Dynamic phenomena and human activity in an artificial society

A. Grabowski,[1,*] N. Kruszewska,[2,†] and R. A. Kosiński[1,3,‡]

[1]*Central Institute for Labour Protection, National Research Institute, 00-701 Warsaw, Poland*

[2]*Institute of Mathematics and Physics, University of Technology and Life Sciences, 85-796 Bydgoszcz, Poland*

[3]*Faculty of Physics, Warsaw University of Technology, 00-662 Warsaw, Poland*

(Received 30 November 2007; revised manuscript received 9 May 2008; published 18 December 2008)

We study dynamic phenomena in a large social network of nearly 3×10^4 individuals who interact in the large virtual world of a massive multiplayer online role playing game. On the basis of a database received from the online game server, we examine the structure of the friendship network and human dynamics. To investigate the relation between networks of acquaintances in virtual and real worlds, we carried out a survey among the players. We show that, even though the virtual network did not develop as a growing graph of an underlying network of social acquaintances in the real world, it influences it. Furthermore we find very interesting scaling laws concerning human dynamics. Our research shows how long people are interested in a single task and how much time they devote to it. Surprisingly, exponent values in both cases are close to -1. We calculate the activity of individuals, i.e., the relative time daily devoted to interactions with others in the artificial society. Our research shows that the distribution of activity is not uniform and is highly correlated with the degree of the node, and that such human activity has a significant influence on dynamic phenomena, e.g., epidemic spreading and rumor propagation, in complex networks. We find that spreading is accelerated (an epidemic) or decelerated (a rumor) as a result of superspreaders' various behavior.

DOI: 10.1103/PhysRevE.78.066110 PACS number(s): 89.75.−k, 89.65.−s

I. INTRODUCTION

In recent years investigations of complex networks have attracted the physics community's great interest. It has been discovered that the structures of various biological, technical, economical, and social systems have the form of complex networks [1–3]. The advent of modern database technology has greatly advanced the statistical study of networks. Because the available data sets are vast, it is possible to use techniques of statistical physics [2].

Studying the statistical properties of social, e.g., friendship, networks remains a challenge. It is possible to assess the form of degree distribution with a survey, as in the case of the web of human sexual contacts [4]. However, it is much more difficult to learn about other important properties of networks, because there are no data on their entirety. A survey often provides data on a small sample only.

Progress in information technology has made it possible to investigate the structure of the social networks of interpersonal interactions maintained over the internet, e.g. e-mail networks [5] and web-based social networks of artificial communities [6]. However, there is still an unexplored area of research. In recent years online games have become increasingly popular and have attracted an increasing number of players, who interact in the large virtual world of massive multiplayer online role playing games (MMORPGs).

A MMORPG is a network game in which players enter a virtual world as characters they have invented—gaining virtual life. This virtual world takes the form of a game server connected to the internet, on which accounts are registered for users who log in through special game client programs. The rules allow players to create more than one character on one account, with each of those characters having its own personality (further in the text we refer to such characters as individuals). Thousands of people can play on one server—they become a virtual society—so they share the common culture, area, identity, and interactions in the network of interpersonal relationships.

Individuals exploring this virtual world can collect funds, trade, organize in groups of different sizes which can make alliances or wage war, etc. All individuals can add, by mutual consent, other people to their databases of friends. In this way an undirected friendship network is formed. Playing time, the network of interpersonal relations (the friendship network), and the network's expansion, are the observables that illustrate the activity of the virtual society and provide an opportunity to study human dynamics. Moreover, a survey carried out among the players can show the influence of this activity on the real world.

In recent years dynamic phenomena in social networks like epidemic spreading and rumor propagation have been investigated with different models of interpersonal interactions [7,8]. Different approaches to the generation of graphs with desirable properties have been used, e.g., degree distribution or correlations between the degrees of nodes [2]. The influence of the heavy-tailed distribution of the intercontact times between susceptible and infected individuals on the spreading of computer viruses has been presented in [9]. We believe that human social activity has a strong influence on various dynamic phenomena in social networks.

The first aim of this work is to introduce a data set describing a large social network of an online game [10], which consists of almost 3×10^4 individuals. The project was started in Poland. We present data collected during its two years of existence. We show that the structure of this network has similar properties to those of other social networks. We

*angra@ciop.pl
†nkruszewska@utp.edu.pl
‡rokos@ciop.pl

TABLE I. Average properties of the whole network and the giant component (GC) and a comparison with a random graph (RG) with the same number of nodes N and the same average degree $\langle k \rangle$.

	N	C	$\langle l \rangle$	$\langle k \rangle$	$k_{\max}$
Network	28011	0.02		1.4	64
RG	28011	0.006	29	1.4	
GC	6065	0.1	4.8	6.4	64
RG	6065	0.006	4.8	6.4	

have found interesting scaling laws concerning human dynamics [11–13], e.g., duration of player activity in the virtual world. To study the structure of the network and of human activity, we analyzed data containing the list of all friends, creation date, last log-in date, and accumulative (total) playing time of each individual. The network under consideration consisted of a collection of individuals (network nodes) connected with one another by friendly relationships (network links).

The second aim of our work is to investigate the influence of human social activity on dynamic phenomena in a social network. In the present work we use data on a social network consisting of 6×10^3 individuals. It is a giant component of a network of individuals who interact in the large virtual world of the aforementioned MMORPG [14]. On the basis of playing time, we calculate the activity A of individuals, i.e., the relative time devoted daily to interactions with others. It has often been assumed in models of epidemic spreading that the intensity of interactions with others is the same for each individual. Our research has shown that the distribution of activity is not uniform and is highly correlated with the degree of the node. On the basis of human activity calculated as above, we investigate epidemic spreading [on the basis of a susceptible-infected-recovered (SIR) model] and rumor propagation in a real social network.

The paper is organized as follows. In Secs. II and IV we describe the structure of a friendship network and calculate parameters of human activity. In Sec. III we present results of a survey. Next, in Sec. V (using results from Sec. IV) we investigate epidemic spreading and, in Sec. VI, rumor propagation in a real social network. Section VII has the conclusions.

II. THE STRUCTURE OF THE NETWORK

Basic network measures of the whole network and the giant component (GC) [1] are presented in Table I. This network consists of 28 011 individuals but for many of them the number of connections k equals zero. This means that those characters have no friends on their lists. Most individuals with $k=0$ are abandoned characters who do not appear in the virtual world and who have therefore lost all their contacts with those still active, or are new characters (cf. Fig. 5, in which one can see numerous individuals who have spent 1–10 days in the game). The GC contains almost all individuals whose degree is greater than zero (6065 characters); only 252 individuals with $k>0$ do not belong to the GC (i.e., only 4%). It is noteworthy that the maximal value of k is 64 individuals, which is the maximal number imposed by game mechanisms. However, this restriction does not influence the form of the degree distribution, because there is only one individual with $k=64$ in the network (and nine individuals with $k>60$).

