Int. J. Appl. Math. Comput. Sci., 2014, Vol. 24, No. 3, 683–696
DOI: 10.2478/amcs-2014-0050

MODELING ACQUAINTANCE NETWORKS BASED ON BALANCE THEORY

VIDA VUKAŠINOVIĆ *, JURIJ ŠILC *, RISTE ŠKREKOVSKI **,***

* Computer Systems Department
Jožef Stefan Institute, Jamova cesta 39, 1000 Ljubljana, Slovenia
e-mail: {vida.vukasinovic,jurij.silc}@ijs.si

**Department of Mathematics
University of Ljubljana, Jadranska ulica 19, 1000 Ljubljana, Slovenia

***Laboratory of Data Technologies
Faculty of Information Studies, Ulica talcev 3, 8000 Novo mesto, Slovenia

An acquaintance network is a social structure made up of a set of actors and the ties between them. These ties change dynamically as a consequence of incessant interactions between the actors. In this paper we introduce a social network model called the Interaction-Based (IB) model that involves well-known sociological principles. The connections between the actors and the strength of the connections are influenced by the continuous positive and negative interactions between the actors and, vice versa, the future interactions are more likely to happen between the actors that are connected with stronger ties. The model is also inspired by the social behavior of animal species, particularly that of ants in their colony. A model evaluation showed that the IB model turned out to be sparse. The model has a small diameter and an average path length that grows in proportion to the logarithm of the number of vertices. The clustering coefficient is relatively high, and its value stabilizes in larger networks. The degree distributions are slightly right-skewed. In the mature phase of the IB model, i.e., when the number of edges does not change significantly, most of the network properties do not change significantly either. The IB model was found to be the best of all the compared models in simulating the e-mail URV (University Rovira i Virgili of Tarragona) network because the properties of the IB model more closely matched those of the e-mail URV network than the other models.

Keywords: social network, small world, balance theory, pheromone infrastructure.

1. Introduction

We still do not understand human social systems well enough, but in this study we try to come one step closer. In the past few decades researchers have extensively investigated networks in technology, biology, and sociology. For instance, real-world networks are include *technological networks* (such as road networks, railways, and electric power grids), *biological networks* (such as neural networks, metabolic paths, and vascular networks in plants), *information networks* (such as citation networks between academic papers and the World Wide Web), and *social networks* (such as e-mail networks, phone-call networks, and research-collaboration networks).

Researchers try to specify the network's structure and understand how networks evolve. Three components in network studies have been identified (Newman, 2008). First, in empirical studies researchers try to picture the connections between individuals by using different techniques. These include interviews, the observation of individuals, and the use of archival records. Then, empirical data can be further investigated using mathematical and statistical techniques. Finally, based on the properties extracted from the network, we can build mathematical models to simulate dependent processes that allow us to predict the behavior of the network (Barabási and Albert, 1999; Davidsen *et al.*, 2002; Erdös and Rényi, 1960; Jiang *et al.*, 2011; Kumpula *et al.*, 2007; Leskovec, 2010; Ludwig and Abell, 2007; Marsili *et al.*, 2004; Péter, 2012; Wang *et al.*, 2005; White *et al.*, 2006; Xiong *et al.*, 2011).

Some of the most interesting networks are social networks that represent the relationships and flows between people, groups, and animals (Freeman, 2004).

The social behavior of animals has been observed for a long time, and the social patterns of simple animal species such as birds, bees, and ants can yield insights into human social behavior (Camazine *et al.*, 2001; Wynne, 2001).

For example, ants can find the shortest path from a food source to the nest. While walking from the food source to the nest, ants lay a chemical substance called pheromone on the trail. This pheromone slowly evaporates. During the food search, each ant probabilistically prefers to follow a direction rich in pheromone (Grassé, 1959). This paradigm was successfully used in combinatorial and continuous optimization (Dorigo *et al.*, 2000; 1996; Korošec, 2006; Korošec *et al.*, 2012) and network analysis (Handl *et al.*, 2006; Martens *et al.*, 2007).

In this paper we combine the idea of pheromone deposition and evaporation with balance theory (Heider, 1946), where interactions between individuals change their connection strength. These interactions between individuals are modeled using pheromone and they appear according to balance theory. Thus, human social systems are living structures where many simultaneous interactions are happening between individuals all the time and each interaction influences future interactions. We present and study a new interaction-based model where each interaction between two individuals influences future interactions in a similar way to how pheromone trails influence the ant's future decision about its path.

2. Social-network properties

The term "social network" describes a social structure determined by the interactions between individuals, groups, and organizations. For a summary of the progress of social networks and social-network analysis, see the work of Freeman (2004). However, the major developments in the field have occurred since the beginning of the twentieth century, when graph theory became more involved with statistics and algorithms. Most real networks possess non-trivial topological features. Their structure changes dynamically and lies between order and perfect randomness.

Large real-world graphs do not only have large numbers of vertices, but they also tend to be sparse, clustered, and have a small diameter. These kinds of graphs are termed small-world graphs (Watts, 1999). Let G be a graph comprising of a set of vertices $V(G)$ together with a set of edges $E(G)$. The diameter d of the graph G is the longest shortest path between any two graph vertices:

$$d(G) = \max_{u,v \in V(G)} d(u,v),$$

where $d(u,v)$ is the distance between the vertices u and v. The average path length is defined as follows:

$$l(G) = \sum_{u,v \in V(G)} \frac{d(u,v)}{n(n-1)},$$

where n is the graph order. Some of the first studies on distances in social networks were made by Rapoport (1957) and later by de Sola Pool and Kochen (1978), while at about the same time Milgram (1967) conducted the first empirical work. In his experiment, Milgram prepared letters addressed to a person in Massachusetts and asked some people in Nebraska to send a letter to an acquaintance who might be closer to the target person. The average number of intermediate steps in the trials that successfully reached the target person was found to lie between five and six. From this result comes the well-known "six degrees of separation" principle.

The local clustering coefficient of a vertex measures how close the vertex neighbors are to being a complete graph. The average clustering coefficient of the graph G is the average local clustering coefficient of the vertices. This coefficient is calculated as

$$c(G) = n^{-1} \sum_{i=1}^{n} \binom{|N(i)|}{2}^{-1} |E(N(i))|,$$

where $N(i)$ is the neighborhood of the vertex i and $|E(N(i))|$ is the number of edges in the graph induced by $N(i)$. For the sake of simplicity, the average clustering coefficient is referred to as the clustering coefficient.

