Evolution of a social network: The role of cultural diversity

A. Grabowski\textsuperscript{1,8} and R. A. Kosinski\textsuperscript{1,2,9}

\textsuperscript{1}Central Institute for Labour Protection—National Research Institute, 00-701 Warsaw, Poland
\textsuperscript{2}Faculty of Physics, Warsaw University of Technology, 00-662 Warsaw, Poland

(Received 7 June 2005; published 27 January 2006)

We present a simple deterministic and based on local rules model of evolving social network, which leads to a network with the properties of a real social system, e.g., small-world topology and assortative mixing. The state of an individual $s_i$ is characterized by the values of $Q$ cultural features, drawn from Gaussian distribution with variance $\sigma$. The other control parameter is sociability $T_c$, which describes the maximal number of connections of an individual. The state of individuals and connections between them evolve in time. As results from numerical computations, an initial diversity of cultural features in a community has an essential influence on an evolution of social network. It was found that for a critical value of control parameter $\sigma_c(Q)$ there is a structural transition and a hierarchical network with small-world topology of connections and a high clustering coefficient emerges. The emergence of small-world properties can be related to the creation of subcultural groups in a community. The power-law relation between the clustering coefficient of a node and its connectivity $C(k) \sim k^{-\gamma}$ was observed in the case of a scale-free distribution of sociability $T_c$ and a high enough cultural diversity in a population.

DOI: 10.1103/PhysRevE.73.016135

PACS number(s): 89.75.Fb, 89.65.−s

I. INTRODUCTION

In recent years it was found that a structure of different biological, technical, economical, and social systems has the form of complex networks [1–4].

The short length of the average shortest-path distance, the high value of the clustering coefficient and scale-free distribution of connectivity are some of the common properties of those networks [2,3,5]. Social networks, which are an important example of complex networks, also have such properties. They are successfully modeled using different approaches [4,6–11], in particular, small-world topology of interpersonal connections [2,3,12,13] and their hierarchical structure [13,14] are taken into account. Small-world networks are networks, which on one hand, have some properties of random graphs (i.e., the length of the average shortest-path distance is short and increases logarithmically with the size of the network); on the other hand, they are similar to regular networks, because the value of the clustering coefficient is high.

Different approaches to generation of graphs with desirable properties, e.g., a degree distribution or correlations between nodes connectivity, were presented [5]. In most of them it is assumed that a new connection can be created between each pair of nodes with a certain probability (e.g., proportional to node connectivity [2,5]), but some models are based on local (i.e., involving a node and its neighbors) rules [15,16].

The dynamic properties of systems which can be described in terms of complex networks attract great interest among physicists—the evolution of different networks (e.g., technical [17] or social [18]) is well described and success-fully modeled. In the case of social networks it is not only the evolution of the structure of the network that is investigated, but also the evolution of the state of the nodes (individuals). The state of an individual changes as a result of interactions with other individuals, e.g., in models of opinion formation, the state of an individual depends on the states of its neighbors [19,20]. In these models it is usually assumed that an individual can have one of two permitted states. This assumption make it possible to simplify analytical calculations and to describe a social system using the Ising model [21]. Individuals can also have additional parameters fixed during time evolution, e.g., authority [22]. However, if we want to simulate cultural evolution [23] the Ising-type approach is too simple. Axelrod’s model for the dissemination of culture is a good example of a more sophisticated approach [6,24,25]. In this model individuals are described by a vector of $Q$ cultural features and each component can take $n$ integer values (cultural traits). The more similar are the neighbors, i.e., the greater the number of the same cultural traits, the greater the probability that the neighbors will interact. As a result of this interaction, a cultural feature which has different values in both individuals is chosen and is modified in such a way that the trait of the second individual is assigned to the first one. In order to perform a statistical analysis of the model the order parameter is introduced, which is the size of the largest homogeneous domain, i.e., a group of individuals sharing the same cultural traits in all cultural features. The calculations of this model in a regular network identify a nonequilibrium phase transition separating an ordered (culturally polarized) phase from a disordered (culturually fragmented) one [26].

An interesting model of social networks is presented in Ref. [27], where individuals are randomly located in social space and the position of an individual is described by a vector of cultural features. In this model the connections between individuals are created with probability dependent on the distance between them in social space—the greater the
distance, the smaller the probability of a connection.

The fixed structure of a social network is the common feature of the abovementioned models of the evolution of the states of individuals. On the other hand, in some models the structure of a network changes in time, while the state of the nodes (e.g., intrinsic variables) is fixed [28]. However, real social networks evolve in both of these levels at the same time. Moreover, it appears that a co-evolution of the states of the nodes and interactions between them can lead to very interesting behavior of these systems [29–31]. To our knowledge, there is no deterministic model of co-evolution of the state of individuals and interactions between them, which would be based on simple local rules and lead to a network with the properties of a real social network.