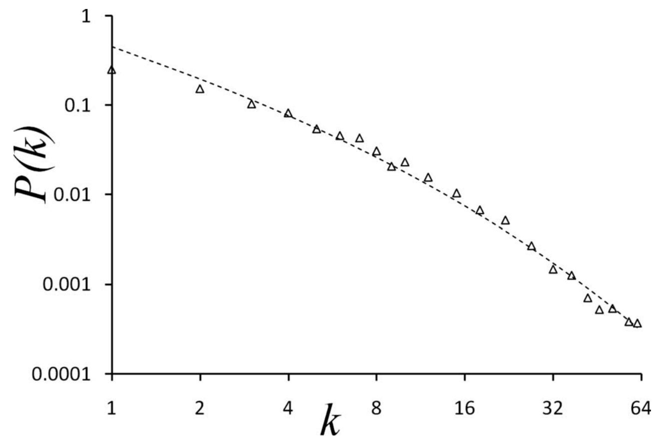

FIG. 1. Degree distribution $P(k)$. The results can be approximated using the fitting function $P(k)\sim k^{-0.65}\exp(-1.4k^{-0.35})$ (dashed line).

The average path length $\langle l \rangle$ in the GC is similar to that in a random graph. A high value of the clustering coefficient and a short average path length $\langle l \rangle$ are characteristic features of social networks [2,15,16]; they are typical for small-world networks [17]. The degree distribution of the network is plotted in Fig. 1. Deviations from power-law behavior at small values of degree are observed in some networks [2,18,19]. In such cases the fitting can be improved by using the functions $P(k)\sim(k+k_0)^{-\gamma}$, e.g., for out-degree distribution of the world wide web (WWW), $k_0=6.94$ and $\gamma=2.82$ [2,18]. However in our case the range of values of the degree is relatively small (less than two orders of magnitude, $k<64$). Therefore, we use two different fitting functions: $P(k)\sim(k+k_0)^{-\gamma_1}$ ($k_0=6.2\pm 0.9$; $\gamma_1=2.9\pm 0.1$—note that the values are very similar to those observed for the WWW network) and $P(k)\sim k^{-\gamma_2}\exp(-\eta k)$, ($\eta=0.048\pm 0.005$; $\gamma_2=1.1\pm 0.1$).[1] Both functions provide good fits to the observed data ($R\approx 0.98$).

The aforementioned results suggest that for such a short range of degree observed in the network under investigation it is possible to find different (two-parameter) functions that fit. Data from the degree distribution cannot explain which function fits better. However, in our case we have additional data on individuals' activity and the relation between an individual's activity and its degree. These data indicate that the degree distribution follows an exponential decay (for more details see Sec. IV).

In the network under investigation, the greater the k, the greater the average degree of nearest neighbors k_{NN}. Hence, the network is assortatively mixed by degree; such a correlation has been observed in many social networks [20]. In social networks it is entirely possible, and is often assumed in sociological literature, that similar people attract one an-

[1]Computations were performed using STATISTICA v.7.

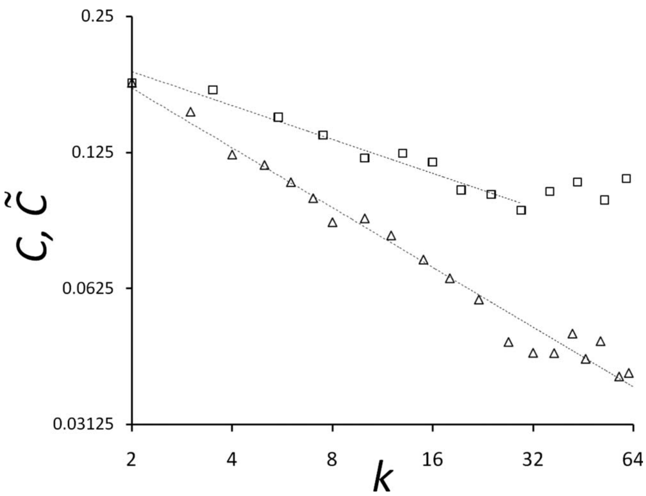

FIG. 2. Relation between the clustering coefficient of a node, C (triangles), $\widetilde{C}$ (boxes), and its degree k and the fit with a power law (dashed line).

other. The relation $k_{\mathrm{NN}}(k)$ can be approximated with the power-law relation $k_{\mathrm{NN}}(k) \sim k^{0.18\pm0.01}$ ($R^2=0.93$). A similar value of the exponent (0.2) has been found in other social networks [21].

The behavior of the clustering coefficient C is an interesting problem [1,2]. We measure the clustering coefficient of the ith node as the number of connections between neighbors E_i divided by the maximal number of possible connections:

$$C = \frac{E_i}{k_i(k_i-1)/2}. \quad (1)$$

The local clustering coefficient $C(k)$ is negatively correlated with the degree of the node k, showing the existence of a power law $C(k) \sim k^{-\alpha}$ with $\alpha=0.44\pm0.02$ ($R^2=0.98$). A slightly lower value of the exponent α has been observed in other social networks, 0.33 [6], and 0.35 in a network consisting of over 1×10^6 nodes [21]. The power-law relation $C(k)$ is similar to the relationship in hierarchical networks [22,23]. However, it has been recently shown [24] that most observed degree dependence of clustering coefficients follows from degree mixing patterns.

To check whether the observed decay $C(k)$ is just a consequence of assortative mixing we calculate the clustering coefficient $\widetilde{C}$ using the algorithm presented in Ref. [24]. According to the definition, $\widetilde{C}$ allows us to quantify the degree among the neighbors of a node, independently of its degree and the degree of its neighbors. The relation between C, $\widetilde{C}$, and the node degree k is shown in Fig. 2. For low degree we do not observe a substantial difference between the two definitions of a clustering coefficient, which provide similar results. However, the greater k, the greater the difference in values of $C(k)$ and $\widetilde{C}(k)$. For $k>60$ the clustering coefficient $\widetilde{C}$ is approximately three times larger than C. This result shows that the variations in the clustering coefficient C with vertex degrees just reflect partially the existence of degree correlations.

III. RESULTS OF A POLL

With the server continually accessible for two years and its software updated, there was an excellent opportunity to create a continually evolving virtual space where a specific local society could come into being. Social interactions with other players are an important part of each MMORPG. On the one hand, such interactions influence the network of acquaintances in the real world. On the other, some preexisting acquaintances from the real world are maintained in the virtual one, too. To investigate the relation between networks of acquaintances in the virtual and real worlds, we carried out a survey among active players. As mentioned in the Introduction, a survey provides data on a small sample of a network.