Small worlds have $O(n \log n)$ edges, the diameter is $O(\log n)$ and the clustering coefficient is a small constant independent of the network size. Watts and Strogatz (1998) found a natural way to build small-world graphs by interpolating between lattice and random graphs. They began with a regular lattice and then rewired every edge with a given probability p. Kleinberg (2000) criticized the Watts and Strogatz (1998) model because it does not consider Milgram's (1967) findings that individuals using local information are collectively effective at constructing short paths between two vertices in a social network. Kleinberg (2000) built a model for which decentralized algorithms are effective.

The strength of division of a network into communities can be measured by a modularity measure. The modularity Q measures the fraction of the edges inside the given communities minus the expected such fraction if the edges were distributed at random. It is defined as

$$Q(D(G)) = \frac{1}{2m} \sum_{i,j} \Big(w_{ij} - \frac{k_i k_j}{2m} \Big) \delta_{c_i, c_j},$$

where $D(G)$ is a division of the graph G into communities, w_{ij} is the weight on the edge ij, $k_i = \sum_j w_{ij}$, c_i is the community to which the vertex i is

assigned, $m = \frac{1}{2}\sum_{ij} w_{ij}$, and δ_{c_i,c_j} is the Kronecker delta of c_i and c_j.

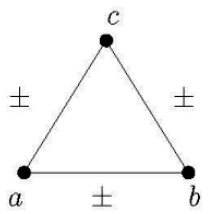
3. Sociological principles

Most of the noticeable interactions that happen to an average person during a day are the interactions between the person and his or her acquaintances or the acquaintances of his or her acquaintance. We can more generally say that in social networks the connections between individuals are formed on the basis of local information about the network (Grossetti, 2005). Based on this principle, Davidsen *et al*. (2002) simulated the evolution of social networks. The dynamics of their model are defined by randomly linking up neighboring vertices: when a vertex does not have enough neighbors, one other randomly chosen vertex is linked.

Interactions between people can cause positive or negative sentiments. These sentiments define how people perceive their relations with other people. Heider's balance theory (1946) says that two people equivalently perceive their relations as either positive or negative. Three people perceive their triadic relations as balanced or imbalanced, depending on the number of positive and negative relations between them. If there are three positive or one positive and two negative relations, they perceive it as balanced. Otherwise, if there are one or three negative relations, they perceive it as imbalanced, as presented in Fig. 1.

From the perspective of a person, the balanced triadic relations can be understood by the following statements: "friend of my friend is my friend", "friend of my enemy is my enemy", "enemy of my friend is my enemy", and "enemy of my enemy is my friend". According to social-balance theory, imbalanced triads are uncomfortable for the people involved. Moreover, they could cause a reorganization of the entire network. The people in the imbalanced triads tend to make them balanced or some connections may break. Some studies on this subject can be found in the works of Antal *et al*. (2005), Marvel *et al*. (2011) and Srinivasan (2011).

Based on later insights, Ludwig and Abell (2007) built an evolutionary social-network model by sequentially attaching randomly selected positive and negative edges to a given set of vertices. During each time step they take care that each vertex keep a small number of imbalanced triads.



Triadic relation	ab	ac	bc
Balanced	+	+	+
	−	+	−
	+	−	−
	−	−	+
Imbalanced	+	−	+
	−	+	+
	+	+	−
	−	−	−

Fig. 1. Balanced and imbalanced triadic relations where every two individuals perceive their relation as being positive or negative.

The differences in the relations are not just in terms of positive and negative perceptions, but also in how close (i.e., strong) the relations are between people. Therefore, weighted networks were introduced: to the edges of these networks are assigned appropriate weights, where the weight quantifies the strength of the given relation. It is natural to expect that weights have an influence on the formation of the network. The weak-link hypothesis (Granovetter, 1973) says that weak ties between people connect two or more communities and keep the network connected, whereas strong ties are mostly inside communities.

Kumpula *et al*. (2007) presented a model where the weights are generated dynamically. They distinguish two mechanisms of tie formation. The first one refers to forming ties with one's network neighbor, while the other refers to forming ties between people who share the same activities, independent of the distance inside the network. In their model these two mechanisms are called local attachment and global attachment. The local attachment process is a local two-step weighted self-avoiding random walk. Here the visited edges and the edge between the first and last vertex in the weighted random walk are reinforced by a small amount. In the global attachment a random edge is established with some probability p. Such a model proves the weak-link hypothesis.

4. Interaction-based model

In this work we are interested in how networks are formed if we consider sociological principles on the level of everyday interactions. We have built a new social network model, termed the Interaction-Based (IB) model, founded on the fact that everyday interactions keep the ties between us. The interactions between individuals cause new connections to be established or the strength of the connections to change. The more interactions happen between two people, the more important this relation is for them on a daily basis. The interactions can be positive or negative, and they appear as described in the previous section. These interactions have a strong influence on how individuals perceive their relations. The more often the interaction between two people is positive, the more friendly their relation is, and therefore their connection is stronger. Negative interactions between friends make their friendship weaker, until the connection breaks or becomes negative.

According to balance theory (Heider, 1946), interactions between individuals in triads tend to be balanced and the relations, as they are, are the results

Algorithm 1. Interaction-based model.

```
while not EndingCondition() do
  for each vertex i do
    GlobalInteractions(i, M_G)
    LocalInteractions(i, M_L)
  end for
  Evaporation(E)
end while
```

Algorithm 2. GlobalInteractions(i, M).

```
for M times do
  choose u with probability (d(u)+1)/(Σ_{v∈V(G)} d(v)+N)
  w_iu = w_iu + δ with probability p_G
  otherwise w_iu = w_iu − δ
end for
```

Algorithm 3. LocalInteractions(i, M).

```
for M times do
  //choose i's neighbor j
  if N(i)! = null then
    choose j ∈ N(i) with probability p_ij = |w_ij|/s_i,
    where
    s_i = Σ_{l∈N(i)} |w_il|
  else
    j is chosen at random
    edge ij is established
  end if
  //choose j's neighbor k
  if |N(j)| > 1 then
    choose k ∈ N(j) \ {i} with probability p_jk =
    |w_jk|/(s_j − |w_ij|)
  else
    k is chosen at random
    edge jk is established
  end if
  for edges e = ij, jk do
    w_e = w_e + δ with probability p_L
    otherwise w_e = w_e − δ
  end for
  according to balance theory w_ik = w_ik ± δ
end for
```

Algorithm 4. Evaporation (E).

```
for each edge e ∈ E do
  if w_e > 0 then
    w_e = w_e − ϵ
  else
    w_e = w_e + ϵ
  end if
end for
```

of balanced interactions between the individuals. It seems that the interactions between people who are connected with stronger ties take place more often, being either positive or negative. There are some connections between people that are not strong, but do they not seem to be affected by anything. These are usually weak ties between people from different communities (Granovetter, 1973). We model the network growth by simulating the interactions between individuals in dyads (global interactions) and triads (local interactions), see Algorithm 1.