In our model an individual is described like in Axelrod’s model with the vector of $Q$ cultural features. The second control parameter $\sigma$ describes initial distribution of cultural traits. The state of an individual changes in time as a result of interaction with its neighbors. Initially the network of acquaintances has the form of a regular network with large average shortest-path ($L$) and the clustering coefficient $C=0$. During the time evolution new connections between pairs of individuals are created and some existing connections die out. These processes of adding and removing the connections of an individual are determined by its state and the states of its neighbors. In the simplest case a new friend belongs to friends of our friends and such a mechanism, observed in real social networks, is applied in our model [32,33]. Our calculations show that in a specific range of control parameters this simple assumption is sufficient to reproduce the nontrivial feature of a social network—small-world topology of interpersonal connections. We calculate a phase diagram in a two-dimensional plane of control parameters $(Q, \sigma)$ describing cultural diversity in a population. To get deeper insight into the structure of a network we also investigate community structure [34].

II. EVOLUTION OF A NETWORK AND INITIAL CONDITIONS

In our model we investigate the evolution of a social network in a population consisting of $N$ individuals, located in a regular lattice. We assume that the number of individuals is fixed, because the timescale on which people change their interpersonal interactions is much shorter than the timescale on which the size of the population changes [32]. The state of an individual is described by a vector $S_i=(S_{i1}, S_{i2}, \ldots , S_{iq}, \ldots , S_{iQ})$ with $Q$ components, called cultural features. Each $S_{iq}$ can take an integer value (cultural trait) and it is assumed that there is an infinite number of cultural traits ($S_{iq}=-2,-1,0,1,2,\ldots$). Initially the values of $S_{iq}$ are independently drawn from probability distribution $P(S_{iq}=x)=\exp(-x^2/2\sigma^2)$ (e.g., Gaussian distribution with mean value zero), where the variance $\sigma$ is a measure of cultural diversity. Note that maximal cultural traits are not limited by a parameter of the model like in Axelrod’s model, but they are characterized by initial distribution. The number of social connections of the $i$th individual is denoted by $k_i$ and their localization evolves in time. To simplify our model we assume that social connections are symmetric and have the same value. Thus, a complex network evolving in time is formed.

All individuals are located in the lattice with size $N=L \times L$ and no periodic boundary conditions are used. In the time $t=0$ each individual is connected only with four nearest neighbors (individuals on the border of the lattice have fewer connections). We assume that those connections can not be removed, because they describe basic connections, e.g., with family (for most people their contacts with closest family are independent of differences in cultural features). Such a network has large average shortest-path distance (which is typical for regular networks) and the clustering coefficient $C=0$. The updating of the state of the network is performed using synchronous dynamics.

Let us describe the rules of the time evolution of the network. The value of $S_i$ changes in time as a result of interactions with other individuals. These changes depend on the states of other individuals connected with $S_i$. The next state of the $i$th individual in time $t+1$ is a result of its state in the current time step and interactions with other individuals:

$$S_{iq}(t+1) = S_{iq}(t) + h_{iq}$$  

(1)

for each component $q=1,2,\ldots ,Q$ the value of each cultural feature is modified independently and $h_i=-1,0,1$. The value of $h_{iq}$ depends on the states of $k_i$ neighbors of the $i$th individual. The components of the vectors $S_i$ of these neighbors can have higher, equal or lower values than the $i$th individual. The new state of the $i$th individual increases by $1$ ($h_{iq}=1$) when the number of the neighbors with higher values of the $q$th component is greater than the number of neighbors with the same value of this component and it is greater than the number of neighbors with lower values of this component (both conditions must be fulfilled). On the other hand, the value $S_{iq}$ decreases ($h_{iq}=-1$) if the number of neighbors with lower values of the $q$th component is greater than the number of neighbors with the same value of this component and if it is greater than the influence of neighbors with higher values of this component. Like in the previous case, both conditions must be fulfilled, otherwise $h_{iq}=0$.

In order to describe the evolution of the network we define the social distance $d_{ij}$ between a pair of individuals $(i,j)$:

$$d_{ij} = \sum_{q=1}^{Q} |S_{iq} - S_{jq}|$$  

(2)

In each time step the current number of connections of a node $k_i$ is compared to its sociability $T_i$—a parameter describing the maximal number of acquaintances, which can be maintained by the $i$th individual [32]. If $k_i > T_i$ one connection of the $i$th individual with the most distant neighbor is removed. On the other hand, if $k_i \leq T_i$, a new connection between the $i$th individual and an individual chosen from neighbors of the $i$th individual is created with respect to the shortest $d_{ij}$ distance [9,27] (in the case of more than one individual with the same social distance $d_{ij}$—the spatially closer one is chosen), which is shown in Fig. 1. In this way the network does not settle in a fixed point (its evolution never stops), which is a characteristic feature.
of living communities. After transient time the distribution of connectivity $k$ is similar to the distribution of sociability $T$. The behavior of the degree of a node as a function of time is not interesting and quickly reaches saturation.