Only 6% of active players were interested in completing our survey (360 people). Moreover, they were players who had spent many months in the game (on average 16.4 months) and had numerous friends inside the game so answers from people for whom the game is an important part of their life dominate the results. We asked five questions. (a) How many people are on your list of friends—k? (b) How long have you played this game—T_L? (c) How many people from your list of friends did you know before you started to play—N_C? (d) How many people whom you got to know in the virtual world have you met (at least once) in the real world—N_D? (e) With how many people that you got to know in the virtual world and have met in the real world do you maintain social contacts in the real world (you contact each other at least once a week)—N_E?

The declared average number of friends was about 18.4 ± 1.6 characters. Many players said that they found out about the game from other people who played this game (the *snowball* effect).

The average number of preexisting acquaintances N_C turned out to account for 13% ($N_C=2.5\pm0.3$ people) of the friendship network in the virtual world. Hence, the network under investigation did not develop as a growing graph of an underlying network of social acquaintances in the real world. The results of the survey as a function of k are shown in Fig. 3.

The average number of contacts established as a result of playing the game, $N_D=6.5\pm1.0$ people, is significantly greater than N_C. The players often meet in the real world (e.g., people from the same city). It is a good opportunity to meet new people. However, the average number of contacts maintained in the real world, $N_E=3.9\pm0.7$ people, is smaller, but still larger than N_C. Hence, players maintain contacts over a long time with only some of the people they met in the real world. The results of the survey indicate that online games may have a strong influence on the network of acquaintances in the real world.

It should be stressed that all players we surveyed were from the same country. In the case of MMORPGs with players from different countries, the influence of interactions in the virtual world on contacts in the real world will probably be smaller (because communication, travel, etc., are difficult).

IV. INVESTIGATION OF HUMAN DYNAMICS

Online games like MMORPGs offer a great opportunity to investigate human dynamics, because much information

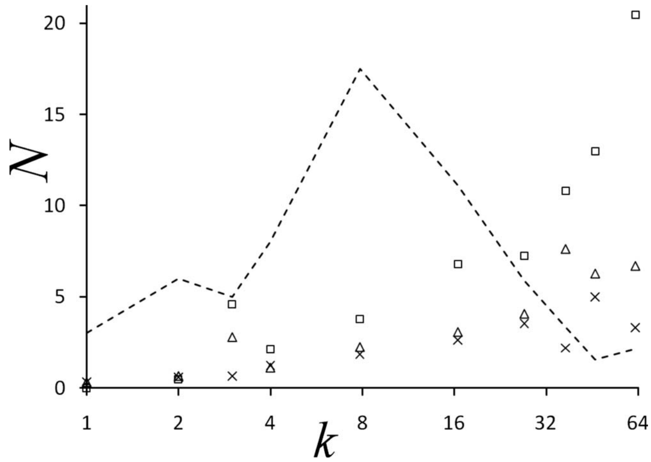

FIG. 3. Results of the survey. The numbers N_C (crosses), N_D (boxes), N_E (triangles), and the number of people who filled out the survey (dashed line) as a function of the number of people in the list of friends, k.

about individuals is registered in databases. To analyze how long people are interested in a single task and how much time they devote to it, we studied cumulative time spent in the virtual world, T_G, which is registered in the game database. The time T_G is the sum of the duration of all connections to the game server. It turned out that the distribution of the probability P_G that an individual devotes T_G hours to the game has the power-law form $P_G(T_G) \sim T_G^{-1.10\pm0.03}$ (Fig. 4). Thus, the probability that a human will devote time t to a single activity has a fat-tailed distribution. This is so because players can lose interest in the game and abandon their characters. The lifespan of an individual, T_L, is defined as the number of days from the time the individual was created to the date of last logging. The distribution of the probability P_L that an individual has the lifespan T_L is shown in Fig. 5. This distribution can be approximated with the power law $P_L(T_L) \sim T_L^{-1.00\pm0.03}$. The average time T_L equals 69 days. For individuals who are active for more than one month, the average time T_L equals as many as 170 days. However, we found such a distribution of lifespan, with a different value of the exponent, in an internet community (www.grono.net)

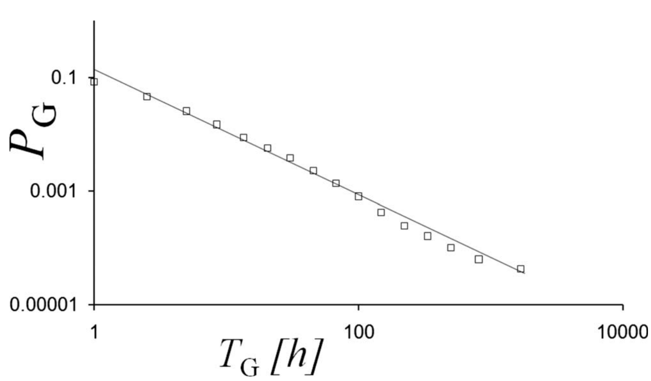

FIG. 4. Probability that an individual spends the time T_G playing the game (boxes) and fit to the power law $P_G(T_G) \sim T_G^{-1.10\pm0.03}$, $R^2=0.98$ (solid line). The data were log-binned to reduce uneven statistical fluctuations.

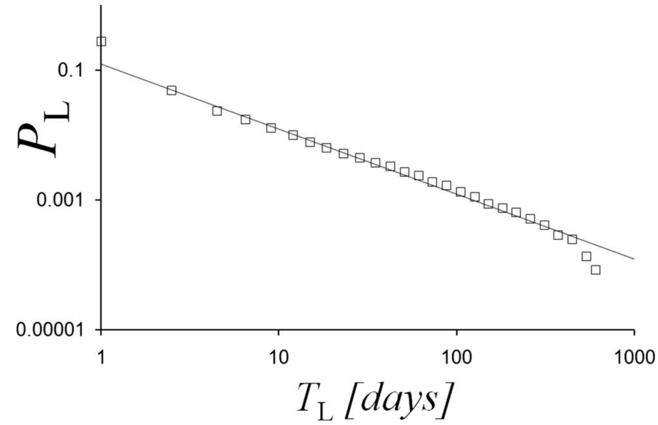

FIG. 5. Probability that an individual's activity in the virtual world lasted T_L days (boxes) and fit to the power law $P_L(T_L) \sim T_L^{-1.00\pm0.03}$, $R^2=0.98$ (solid line). The data were log-binned to reduce uneven statistical fluctuations.

of over 1×10^6 users [21]. The behavior of the number of active individuals as a function of time is not interesting and quickly (after few months) reaches saturation.

It is very interesting that both distributions, $P_G(T_G)$ and $P_L(T_L)$, have the power-law form with exponents close to -1. Similar relations concerning human dynamics have also been observed elsewhere [21] and can be a consequence of a decision-based queuing process. A model of such a process was recently proposed by Barabási [11–13]. It indicates that scale-free distributions are common in human dynamics. The distribution of the time devoted to entertainment (i.e., a character's lifespan) has a similar form to the distribution of the timing of many other human activities [11]. We suggest that different forms of activities (e.g., different forms of entertainment) compete with one another and can be ranked according to some perceived priority, like tasks in Barabási's model of a decision-based queuing process [11]. When a task (a form of activity) is added to a list, it is executed as long as it is on the list. The lower the task priority, the greater the probability that the task will be removed from the list. However, the model should be modified in order to obtain power-law distributions with different values of the exponents observed in real systems [21] (a detailed description of the model will be presented elsewhere). Therefore further work in this field is needed.