With the interactions in dyads we simulate the interactions that happen between people that have the same interests and their interaction is the result of the common interest, e.g., attending the same language course, see Algorithm 2. These interactions are independent of the local knowledge, which means they can happen between some acquaintances or between complete strangers. However, the people with more acquaintances have a greater ability to interact and make new connections than the others.

With the interactions in triads we simulate the interactions that are influenced by the knowledge of a local network for a given individual $i \in V(G)$, see Algorithm 3. The interactions happen more often between individuals whose absolute value of the weights of their connections are higher. In these triads usually at least two people know each other.

In addition, people tend to forget, and if there are no interactions the connections lose their strength; being positive or negative. This is done with the function Evaporation() described by Algorithm 4. Algorithm 1 ends after t time steps.

We consider a fixed size network of N vertices. At every time step each vertex establishes M_{G} new interactions using the preferential random interaction rule. The probability of choosing a particular vertex increases with the vertex degree value. The weight of the connection is increased by δ with the probability p_{G} and decreased otherwise.

Furthermore, at each time step, for each vertex a simulation of the interactions influenced by the knowledge of the local network is repeated M_{L} times, see Fig. 2. The vertex i chooses its neighbor j with the probability $|w_{ij}|/s_i$, where w_{ij} is the weight of the edge between the vertices i and j, and $s_i = \sum_{j\in N(i)} |w_{ij}|$ is the strength of i. The weight w_{ij} is increased by δ with the probability p_{L} and it is decreased by δ with the probability $1 - p_{\mathrm{L}}$. If the chosen vertex j has other neighbors besides i, it chooses from them a vertex k with the probability $|w_{jk}|/(s_j - |w_{ij}|)$. Like before, w_{jk} is increased by δ with the probability p_{L} and it is decreased by δ with the probability $1 - p_{\mathrm{L}}$. According to balance theory, the link between i and k is established or reinforced by a positive or negative amount of δ. If the vertex i or j

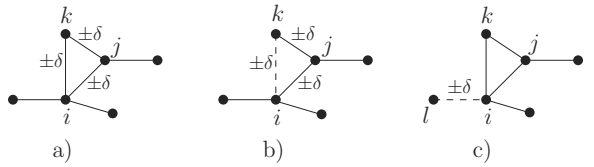


Fig. 2. Simulation of local interactions and preferential random interaction rules. A simulation of local interactions starts at vertex i, proceeds to vertex j and then to k. Weights on ij and jk are increased by δ with probability p_{L} and decreased by δ with probability $1 - p_{\mathrm{L}}$. Weight on ik is established or refreshed by $\pm\delta$ considering balance theory: vertices j and k are i's neighbors (a), if vertex k is not i's neighbor, a new edge is established (b). A vertex l is chosen by the preferential random interaction rule and the edge il is established or refreshed by δ with probability p_{G} and by $-\delta$ with probability $1 - p_{\mathrm{G}}$: example where vertices i and l are not neighbors (c).

has no neighbors, it creates a connection to a randomly chosen vertex, where the connection is positive with the probability p_{L} and negative otherwise.

After each time step, the absolute values of the weights on the edges are decreased by some amount ε and the connections break if their weight is smaller than ε. Thus, if the weight is positive, then it is reduced by ε, and if the weight is negative, then it is increased by ε. The weights of the edges lose their strength in a similar way in which the pheromone amount on an ant's trails evaporates.

5. Model evaluation

The basic version of the model, where $p_+ = p_{\mathrm{L}} = p_{\mathrm{G}}$, $M_{\mathrm{L}} = M_{\mathrm{G}} = 1$, and $\varepsilon = \delta/2$, is presented by Vukašinović *et al.* (2012a; 2012b).

In this paper we study the interaction-based model from three different aspects. First, we provide a detailed evaluation of the various parameters and we observe their influence on the network structure obtained in the mature phase. Second, we are interested how resistant the IB model is to edge malfunctions. We did the experiments on the IB model's resistance by adding a large amount of negative interactions and by deleting the edges between vertices in the network. Third, we compare the IB network's structure to that of some other models and a real-world network.

The proposed IB model is studied by simulations that start from an empty graph of N vertices. The model was tested for networks of the order $N = 1000$ and different parameter values of M_{L}, M_{G}, ε, p_{L}, p_{G}, and t. The meanings of the parameters are as follows:

- M_{L}, number of times the simulation of local interactions is repeated for each vertex in LocalInteractions(i, M_{L}),
- M_{G}, number of times the simulation of global interactions is repeated for each vertex in GlobalInteractions(i, M_{G}),
- ε, number of the evaporation factor,
- p_{L}, number of the probability that the interaction in the simulation of the local interactions process is positive,
- p_{G}, number of the probability that the interaction in the simulation of the global interactions process is positive,
- t, number of the number of time steps.

For each set of parameters, we carried out 50 runs and computed the average properties of the networks.

Figure 3 shows the influence of the parameter M_{L} on the network formation at $\varepsilon = 0.5\delta$, $p_{\mathrm{L}} = 0.6$, $p_{\mathrm{G}} = 0.9$, $t = 150$ and different values of the parameter M_{G}. The number of edges slowly increases with the number of repeated weighted local searches for new acquaintances M_{L}, see Fig. 3(a). As M_{L} increases, the clustering coefficient also increases, see Fig. 3(b). Eventually, its value starts to decrease. This trend is very clear when $M_{\mathrm{G}} = 0$, but is also expected to a lesser extent at higher values of M_{L} when $M_{\mathrm{G}} = 1$ and $M_{\mathrm{G}} = 5$. At the point when the slope of the curve becomes negative, we can expect that certain edges are reinforced so much that new edges and edges with a smaller absolute value of weight have a low probability of being chosen. The diameter does not change significantly with M_{L} at $M_{\mathrm{G}} = 1$ and $M_{\mathrm{G}} = 5$, while its dynamics are more unpredictable at $M_{\mathrm{G}} = 0$, i.e., in the case with no global interactions (Fig. 3(c)). In this case it reaches its minimum at $M_{\mathrm{L}} = 3$. The number of triangles and balanced triangles increases almost linearly with M_{L}, see Fig. 3(d).

