelomimetic cluster systems such as colloids and galaxies $(\alpha \approx 0)$.

Cluster systems that are formed by humans such as slums, cities, and business firms are associated with a wide range of α values $(0.1 \leqslant \alpha \leqslant 0.9)$. However, their $D(s)$ plots are restricted within a limited spectrum of τ values $(1.4 \leqslant \tau < 2.16)$. Human beings have developed (by evolution and learning from past mistakes) the capability to make decisions that consider group pressure, on the one hand, and merit of available information on the other.

Slum areas which are antisocial and often illegal entities are driven more by collective action of informal settlers which is possible among strongly allelomimetic individuals. Collective action has the advantage of strength that is gained by numbers. On the other hand, business firms and the cities in Germany and Japan are formed deliberately based on information-driven plans and project studies. The cluster-size distribution of slum areas tend to be uniform while those of business firms are likely to be biased towards the small and medium sized (in terms of employee number).

In cluster formation that arises from information-use copying, the cluster-size frequency distribution is weakly sensitive to slight α variations $(\alpha \propto \tau)$. The cluster systems are quite robust—a significant shift in the allelomimetic tendency of the component agents only results in a slight change of the cluster-size frequency distribution.

Allelomimetic adaptation also reveals that the most stable cluster size is formed by three strongly allelomimetic individuals $(s=3)$. The prediction is supported by field observations taken from killer whales and marmots. The cluster size $s=3$ is stable because it is the minimum (hence the most economical to maintain) cluster size where the concept of majority-based decision remains meaningful.

V. CONCLUSIONS

Allelomimesis is a generic interaction mechanism between adaptive agents that could accurately explain the richness of τ values observed in real-world cluster systems. Allelomimetic adaptation can be described by few and simple local rules. The absence of details makes the self-aggregation model effective to a wide variety of real-world cluster systems.

We have generated an $\alpha(\tau)$ curve that rationalizes the broad τ spectrum observed in real-world cluster systems. The curve can be utilized to formulate effective strategies in wildlife conservation, urban planning, and even product marketing. Allelomimesis explains why the cluster size of $s=3$ is preferred among strongly allelomimetic animals such as killer whales and marmots.

In the real world, it is not unusual to find several cluster systems occupying a common habitat. That each of them can be analyzed by one and the same self-aggregation model is proof of the underlying interconnectivity of animate and inanimate clusters. The existence of a *basic* mechanism for adaptation is vital in the formulation of effective strategies in wildlife preservation, environmental management, urban planning, economics, and even politics.

Allelomimesis in endangered species may be enhanced to favor large-cluster formation where reproductive success is greater since survivorship is directly related to group size [10]. Urban overcrowding may be reduced at minimum economic and social cost, with initiatives that discourage blind copying among humans. On the other hand, an efficient army or a successful beauty product may be developed by strategies that favor blind obedience and mass mimicry, respectively.

[1] E. Bonabeau and L. Dagorn, Phys. Rev. E **51**, R5220 (1995).
[2] E. Bonabeau, L. Dagorn, and P. Fréon, Proc. Natl. Acad. Sci. U.S.A. **96**, 4472 (1999).
[3] T. Jarrett, 2MASS Galaxy cluster catalog, Infrared Processing and Analysis Center, Jet Propulsion Laboratory, California Insititute of Technology, 1998, available online at, http://www.spider.ipac.caltech.edu/staff/jarrett/2mass/.
[4] H. E. Stanley *et al.*, Physica A **231**, 20 (1996).
[5] D. Johansson, Ph.D. thesis, Stockholm School of Economics, Stockholm, Sweden, 1997.
[6] R. Axtell (unpublished).
[7] R. Axtell, Science **293**, 1818 (2001).

[8] M. Huynen and E. van Nimwegen, Mol. Biol. Evol. **15**, 583 (1998).
[9] R. May, Science **303**, 790 (2004).
[10] J. Parrish and L. Edelstein-Keshet, Science **284**, 99 (1999).
[11] D. E. Juanico, C. Monterola, and C. Saloma, Physica A **320**, 590 (2003).
[12] V. Grimm, Oikos **102**, 124 (2003).
[13] R. Baird and L. Dill, Behav. Ecol. Sociobiol. **7**, 408 (1996).
[14] R. Wagner and E. Danchin, Anim. Behav. **65**, 405 (2003).
[15] C. Saloma *et al.*, Proc. Natl. Acad. Sci. U.S.A. **100**, 11947 (2003).
[16] H. Makse, S. Havlin, and H. E. Stanley, Nature (London) **377**, 608 (1995).
[17] P. Higgs, Proc. R. Soc. London, Ser. B **267**, 1355 (2000).
[18] M. Milinski, D. Semmann, and H. J. Krambeck, Nature (London) **415**, 424 (2002).
[19] P. Armienti and S. Tarquini, Lithos **65**, 273 (2002).
[20] Y. Ito, S. Yamane, and K. Tsuchida, J. Ethol. **20**, 3 (2002).
[21] M. Sjoberg, B. Albrectsen, and J. Hjalten, Ecol. Lett. **3**, 90 (2000).
[22] B. Malamud, G. Morein, and D. Turcotte, Science **281**, 1840 (1998).
[23] P. Perkins and E. Edwards (unpublished).
[24] S. Altmann and J. Altmann, *Baboon Ecology: African field research* (University of Chicago Press, Chicago, 1970).
[25] J. Kadel, A. Dukeman, and G. Patton (unpublished).
[26] G. B. Schaller, *The Serengeti Lion: A study of predator-prey relations* (University of Chicago Press, Chicago, 1972).
[27] F. Sobreira and M. Gomes (unpublished).
[28] B. van Ark and E. Monnikhof (unpublished).
[29] T. Brinkhoff, City population: principal cities and agglomerations of the world, 2004, available online at, http://www.citypopulation.de
[30] H. Simon and C. Bonini, Am. Econ. Rev. **48**, 607 (1958).
[31] C. Domb, *Phase Transitions and Critical Phenomena* (Academic Press, New York, 1973), Vol. 8.
[32] J. Hoshen and R. Kopelman, Phys. Rev. B **14**, 3438 (1976).
[33] S. Burroughs and S. Tebbens, Pure Appl. Geophys. **158**, 741 (2001).
[34] B. Culik, Orcinus orca: Killer whale, 2003, available online at, http://www.wcmc.org.uk.
[35] K. Armitage and O. Schwartz, Proc. Natl. Acad. Sci. U.S.A. **97**, 12149 (2000).
[36] S. Smith, The biogeography of the yellow-bellied marmot, San Francisco State University, Dept. of Geography, 2001, available online at, http://bss.sfsu.edu.224/courses/.