Knowing the accumulative time spent by the user in the virtual world T_G and the lifespan T_L, we can calculate the average time devoted daily to interactions in the virtual environment. By dividing this average time by 24 h we obtain the activity A of a character. Thus, the activity A_i denotes the probability that the ith character exists in the virtual world,

$$A = \frac{T_G}{24T_L}. \tag{2}$$

The value of activity relates to the whole lifespan of an individual. However, the activity is not biased by the lifespan of an individual, because the average time devoted daily for playing the game is approximately independent of T_L. It should be noted that similar results (e.g., lack of correlations

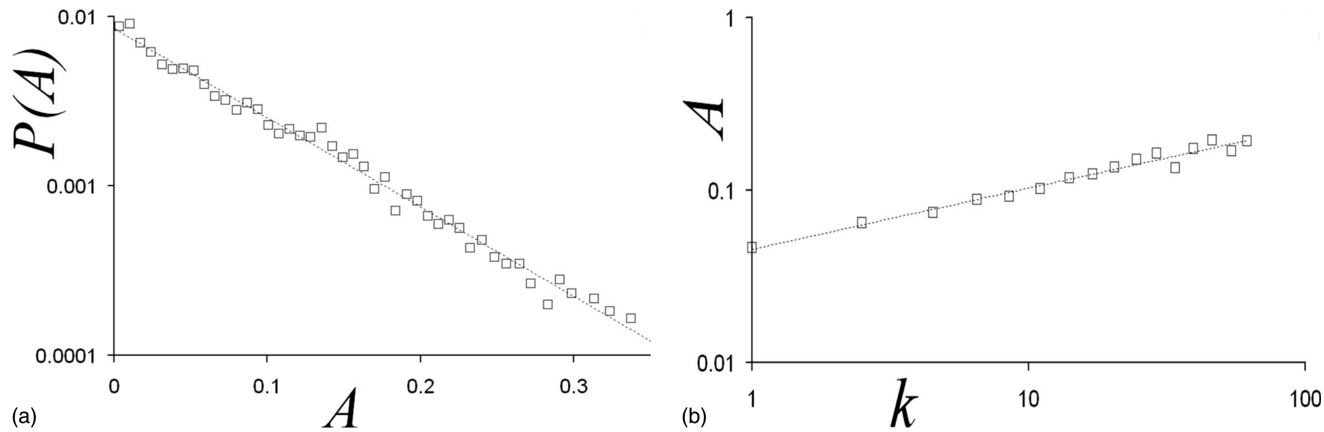

FIG. 6. Activity distribution (boxes) and fit to the exponential form (dashed line) $P(A) \sim \exp(-\mu A)$, where $\mu = 12.0 \pm 0.2$, $R^2 = 0.98$ (a). The relation between the degree of an individual and its activity (b). Results can be approximated with the power law $A(k) \sim k^{0.35}$ (dashed line), $R^2 = 0.97$.

between A and T_L) were found in analyzing the behavior of 5×10^6 players in the web service www.xfire.com.

The activity distribution $P(A)$ is exponential [Fig. 6(a)]. The maximal value A is two orders of magnitude greater than the minimal value of the activity. The relation between the degree of a character and its activity is shown in Fig. 6(b); the greater the k, the greater the A. Hence, the activity of a character is positively correlated with its degree and the results can be approximated with the power law

$$A(k) \sim k^{\eta} \tag{3}$$

where $\eta = 0.35 \pm 0.02$ [see dashed line in Fig. 6(b)]. The two relations $P(A)$ and $k(A)$ can be fitted using one-parameter fitting functions. On the basis of those relations we can calculate the form of the degree distribution: $P(k) = P(A(k))dA/dk$. After calculations we obtain $P(k) \sim k^{\eta-1} \exp(-\xi k^{\eta})$, where $\xi = 1.40 \pm 0.05$ ($R^2 = 0.99$). Hence, in the network under investigation, the degree distribution decays as a stretched exponential for large k (see Fig. 1). It is difficult to judge the validity of the approximation by R^2 alone. Therefore, to compare the goodness of fit to power-law and exponential decay we calculate the Kolmogorov-Smirnov ($\mathcal{K}$) statistic for three competing distributions (the lower the value of $\mathcal{K}$, the better the goodness of fit) [19]. The results for R^2 and the Kolmogorov-Smirnov statistic ($\mathcal{K} = 0.02$, 0.04, and 0.02 for shifted power law, power law with cutoff, and stretched exponential, respectively) indicate that the goodness of fit of the stretched exponential distribution is better than in the case of the other distributions.

Casual gamers spend their spare time playing computer games—meeting and cooperating with others there, as people who meet with others in real life spend their spare time together, e.g., going for a walk, in a restaurant, in a cinema, or at a marketplace. The characteristics of a friendship network are similar to those we can observe for real life social networks (e.g., children in one school) [1,16,25]. Thus, we can suppose that people's behavior in establishing and maintaining social contacts in an artificial society is similar to that in the real world. Therefore, we believe the relation between the time devoted to interactions with others (social activity) and the number of friends is similar in both worlds. Knowing the activity of individuals, we can start to investigate epidemic spreading and rumor propagation in the network.

V. EPIDEMIC SPREADING

In the literature there are many models of epidemic spreading with different mechanisms of contagion [26,27]. However, to understand better the influence of human activity on the process of spreading, we used a simple SIR model [7,28]. In our model, each individual is in one of three permitted states: healthy and susceptible (S), ill (I), healthy and unsusceptible or isolated from the rest of the population (R). The individual's state evolves in time and depends on their previous state and the connections or random contacts with other individuals. The probabilities of transitions between different states in one time step are described with the following parameters: $W_{S\to I}$, the probability that a susceptible individual will be infected by an ill individual (this also denotes how contagious the disease is); $W_{I\to R}$, the probability that an ill individual will recover or be isolated from the rest of the population (e.g., in a hospital).