Figure 4 shows the influence of the parameter M_{G} on the network formation at $\varepsilon = 0.5\delta$, $p_{\mathrm{L}} = 0.6$, $p_{\mathrm{G}} = 0.9$, $t = 150$ and different values of the parameter M_{L}. The number of edges and the number of positive edges increase linearly with M_{G}, see Fig. 4(a), while the number of triangles and the number of balanced triangles have superlinear growth, see Fig. 4(d). The clustering coefficient (Fig. 4(b)) decreases until it stabilizes. The diameter d and the average path length l decrease while M_{G} increases, see Fig. 4(c).

M_{G} has a larger influence on the number of edges than M_{L}. With global interactions a lot of new connections are established, while with local interactions mostly old connections are reinforced, see Figs. 3 and 4.

Figure 5 shows the influence of the parameter p_{L} on the network formation at $\varepsilon = 0.5\delta$, $t = 150$, $M_{\mathrm{L}} = 5$, $M_{\mathrm{G}} = 1$ and different values of the parameter p_{G}. The results show that, while we change the parameter p_{L}, the different values $p_{\mathrm{G}} = 0.6$ and $p_{\mathrm{G}} = 0.9$ do not

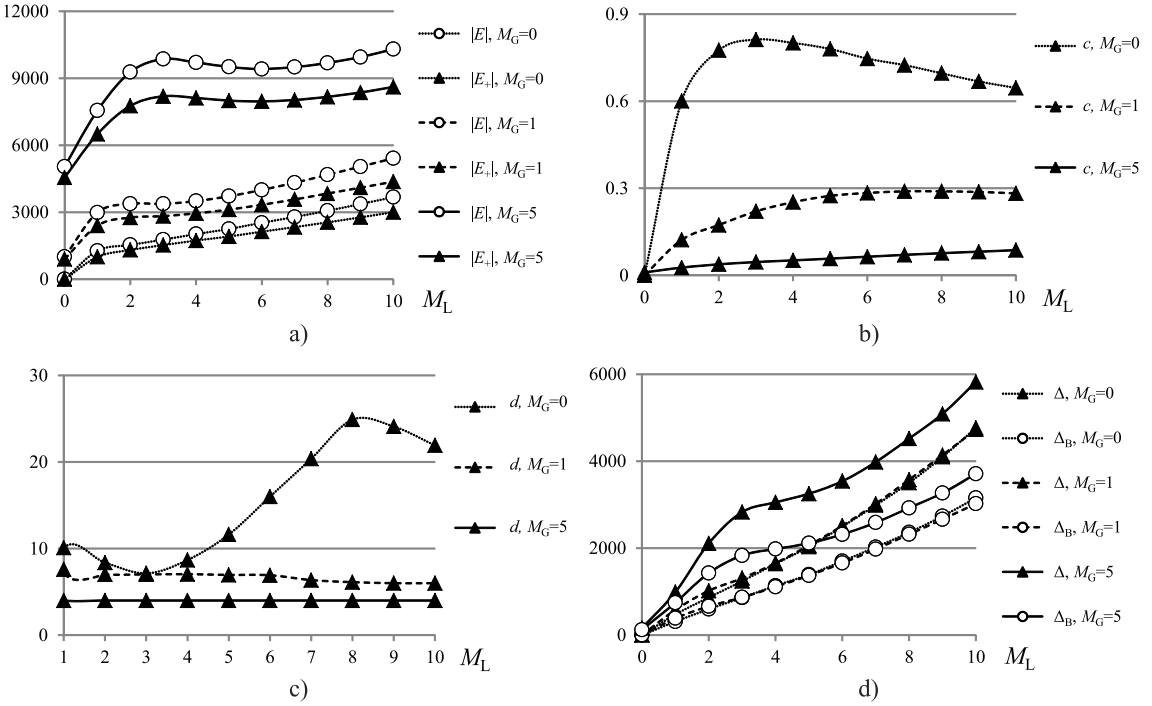


Fig. 3. Influence of the parameter M_{L} on the IB model network formation: number of edges $|E|$ and number of positive edges $|E_+|$ (a), clustering coefficient c (b), diameter d (c), number of triangles Δ and number of balanced triangles Δ_{B} (d). The results are averaged over 50 realizations for $N = 1000$ networks.

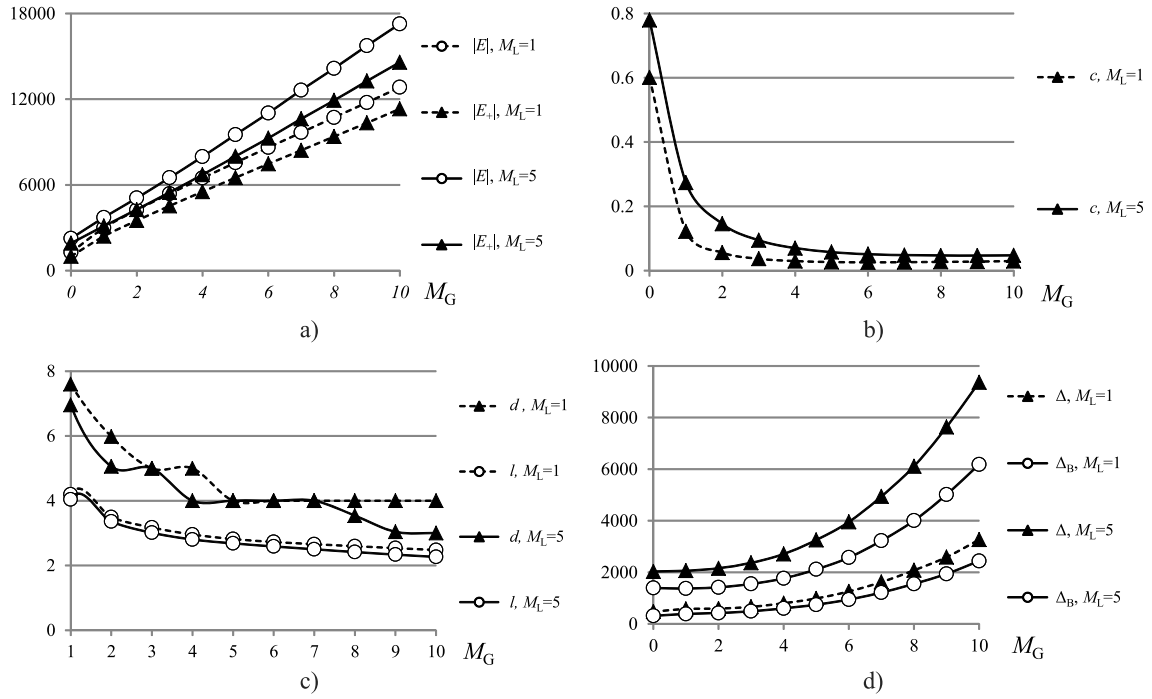


Fig. 4. Influence of the parameter M_{G} on the IB model network formation: number of edges $|E|$ and number of positive edges $|E_+|$ (a), clustering coefficient c (b), diameter d and average path length l (c), number of triangles Δ and number of balanced triangles Δ_{B} (d). The results are averaged over 50 realizations for $N = 1000$ networks.