In our work two different distributions of sociability were used: Uniform distribution ($T_i=24$ for all individuals—in this way the results are easily comparable with a regular network where an individual is connected to all individuals in the radius two) and scale-free distribution, where $P(T)\sim T^{-\gamma}$ ($\gamma=3$, as in Barabási-Albert network with linear preferential attachment [2], was used in most computations). In order to obtain better comparable results the average sociability was the same in both distributions (in the case of scale-free distribution maximal sociability $T_{\text{max}}=100$ was used).

III. RESULTS AND DISCUSSION

For the initial conditions described in Sec. II and both types of $T_i$ distributions the time evolution of the network is calculated and it is terminated when the time dependence of the parameters characterizing the structure of the network $\langle L \rangle$ and $C$ reaches a plateau. A typical time evolution of these parameters is depicted in Fig. 2. One can see that the length of the average shortest-path distance decreases significantly, while the value of the clustering coefficient increases. These changes are connected with the processes of removing and creating new connections between individuals. The structure of the network after simulation depends significantly on the values $\sigma$ and $Q$ used in computations (see Fig. 3). With an increase of the values of these parameters $\langle L \rangle$ decreases. This means that the more complex the initial state of the network (the cultural features are more diversified in the population), the smaller the $\langle L \rangle$. For small values of $\sigma$ and $Q$ all individuals have similar states $S_i$ and new connections are created with spatially close individuals. For greater values of $\sigma$ and $Q$ the number of individuals with a state different from the...
FIG. 4. The relation between the length of the average shortest-path \( \langle L \rangle \) and the size of the network \( N \) for \( Q = 3 \) and different values of \( \sigma \) (a)—for a large enough \( \sigma = \sigma_c \), the length of the average shortest-path increases linearly with logarithm of size of the network \( \langle L \rangle - \log N \). In order to obtain a phase diagram (b) we calculate critical values \( \sigma_c \) for each value of \( Q \). It is visible that for a large enough \( \sigma \) and \( Q \) the network has small-world properties. It is interesting that for simplest case \( Q = 1 \) the network has not small-world properties even for very large values of \( \sigma \).

For the smallest values of \( \sigma \), the clustering coefficient is relatively small. This is so because all individuals are in the same state and an individual creates connections with spatially closest individuals. For an increase in \( \sigma \), there is an increase in the diversity of the states of individuals. In consequence, there is also an increase in \( C \). This can be explained using Fig. 5(a), which illustrates the network with clusters formed by individuals in the same states and where a set of connections (black lines) of a chosen individual is shown. A large number of connections between individuals in one cluster results in a high value of \( C \). For a greater value of \( Q \) the clustering coefficient increases. This is so because individuals create—from the point of view of cultural features—groups that are more hermetic, i.e., groups of highly interconnected individuals with a number of cultural features with similar values.

For \( \sigma > \sigma_{\text{max}} \) the value of the clustering coefficient decreases with \( \sigma \) as a result of an increasing randomness in the spatial localization of connections. This is so because, for high values of \( \sigma \), there are many individuals, whose state is different from the mean value. The average value of the social distance in groups formed by those individuals is greater. In consequence, the formed groups are less hermetic and the spatial distribution of connections in the network is more random. However, groups of highly interconnected individuals remain in the network. They are formed by spatially distant individuals with similar states, as is shown in Figs. 5(b) and 5(c). Therefore, the clustering coefficient is smaller but still large.

For large enough values of \( \sigma \) and \( Q \), when the diversity of the values of cultural features is high enough, there is a large number of individuals who belong to many different groups, because each of them has cultural features that fit more than one group. In consequence, due to the existence of such individuals, these groups become less hermetic and the clustering coefficient decreases with increasing \( Q \).