In SIR models based on differential equations it is often assumed that an increase in the number of ill individuals N_I is linearly proportional to N_I [7,29]. We use a similar assumption in our model. We assume that the probability of an infection of an individual by one of k neighbors in one time step (one day) is a simple function of the number of ill neighbors. However, to distinguish the effectiveness of interactions between individuals we take into account the human activity A [see Eq. (2)]:

$$p_i = 24 W_{S\to I} A_i \sum_{j}^{k_i^I} A_j \tag{4}$$

where p_i is the probability of infection per one day, $W_{S\to I}$ is probability of infection per one hour of contact, k_i^I is the number of ill neighbors of the ith individual, and A_i is the activity of the ith individual. The probability of individuals

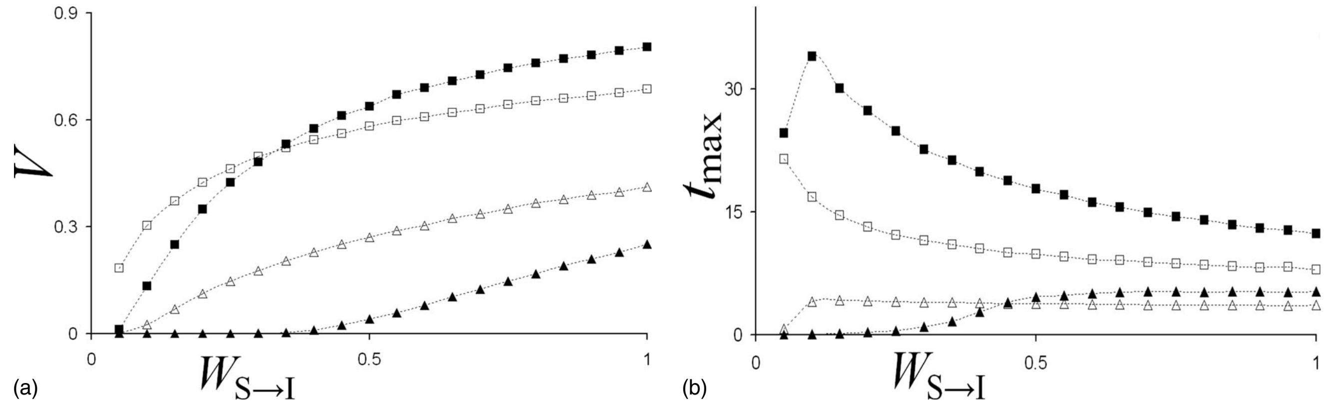

FIG. 7. Influence of the parameter $W_{S\to I}$ on the magnitude of the epidemic, V (a), and the time t_{max} (b) for different values of $W_{I\to R}$ (0.1 boxes and 0.9 triangles). Black and white markers correspond to uniform and real distributions of activity, respectively. Results were averaged over 10^4 independent simulations.

becoming infected is proportional to their activity; less active individuals (i.e., individuals who spend less time in the virtual world) have fewer opportunities to become infected than more active ones. Similarly, we take into account the neighbors' activity. The greater the activity of ill neighbors, the greater the probability of infection. Note that for A_i=const the probability p_i increases linearly with increasing k_i^I. Other probabilities of a transition between states X, Y in one time step do not depend on the structure of the network and human activity and they are described with the parameters $W_{X\to Y}$.

It was recently found that the heavy-tailed distribution of the intercontact times between susceptible and infected individuals has significant influence on the spreading of computer viruses [9]. However, it is noteworthy that the influence of human dynamics on the spreading process in the virtual world of MMORPGs is different than in the case of sending e-mails [9]. Sending an e-mail takes only a few minutes, but the players after logging in spend much longer times before they log off from the game (on average they devote daily approximately 3 h to playing the game). Most of the players (especially those with high connectivity) enter the virtual world every day (often at the same hour) and the distribution of interevent times (i.e., times between successive log-ins) cannot be approximated with a power law. Moreover, it significantly depends on the degree of an individual (players having more friends spend more time in virtual world; see Fig. 6). Next, in the process of propagation of information or a pathogen the most important factor is for how long people contact or talk with each other. In the case of computer viruses, the most important is the number of e-mails that were sent. Therefore inclusion of the distribution of interevent times within the proposed model would introduce nonessential complexity without giving better results.

Computations were performed for the initial conditions with one ill (I) and randomly located individual and the rest of the population healthy and susceptible (S). Synchronous dynamics were used with the assumption that individuals can change their state only once in each time step. To investigate the dynamics of the spreading process and the range of an epidemic, we introduced two observables: the time t_{max} when the maximal number of ill individuals is reached and the magnitude of the epidemic, V, defined as the relative number of individuals who had the disease during the epidemic.

The relation between the control parameters describing a disease and the observables (V and t_{max}) is shown in Fig. 7. To investigate the influence of human activity on the spreading process we made computations for two different distributions of activity, real and uniform A_i=const. To obtain more comparable results the average activity was the same in both distributions.

For large values of $W_{I\to R}$ the magnitude of the epidemic V increases with an increase in $W_{S\to I}$. In the case of real distribution of the activity, V is much larger and the value of the time t_{max} is lower. Hence, the epidemic spreads faster and reaches a larger part of the network. This is a result of the presence of very active superspreaders [30], i.e., individuals with large degree and a high value of social activity A. The process of the epidemic is highly influenced by superspreaders, because the probability that they are connected is high (the network is assortatively mixed by degree). Because of those individuals the epidemic reaches distant parts of the network very fast, even in the case of diseases that are not very contagious. For a uniform distribution of activity, superspreaders are less effective, because their activity is lower (the activity of a node is positively correlated with its degree; see Fig. 6). Therefore the epidemic cannot spread in the network (the magnitude of the epidemic is close to zero for low values of $W_{S\to I}$).

For low values of $W_{I\to R}$ a decrease in the time t_{max} is visible; hence the epidemic spreads faster with an increase in $W_{S\to I}$ [see Fig. 7(b)]. The magnitude of the epidemic remains great, even in the case of diseases that are not very contagious. As a result of the presence of more active superspreaders in the case of a real distribution of activity, the magnitude of the epidemic is greater.

However, for large values of $W_{S\to I}$ the opposite situation is true—the magnitude of the epidemic is greater in the case of a uniform distribution of the activity. The value of degree is positively correlated with the degree of the node. On the one hand, the social activity of an individual with a small number of connections is low, therefore the probability that such an individual will be infected is very low. On the other hand, the hubs, which are highly interconnected (the network

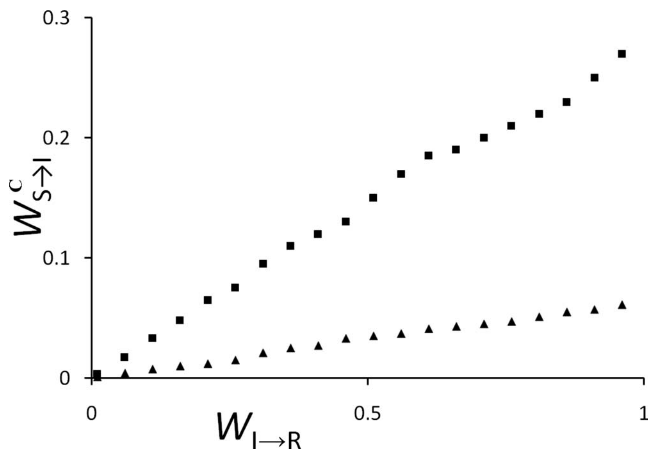

FIG. 8. Values of the parameters for which the number of infected individuals is approximately 1. Boxes and triangles correspond to uniform and real distributions of activity, respectively. Results were averaged over 10^4 independent simulations.

is positively correlated by degree) and very active, quickly spread the disease to other hubs. Thus, the epidemic spreads faster [see Fig. 7(b)], but dies out quickly. As a result, the probability that an individual with low k and low A will be infected during the epidemic is low. Therefore, the magnitude of the epidemic is smaller in the case of a real distribution of the activity, because the number of such individuals is high (see Fig. 1).