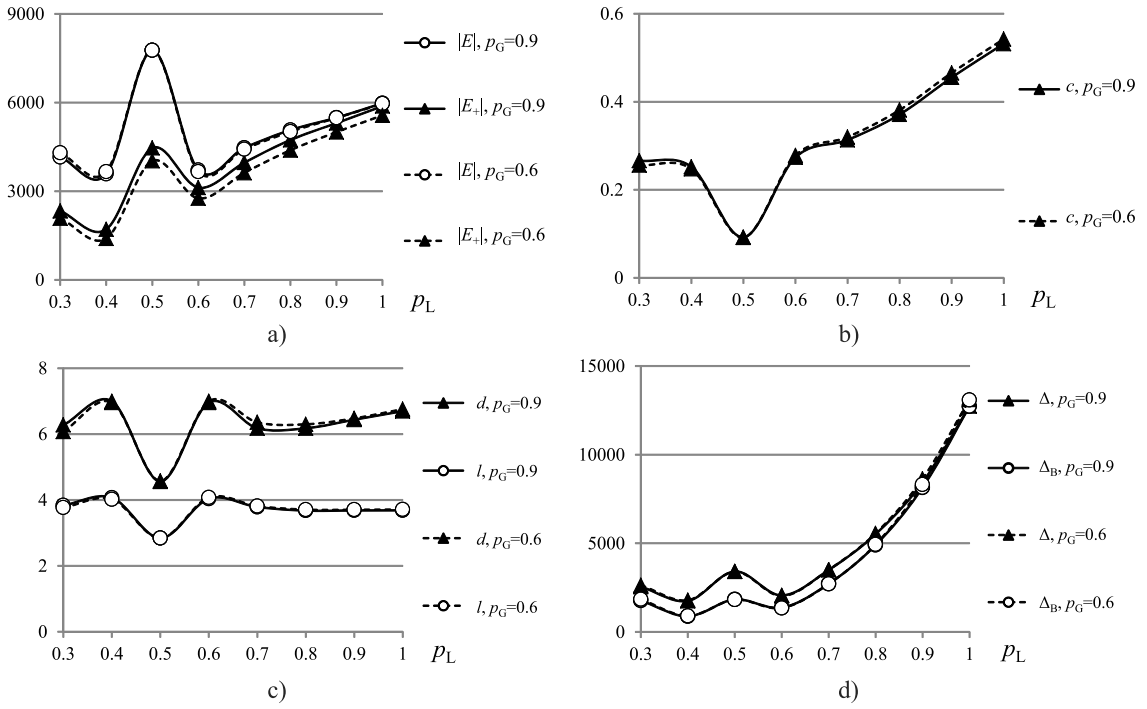


Fig. 5. Influence of the parameter p_L on the IB model network formation: number of edges $|E|$ and number of positive edges $|E_+|$ (a), clustering coefficient c (b), diameter d and average path length l (c), number of triangles Δ and number of balanced triangles Δ_B (d). The results are averaged over 50 realizations for $N = 1000$ networks.

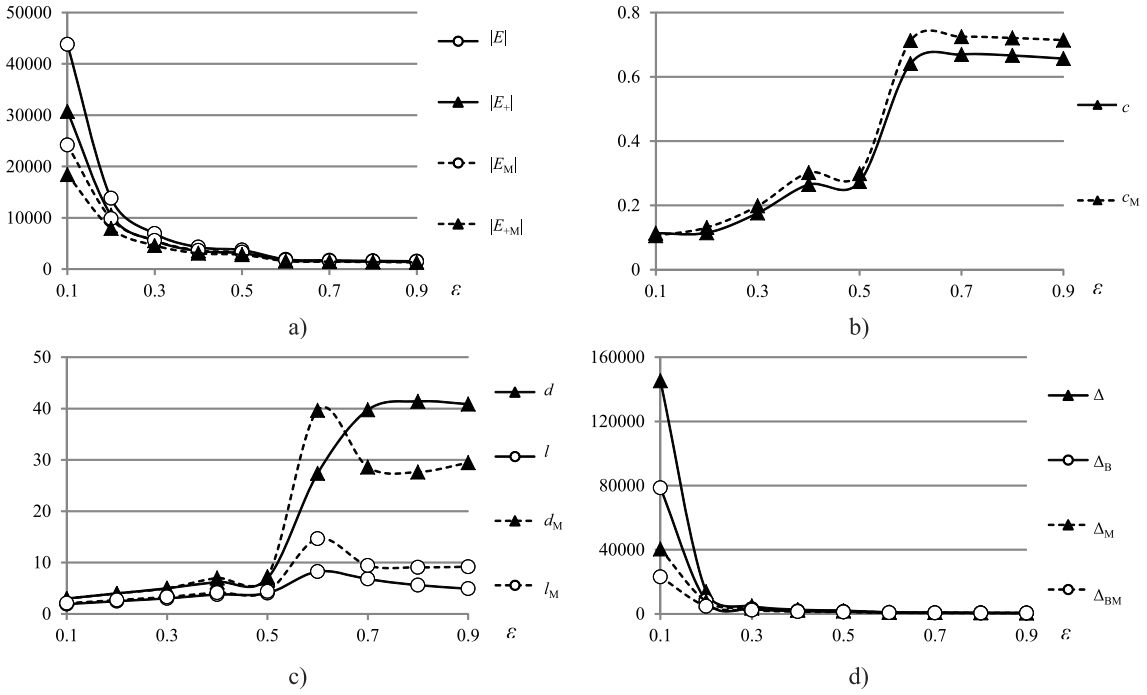


Fig. 6. Influence of the parameter ε on the IB model network formation (at $t = 150$ time steps and in the mature phase, where properties are indexed by M): number of edges $|E|$ and number of positive edges $|E_+|$ (a), clustering coefficient c (b), diameter d and average path length l (c), number of triangles Δ and number of balanced triangles Δ_B (d). The results are averaged over 50 realizations for $N = 1000$ networks.