It can be seen in Fig. 3 that, for the case of scale-free distribution of sociability, the influence of the values of the parameters \( Q \), \( \sigma \) on \( \langle L \rangle \) and \( C \) is similar to these relations obtained for \( T_l = \text{const} \). However, there are some discrepancies: The average shortest-path \( \langle L \rangle \) and the clustering coefficient \( C \) are smaller. Individuals with high connectivity play the role of hubs, which connect individuals from different and spatially distant parts of the network and shorten the average shortest-path. Moreover, hubs are more likely to
The behavior of the clustering coefficient $C$ for our model and for real networks is an interesting problem. In the case of $Q=1$ the clustering coefficient of a node decreases exponentially with its number of connections; however for $Q>1$ the power-law relation $C(k) \sim k^{-\beta}$ is visible (see Fig. 6) [1,17,35]. The value of $\beta$ slightly depends on the values of $Q$ and $\sigma$ and equals approximately $\beta=0.7$ for a wide range of values of these parameters. Such a relation is observed in some real social networks [1,18]. It is interesting that the value of parameter $\beta$ is close to the value obtained for the Internet at an autonomous system level [17], where $C(k) \sim k^{-0.75}$. The power-law relation $C(k)$, obtained in our calculation, is similar to the relation observed in hierarchical networks [1]. Such power laws hint at the presence of a hierarchical architecture: when small groups organize into increasingly larger groups in a hierarchical manner, local clustering decreases on different scales according to such a power law. This may be connected with the fact that individuals create connections on the basis of social distances—spatial distance between individuals is much less important. The influence of spatial distance between nodes in Euclidean growing scale-free networks on $C(k)$ relation is described in Ref. [36].

Let us discuss the time evolution of cultural features in a community. The initial, Gaussian distributions of cultural traits evolve in time and some cultural traits became more favorable, which is shown in Fig. 7 for networks presented in Fig. 5. During the time evolution of a network, peaks corresponding to some cultural traits appear in the distribution of cultural features. The main cultural trends in the community are represented by peaks near the cultural trait $x=0$. Additional peaks, distant from the main cultural trend are connected with certain cultural traits (e.g., $x=9$ or $x=-8$ in Figs. 7(a) and 7(b), respectively), which can be interpreted as a creation of certain subcultures in the community. They are accompanied by a shortening of the length of the average shortest-path distance in social network. It can be seen that individuals forming a subculture group are randomly distributed in the community, which can be seen in Figs. 5(b) and in 5(c) where the connections of an individual with the cultural trait $x=9$ and $x=-8$ are shown, respectively. The number of subcultures depends on the initial cultural diversity in a population; the greater the $\sigma$, the greater the number of subcultures. Note that the obtained distribution of cultural features deviates from Gaussian as the clustering of individuals with similar cultural traits creates certain subcultures.
to the networks from Figs. 5 results obtained after applying the above-mentioned algorithm. Figures 8 hierarchical tree describing the partition of a network into

This algorithm makes it possible to evaluate the community structure in a network by generating a dendrogram—a hierarchical tree describing the partition of a network into communities. Dendrograms are cut when the maximal value of modularity $M$ is reached; $M=0.37$ and $0.35$ for (a) and (b), respectively. A community is distinguished in (a), because it is the subculture of individuals having for the value of the cultural trait of the first cultural feature $x=9$ [see Fig. 7(a)].

$$T=\text{const.} \text{ Thus, in the case of real heterogeneous social networks it can be difficult to find subcultures (subgroups) when analyzing only the structure of the network. It is interesting that in the network some small communities emerge (the number of individuals is similar to the average connectivity), which are very hermetic and individuals inside of such a community have the same values of all cultural features. The number of such communities is much greater in the case of uniform distribution of sociability [see Fig. 8(a)] than in the case of scale-free distribution of sociability [in Fig. 8(b) only one such community is visible].}

IV. CONCLUSIONS

In our model we investigate a network of interpersonal contacts connecting individuals who are characterized by the vector of cultural features and sociability $T$. Both the states of the individuals and the interpersonal interactions evolve in time. As results from our computations the initial diversity of cultural features in the community, measured with the parameter $\sigma$, has an essential influence on the evolution of a social network. An increase in $\sigma$ results in a decrease of the length of the average shortest-path distance. The number of cultural features $Q$ is also important. The linear relation between the average shortest-path and the logarithm of the size of the network $\langle L \rangle \sim \log N$ is observed only for large enough $\sigma$ and $Q$. This means that small-world topology of interpersonal connections appears in the social network for high enough cultural diversity. It should be stressed that initially in our model there were no long-range connections—they are created due to the grouping of the spatially distant individuals with similar cultural features. This phenomenon is connected with the emergence of subculture groups—a process observed in living societies. As a result of this grouping process, the network with a high value of the clustering coefficient emerges. These properties are typical for social networks [3,12].

It was found that in a network with scale-free distribution of connectivity and the same number of connections, the
average shortest-path and the clustering coefficient are smaller than in the case of uniform distribution of connectivity. This is so, because individuals with high connectivity connect with many other individuals from distant parts of the network. On the other hand, these individuals have a significantly lower value of the clustering coefficient. The power-law relation between the clustering coefficient of a node and its number of connections $C(k) \sim k^{-\rho}$, obtained in our calculation, reveals a hierarchical structure of a social network. Assortative mixing by degree is another property of the considered networks.

It should be noted that a similar evolution of a network, especially the creation of new connections, is observed in many other systems, e.g., the evolution of the network of WWW pages, where most links from a page are connected to pages concerning the same subject.