The SIR model [where the probability of infection with k^I ill neighbors equals λk^I; λ is the microscopic spreading (infection) rate, and infected individuals decay into the removed class at the rate $W_{I\to R}=1$] shows that the expression for the critical threshold is a function of the moments of the degree distribution [29] $\lambda_C=\langle k\rangle/\langle k^2\rangle$. If the value of λ is above λ_C, the disease spreads and infects a finite fraction of the population. On the other hand, when λ is below the threshold, the total number of infected individuals is infinitesimally small in the limit of very large populations. In networks with a strongly fluctuating degree distribution, the epidemic threshold approaches zero for increasing sizes of networks [29,31].

To investigate the epidemic threshold in our network we calculated the critical value of the parameter $W^C_{S\to I}(W_{I\to R})$ defined as follows: for $W_{S\to I}<W^C_{S\to I}$, the average number of individuals infected by an initially ill individual n_{I} is smaller than 1; and $n_{\mathrm{I}}>1$ for $W_{S\to I}>W^C_{S\to I}$. Even if we increase the size of the network, the value of $W^C_{S\to I}$ remains the same. The results for both distributions of activity are presented in Fig. 8. In the case of a real distribution of the activity $W^C_{S\to I}$ increases more slowly and takes smaller values (on average it is five times smaller) than in the case of a uniform distribution of the activity.

The results of numerical simulations can be compared to analytical calculations. In the case of a uniform distribution of the activity the probability of an individual becoming infected has the form

$$p=24W_{S\to I}\langle A\rangle^2 k^I=\lambda k^I, \tag{5}$$

where $\langle A\rangle$ is the average activity. Knowing that for $W_{I\to R}=1$ the critical value $W^C_{S\to I}\approx 0.28$ (see Fig. 8), we can calculate the critical value of the infection rate $\lambda_C\approx 0.04$. On the other hand, $\lambda_C=\langle k\rangle/\langle k^2\rangle$, hence in our network we obtain $\lambda_C=0.06$. The two results are similar. We suggest that the discrepancy is a result of the finite size effect.

VI. RUMOR PROPAGATION

The next phenomenon that we study in this paper is rumor propagation in a social network. We investigate the impact of human dynamics and interaction rules on the efficiency of rumor propagation. The model of the rumor is defined as follows. Each of the N individuals can be in three different states. Following the original terminology [32,33], those three classes correspond to ignorant (IG), spreader (SP), and stifler (ST) individuals. Ignorants are those individuals who have not heard the rumor and hence are susceptible to being informed. Spreaders are active individuals who spread the rumor. Stiflers know the rumor but are no longer interested in spreading it. As result of interactions with spreaders, an ignorant individual turns into another spreader with the probability p^{SP},

$$p_i^{\mathrm{SP}}=W_{\mathrm{IG}\to\mathrm{SP}}A_i\sum_j^{k_i^{\mathrm{SP}}}A_j \tag{6}$$

where k_i^{SP} is the number of neighbors in the state SP and $W_{\mathrm{IG}\to\mathrm{SP}}$ is the parameter that describes how interesting the rumor is.

Propagation may decay because of the mechanism of "forgetting" or because spreaders learn that the rumor has lost its "news value." In the first case rumor propagation is similar to epidemic spreading, which is described in Sec. V. The second assumption seems to be more plausible. We study two different interaction rules: (1) spreaders become stiflers as a result of interactions with other active individuals with the probability

$$p_i^{\mathrm{ST-1}}=W_{\mathrm{SP}\to\mathrm{ST}}A_i\sum_j^{k_i^{\mathrm{SP}}}A_j \tag{7}$$

and (2) spreaders become stiflers if they encounter other spreaders or stiflers with the probability

$$p_i^{\mathrm{ST-2}}=W_{\mathrm{SP}\to\mathrm{ST}}A_i\left(\sum_j^{k_i^{\mathrm{SP}}}A_j+\sum_j^{k_i^{\mathrm{ST}}}A_j\right) \tag{8}$$

where k_i^{ST} is the number of neighbors in the state ST and $W_{\mathrm{SP}\to\mathrm{ST}}$ is the parameter that describes how fast the rumor loses its attractiveness.

Computations were performed for the initial conditions with one randomly located spreader and the rest of the population in the state IG. Synchronous dynamics were used with the assumption that individuals can change their state only once in each time step. As in the previous case we introduce two observables: the time $t_{\max}$ when the maximal number of spreaders is reached and the relative number of individuals affected by the rumor V. The results of simulations are shown in Fig. 9. In all simulations the value of V increases with an increase in $W_{\mathrm{IG}\to\mathrm{SP}}$ and a decrease in $W_{\mathrm{SP}\to\mathrm{ST}}$.