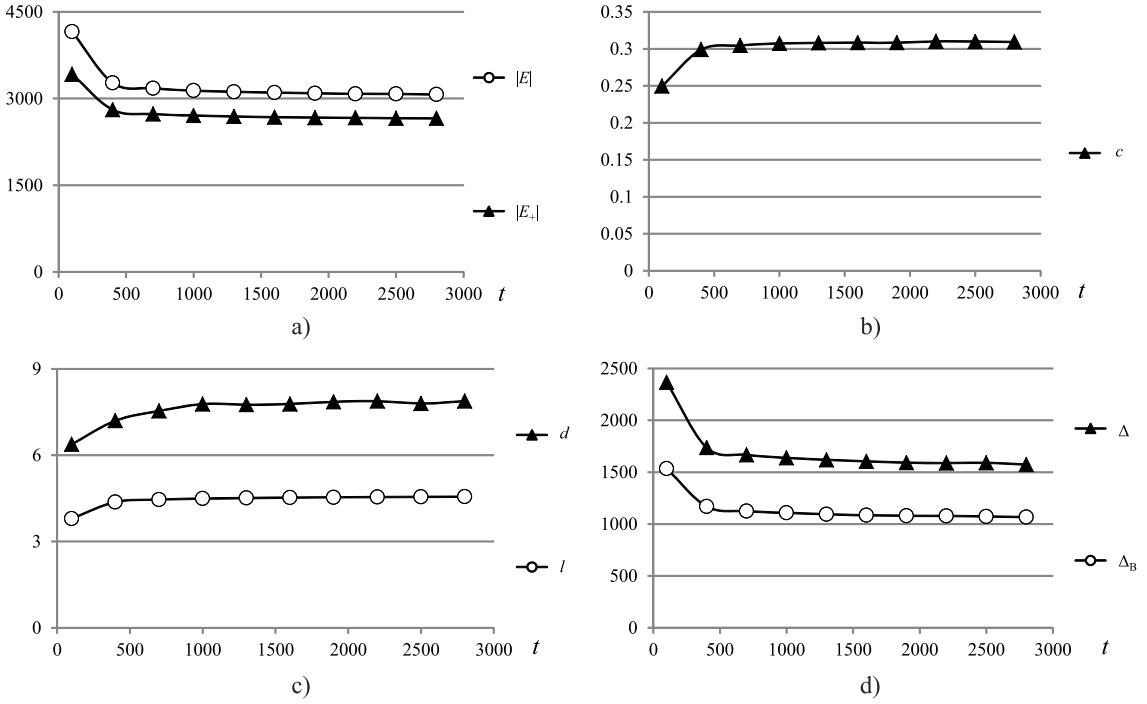


Fig. 7. IB model properties at different time steps t: number of edges $|E|$ and number of positive edges $|E_+|$ (a), clustering coefficient c (b), diameter d and average path length l (c), number of triangles Δ and number of balanced triangles Δ_{B} (d). The results are averaged over 50 realizations for $N = 1000$ networks.

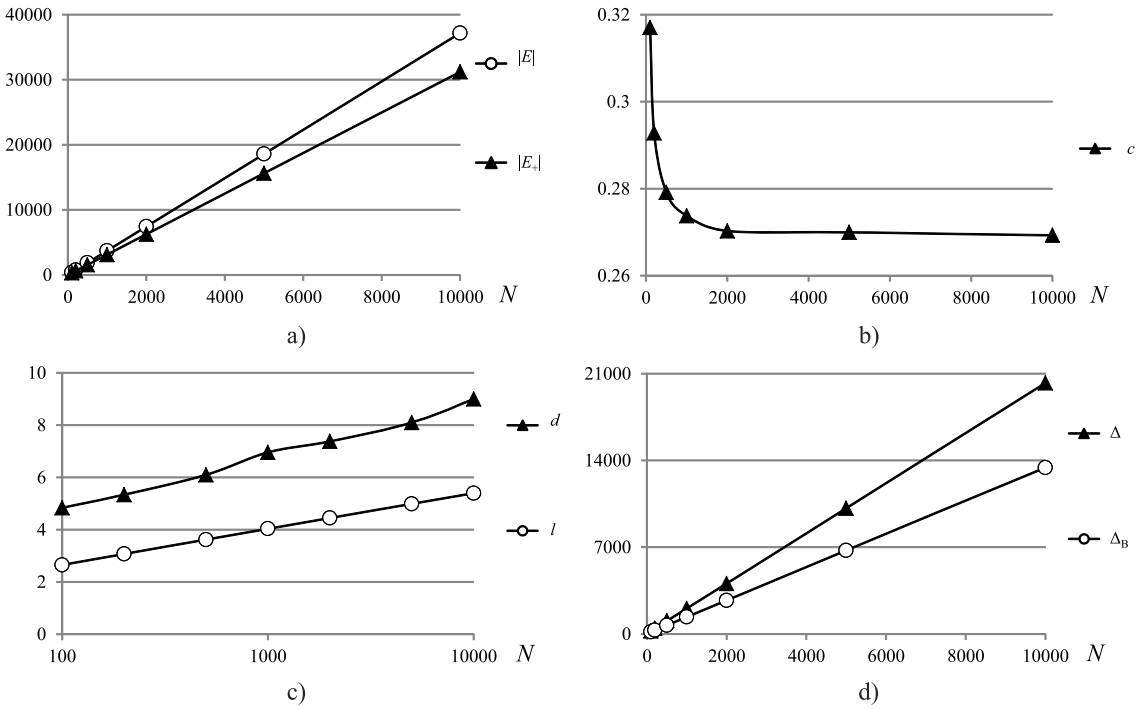


Fig. 8. IB model properties as a function of network order: number of edges $|E|$ and number of positive edges $|E_+|$ (a), clustering coefficient c (b), diameter d and average path length l (c), number of triangles Δ and number of balanced triangles Δ_{B} (d). The results are averaged over 50 realizations for $N = 1000$ networks.