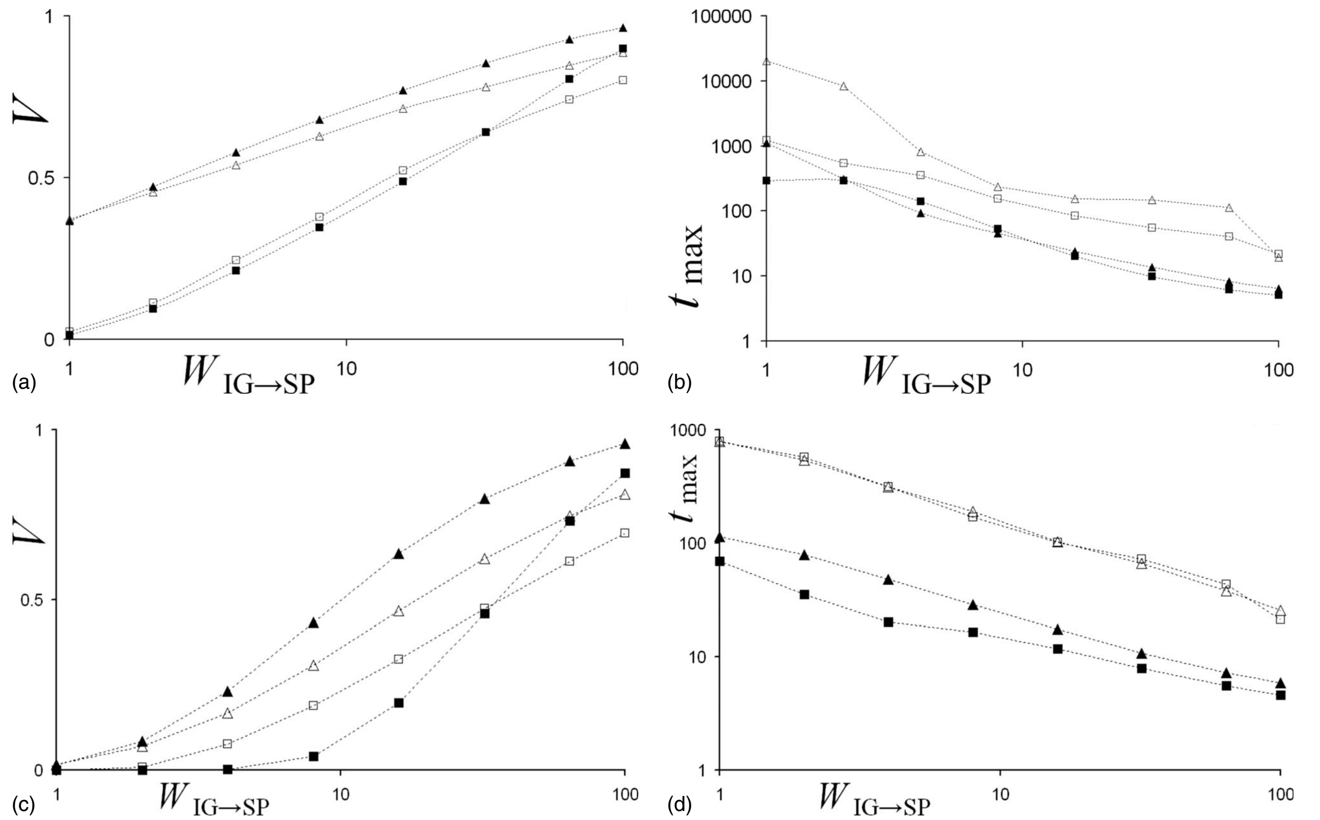

FIG. 9. Influence of the parameter $W_{IG \to SP}$ on the relative number of individuals affected by the rumor V (a), (c) and the time t_{max} (b), (d) for different values of $W_{SP \to ST}$ (5 triangles and 40 boxes). Black and white markers correspond to uniform and real distributions of activity, respectively. (a), (b) and (c), (d) correspond to interaction rules 1 and 2, respectively. Results were averaged over 10^4 independent simulations.

In the case of interaction rule 2 the probability that a spreader becomes a stifler is higher and the average time in which an individual can spread the rumor is lower. As a result of positive correlation by degree, spreaders with a high degree (superspreaders) become stiflers even faster, because the probability that they are connected to other superspreaders or stiflers with a high degree is higher. In this way superspreaders are quickly deactivated. Hence, the final number of individuals affected by the rumor is lower than in the case of the interaction rule 1. The time t_{max} is also lower, because the rumor cannot spread freely as a result of superspreaders' quick deactivation. The discrepancy in the results is more visible for low values of $W_{IG \to SP}$ (see Fig. 9), because for high values of $W_{IG \to SP}$ the probability that an ignorant will be affected by the rumor is high enough and superspreaders can spread the rumor to distant parts of the network before they turn into stiflers.

In contrast to epidemic spreading, the influence of human activity on V is less visible (cf. Fig. 7). Moreover, in the vast range of values of control parameters, it has an opposite effect; the final number of individuals affected by the rumor is higher in the case of a uniform distribution of activity. Similar changes are observed in the dynamics of spreading, the rumor spreads much slower (especially for interaction rule 1) in the case of a real distribution of the activity.

The system behaves like that because of the positive correlation between the degree k and the activity A of a node observed in a real distribution of human activity (see Fig. 6). Because individuals with a high degree have a high value of A, they more effectively interact with their neighbors than in the case of a uniform distribution $P(A)$. Such individuals in the state SP very quickly (but faster in the case of interaction rule 2) learn that the rumor has lost its news value and turn into stiflers.

The influence of degree on the average lifespan of spreaders, i.e., the average number of time steps before a spreader turns into a stifler, is shown in Fig. 10(a). The length of a lifespan depends significantly on degree. However, the pattern of this relation is different for different distributions of the activity. In the case of the real distribution $P(A)$, the spreaders' lifespan is much higher in the range of low degree and lower for high k. Individuals with large degree (e.g., superspreaders) remain approximately one time step in the spreader state. Hence, individuals with low k and low A mainly spread the rumor and the rate of spreading is much slower. Similar results have been observed in some models of epidemic spreading where high-degree nodes are vaccinated [26]. For a number of preventively vaccinated hubs slightly lower than the critical value, there was an abrupt increase in the time t_{max} and the magnitude of the epidemic was smaller.

In such conditions, when superspreaders do not spread a rumor, many nodes with low degree (and low A) which are connected to other parts of the network with hubs cannot

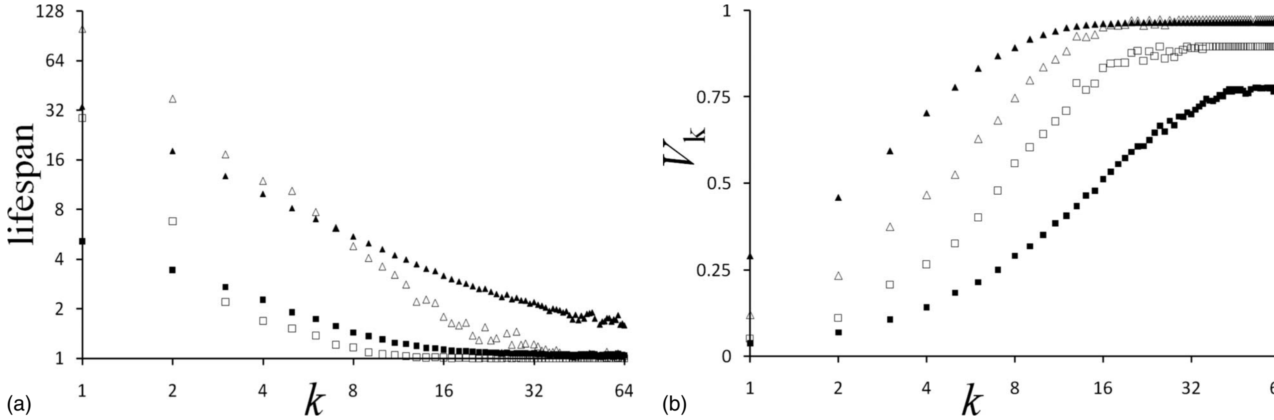

FIG. 10. Relation between degree of an individual k and the average lifespan of spreaders (a) and the relative number of individuals affected by the rumor V_k (b) for different values of $W_{SP\to ST}$ (5 triangles and 40 boxes) and for the second interaction rule. Values of other parameters: $W_{IG\to SP}=16$. Black and white markers correspond to uniform and real distributions of activity, respectively. Results were averaged over 10^4 independent simulations.

hear the rumor. The probability that an individual hears the rumor decreases with decreasing k, and is lower in the case of the real distribution $P(A)$ in the large range of control parameters [see Fig. 10(b)]. Because most individuals have low degree, the number V is lower in the case of a real distribution of the activity [see Figs. 9(a) and 9(c)].

However, for low values of $W_{IG\to SP}$, and low values of $W_{SP\to ST}$, the value of V is slightly higher in the case of a real distribution than in the case of a uniform distribution of $P(A)$. This is so because highly interconnected hubs are very active and can spread the rumor to other hubs, before they turn into stiflers (even if the rumor is not very interesting, low $W_{IG\to SP}$). The rumor spreads mainly among individuals with large degree. In the range of low k the relative number V_k of individuals with the degree k affected by the rumor has similar values for both distributions of the activity. For a uniform distribution of $P(A)$ the social activity of hubs is lower; hence, the probability that a hub will hear the rumor is also lower (see Fig. 10). The discrepancy in V_k increases with increasing k, before the $V_k(k)$ relation reaches saturation.

VII. CONCLUSIONS

We have shown that a friendship network maintained in a virtual world has similar properties (e.g., large clustering, a short average path length, assortative mixing by degree) to other social networks. On the basis of the playing time of each individual recorded on the server, we have presented results concerning human dynamics. The power-law form of the distributions $P_G(T_G)$, $P_L(T_L)$, and $A(k)$ and other authors' results [11] indicate that such a scaling law is common in human dynamics and should be taken into account in models of the evolution of social networks and of human activity. We have also shown that meeting people in a virtual society has a much greater influence on the real society than the other way round.

We have found that the distribution of the activity A of individuals (i.e., the relative time daily devoted to interactions with others) is exponential. It should be stressed that the activity of an individual is positively correlated with its degree.

The next conclusion is that human activity has a significant influence on the dynamic processes in a social network. The process of epidemic spreading in a social network has been investigated numerically.

It turned out that an epidemic spreads faster and for a large range of values of the control parameters the magnitude of an epidemic is greater in the case of a real distribution of activity, as a result of the presence of very active superspreaders (individuals with a high degree and a large value of social activity) in the social network. In the case of an epidemic in a real population the pattern of human behavior will change. In times of severe epidemics, people decrease the time devoted daily to interactions with others (activity A) to avoid infection. However, we suggest that in the case of limited epidemics (low values of $W_{S\to I}$), when the number of ill individuals is relatively low (e.g., annual influenza epidemics), the change in social activity is low too. Therefore, models that take into account data on human social activity seem to be more plausible for modeling epidemic spreading in the human population than models that do not take into account human dynamics, e.g., models of computer viruses [9].

Moreover, rumor propagation in a real social network has been investigated numerically. We studied two different interactions rules: (1) spreaders become stiflers as a result of interactions with other active individuals, and (2) spreaders become stiflers if they encounter other spreaders or stiflers. For both interaction rules, the influence of a real distribution of activity is the same. In contrast to epidemic spreading, rumor propagation is slower and for a vast range of values of control the final number of individuals affected by the rumor is lower.

We have found that taking into account a real distribution of social activity speeds up epidemic spreading; however, it decreases rumor propagation. Our results indicate that the influence of human social activity on dynamic phenomena in social networks significantly depends on the types of phenomena and interaction rules involved.

In future work, we will consider investigating the influence of human social activity on epidemic spreading for different mechanisms of contagion. Next, of further interest would be a more careful exploration of rumor propagation and opinion formation in on-line communities taking into account the lifespan T_L of an individual.

ACKNOWLEDGMENTS

We wish to thank Maciej Krzysztofik and Agnieszka Taper, the heads of the Allseron Team, for help in the preparation of this paper. N.K. wishes to thank A.Gadomski for help and useful discussions. A. G. acknowledges financial support from the Foundation for Polish Science (FNP).

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[2] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks* (Oxford University Press, Oxford, 2004).
[3] D. B. West, *Introduction to Graph Theory* (Prentice-Hall, Englewood Cliffs, NJ, 2001).
[4] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Aberg, Nature (London) **411**, 907 (2001).
[5] H. Ebel, L. I. Mielsch, and S. Bornholdt, Phys. Rev. E **66**, 035103(R) (2002).
[6] G. Csanyi and B. Szendroi, Phys. Rev. E **69**, 036131 (2004).
[7] N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases* (Springer, Berlin, 1993).
[8] P. S. Dodds and D. J. Watts, Phys. Rev. Lett. **92**, 218701 (2004).
[9] A. Vazquez, B. Rácz, A. Lukács, and A.-L. Barabási, Phys. Rev. Lett. **98**, 158702 (2007).
[10] The Allseron Team kindly provided research data (www.allseron.com).
[11] A.-L. Barabási, Nature (London) **435**, 207 (2005).
[12] J. G. Oliveira and A.-L. Barabási, Nature (London) **437**, 1251 (2005).
[13] A. Vázquez, Phys. Rev. Lett. **95**, 248701 (2005).
[14] A. Grabowski and N. Kruszeska, Int. J. Mod. Phys. C **18**, 1 (2007).
[15] M. E. J. Newman and J. Park, Phys. Rev. E **68**, 036122 (2003).
[16] M. C. González, P. G. Lind, and H. J. Herrmann, Phys. Rev. Lett. **96**, 088702 (2006).
[17] S. H. Strogatz, Nature (London) **410**, 268 (2001).
[18] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).
[19] A. Clauset, C. R. Shalizi, and M. E. J. Newman, e-print arXiv:0706.1062.
[20] M. E. J. Newman, Phys. Rev. E **67**, 026126 (2003).
[21] A. Grabowski, Physica A **385**, 363369 (2007).
[22] E. Ravasz and A. L. Barabási, Phys. Rev. E **67**, 026112 (2003).
[23] L. F. Costa, Phys. Rev. Lett. **93**, 098702 (2004).
[24] S. N. Soffer and A. Vazquez, Phys. Rev. E **71**, 057101 (2005).
[25] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
[26] A. Grabowski and R. A. Kosinski, Phys. Rev. E **70**, 031908 (2004).
[27] D. Centola, V. M. Eguiluz, and M. W. Macy, Physica A **374**, 449 (2007).
[28] W. O. Kermack and A. G. McKendrick, Proc. R. Soc. London, Ser. A **115**, 700 (1927).
[29] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, Eur. Phys. J. B **26**, 521 (2002).
[30] M. Barthelemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. **92**, 178701 (2004).
[31] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **65**, 035108(R) (2002).
[32] D. J. Daley and D. G. Kendall, Nature (London) **204**, 1118 (1964).
[33] D. J. Daley and J. Gani, *Epidemic Modeling* (Cambridge University Press, Cambridge, U.K., 2000).