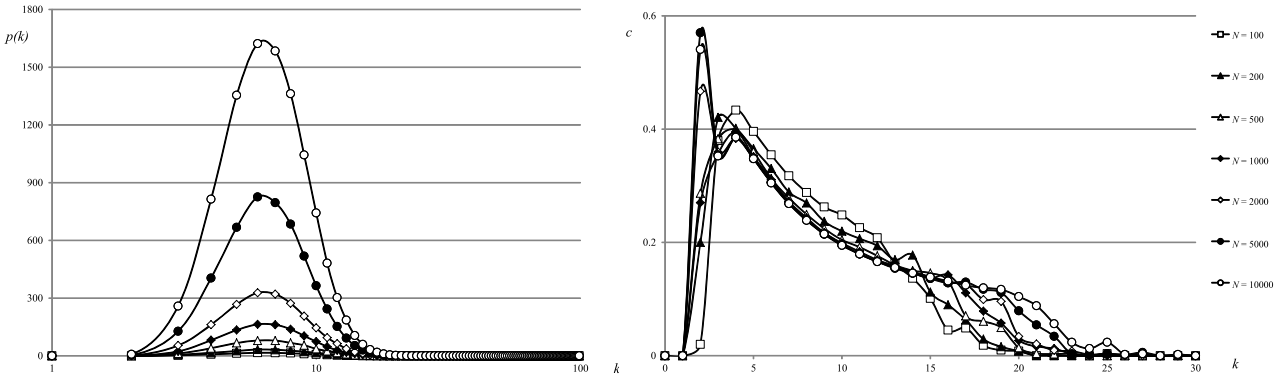


Fig. 9. Degree distributions $p(k)$ of the IB model are skewed and the position of the peak of the distribution does not change as the number of vertices grows. In order to capture the differences between the simulated models corresponding to different graph order, the degree values are plotted on a logarithmic scale (left). The clustering coefficient c of the IB model as a function of the vertex degree k. The results are averaged over 50 realizations for $N = 1000$ networks (right).

have much influence on the network properties, except for the number of positive edges. The network-formation behavior at different p_{L} values shows an interesting hop at $p_{\mathrm{L}} = 0.5$. This is the case when the local interactions are, with equal probabilities, positive or negative. The increase in the number of edges (Fig. 5(a)) and the decrease in the clustering coefficient and the diameter values (Figs. 5(b) and 5(c)) are indicated.

Figure 6 shows the influence of the parameter ε on the network formation at $p_{\mathrm{L}} = 0.6$, $p_{\mathrm{G}} = 0.9$, $M_{\mathrm{L}} = 5$, and $M_{\mathrm{G}} = 1$. The evaluations at $t = 150$ and the evaluations in steady state are shown. The steady state is indicated when the number of edges in 5 repetitions of adding 50 time steps does not change by more than 5%. In the sparse networks, modeled by high ε values, the steady state is established after a smaller number of time steps than with denser networks modeled by small ε values, see Table 1. As expected, ε has a significant influence on the number of edges and the number of triangles in the network (Figs. 6(a) and 6(d)). The numbers of edges and triangles decrease rapidly as ε increases from 0.1 to 0.6. For larger values of ε, the decrease in the number of edges and triangles stabilizes. Notice that the threshold $\varepsilon = 0.5$ is the point where that of the local interactions becomes larger than the impact of the global interactions. If $\varepsilon > 0.5$, the weight of the connections becomes smaller than ε for the edges, which are established and their value is not additionally reinforced in each time step. This means that these edges are not truly established and they indicate only the interactions that are quickly forgotten. As a result, the clustering coefficient and the diameter become significantly larger, see Figs. 6(b) and 6(c).

Table 1 indicates how the networks' properties change with time in the mature phase. Here, the networks' properties in the mature phase and their change in percentage after an additional 400 time steps are shown. As indicated in Table 1, in the mature phase the networks' properties mostly vary beyond 5%. An exception can be found in the behavior of the diameter d and the average path l at ε values greater than 0.5. As $\varepsilon > 0.5$, d and l remain unstable. We calculated the modularity of the communities detected by the Walktrap algorithm (Pons and Latapy, 2007). It is based on random walks and uses the idea that the vertices encountered on any given random walk are more likely to be within a community than not. We compared the community structure of the networks in the mature phase with that before and after an additional 400 time steps with the cc measure introduced by Rand (1971). Networks formed with higher values of ε and therefore a smaller number of edges have a higher modularity, but all over the tested ε, cc values indicate that the community structure does not change much. For comparison, the average modularity of 50 Erdös–Rényi (ER) networks with $N = 1000$ and $|E| = 24000, 3300, 1500$ was $Q = 0.083, 0.297, 0.515$, respectively. We also made 50 community comparisons of ER networks with $N = 1000$ and $|E| = 24000, 3300, 1500$ and obtained $cc = 0.779, 0.903, 0.539$ on average, respectively.

Figure 7 shows the influence of the time steps t on the network formation at $p_{\mathrm{L}} = 0.6$, $p_{\mathrm{G}} = 0.9$, $\varepsilon = 0.5\delta$, $M_{\mathrm{L}} = 5$, and $M_{\mathrm{G}} = 1$. In 433 time steps the model comes into the mature phase, i.e., the state when the number of edges does not change significantly, see Fig. 7(a). Figure 7 shows that at that time the internal dynamics of the network formation also stabilize, i.e., the number of positive edges, the clustering coefficient, the diameter, the average path length, the number of triangles, and the number of balanced triangles do not change significantly.

In addition, we tested the model on $N \in \{100, 200, 500, 1000, 2000, 5000, 10000\}$ vertices at $t = 150$. Figure 8 shows how the network properties change

with the number of vertices. The parameters were set to $p_{\mathrm{L}} = 0.6$, $p_{\mathrm{G}} = 0.9$, $\varepsilon = 0.5\delta$, $t = 150$, $M_{\mathrm{L}} = 5$, and $M_{\mathrm{G}} = 1$. The number of edges, the number of positive edges, the number of triangles, and the number of balanced triangles increase linearly with the number of vertices, see Figs. 8(a) and 8(d). The clustering coefficient for the network order $N = 2000$ stabilizes at the value $c = 0.27$, see Fig. 8(b). The average path length (Fig. 8(c)) increases in proportion to the logarithm of the number of vertices N, which is again the small-world property (Watts, 1999). The degree distributions $p(k)$ (Fig. 9(a)) are slightly right-skewed, and the position of the peak of the degree distribution does not change as the number of vertices increases. The clustering coefficient c as a function of the vertex degree k decreases gradually for all the tested network orders, see Fig. 9(b).

Table 2 shows the numerical values of the tested models at $p_{\mathrm{L}} = 0.6$, $p_{\mathrm{G}} = 0.9$ $\varepsilon = 0.5\delta$, $t = 150$, $M_{\mathrm{L}} = 5$, and $M_{\mathrm{G}} = 1$ and the different network order.

The IB model by itself does not have the property of dynamically adjusting the organization of communities depending on the external conditions, but it is interesting to know how capable the IB modeling mechanism is at reconstructing the primary network structure after some edge malfunctions happen.

We performed two tests on the IB model's behavior in extreme situations. In the fist one, we simulated a large number (25000) of negative interactions between 50% of all the vertices in the networks in the mature phase. We measured the model properties, when networks came into the mature phase, after the simulation of negative interactions, and after the re-stabilization, at different ε values. The simulation can reflect the model's response in riots, i.e., when in a small time period a lot of negative interactions happen in crowds. The results in Table 3 indicate a large increase in the number of edges immediately after the simulation of negative interactions and its subsequent re-stabilization. Similarly, the average path length l, the clustering coefficient c, and the modularity Q changed, but afterwards their values re-stabilized with the trend of the model's behavior in the mature phase, see Table 1. The high cc values indicate that the community structure before and after the simulation of negative interactions remains very similar.

With the second test, we simulated the deletion of edges between 50% of all the vertices in the networks in the mature phase, see Table 4. Like before, we measured the model properties in the networks' mature phase, after the simulation, and after the re-stabilization, at different ε values. This simulation can reflect the model response after natural disasters, when many connections are roughly broken. After the re-stabilization the networks have properties similar to those before the deletion. The deletion causes the networks temporarily to have a smaller clustering coefficient, but higher modularity values. This indicates that an individual's neighbors are less connected, but the fraction of connections inside communities with respect to the fraction of connections if they are distributed at random is now larger. This is an important observation that is in contrast to the effect of adding negative interactions, where both the clustering coefficient and the modularity values decrease, see Table 3.

The results from both tests show that the IB model has high self-repairing abilities when some edge malfunctions happen.

Furthermore, we compared the different topological properties of our interaction-based networks, Barabási–Albert (BA) networks, Erdös–Rényi networks and a real e-mail URV network (Guimerà *et al.*, 2003). The e-mail URV is a network containing 1133 members of the University Rovira i Virgili of Tarragona in Spain, and its edges are e-mail interchanges between the users. We observed the average results of 50 runs of the BA model, the ER model and the IB model with 1133 vertices. The other parameters of the IB model were set to $M_{\mathrm{L}} = 5$, $M_{\mathrm{G}} = 1$, $\varepsilon = 0.4\delta$, $p_{\mathrm{L}} = 0.6$, $p_{\mathrm{G}} = 0.9$, and $t = 150$. We observed the BA model that starts with a complete graph with five vertices and at each time step a new vertex with degree five is connected to the existing vertices with a probability that it is proportional to their number of edges. In the ER model the expected number of edges was set to $4.8N$.

Table 5 shows the number of edges $|E|$, the diameter d, the average path length l, the clustering coefficient c, the number of triangles Δ of the tested models, and Q, which is the modularity value of the network divided into communities achieved using the Walktrap algorithm (Pons and Latapy, 2007). The IB networks, as well as the e-mail URV network clustering coefficient, are relatively high, while the BA networks and the ER networks are not clustered. Similarly, the number of triangles in the IB networks and the e-mail URV network are much higher than in the BA networks and the ER networks. Although that of the IB model is higher than the clustering coefficient of the e-mail URV network, it has a smaller number of triangles, which is expected as the IB networks are sparser. The modularity values of the IB networks fit the modularity value of the URV network, while the modularity values of the BA and the ER networks are much lower. The values of the average path length in all the tested networks are close to each other.

According to the classification of social network models by Toivonen *et al.* (2009), the IB model falls into the category of dynamic-network-evolution models. The representative models of this category are described by Davidsen *et al.* (2002), Marsili *et al.* (2004) and Kumpula *et al.* (2007). For the basic statistics of these models fitted to the e-mail URV network, see the work of Toivonen *et al.* (2009).

Table 1. Main properties of the interaction-based model in the mature phase and their differences after an additional 400 time steps in percentages.

ε	t	$\lvert E\rvert$		$\lvert E_+\rvert$		d		l		c		Q		cc
0.1	623	24163.6	-8.9	18410.3	-6.8	3.00	0	2.08	2.6	0.106	-1.4	0.154	7.0	0.827
0.2	512	9822.4	-6.3	7926.4	-4.9	4.00	0	2.70	1.5	0.131	0.4	0.301	3.7	0.892
0.3	471	5557.1	-4.8	4606.6	-3.8	5.00	0.8	3.31	1.9	0.198	1.6	0.449	2.8	0.934
0.4	457	3651.1	-3.9	3089.4	-3.2	7.00	1.4	4.14	2.5	0.302	2.6	0.560	2.6	0.957
0.5	433	3272.7	-3.2	2804.9	-2.8	7.28	5.8	4.37	2.2	0.299	2.5	0.618	2.1	0.960
0.6	394	1755.7	-2.1	1575.9	-2.2	39.66	-30.1	14.65	-37.2	0.713	3.2	0.957	1.3	0.985
0.7	373	1610.2	-2.0	1458.4	-2.1	28.60	-40.6	9.44	-51.8	0.725	3.3	0.971	1.2	0.988
0.8	378	1523.0	-1.9	1383.7	-2.0	27.64	-37.1	9.08	-50.5	0.721	3.1	0.974	1.0	0.987
0.9	379	1455.9	-2.1	1325.4	-2.2	29.44	-34.0	9.18	-39.5	0.714	2.4	0.974	0.8	0.985

Table 2. Main properties of the IB model for different network orders.

N	$\lvert E\rvert$	d	l	c	Δ
100	374.8 ± 14.6	4.84 ± 0.37	2.65 ± 0.05	0.317 ± 0.017	250.3 ± 23.1
200	747.3 ± 23.5	5.34 ± 0.47	3.08 ± 0.05	0.293 ± 0.011	454.7 ± 33.3
500	1869.5 ± 33.6	6.10 ± 0.30	3.62 ± 0.04	0.279 ± 0.008	1067.2 ± 48.6
1000	3718.5 ± 45.5	6.96 ± 0.20	4.04 ± 0.03	0.274 ± 0.005	2063.2 ± 63.8
2000	7428.6 ± 77.0	7.38 ± 0.49	4.45 ± 0.03	0.270 ± 0.003	4064.8 ± 95.2
5000	18586.0 ± 87.8	8.10 ± 0.30	4.99 ± 0.01	0.270 ± 0.002	10144.0 ± 130.9
10000	37171.0 ± 134.3	9.00 ± 0.00	5.40 ± 0.01	0.269 ± 0.002	20252.0 ± 177.7

6. Conclusion

In this paper we introduced a novel social-network model called the interaction-based model, which involves well-known sociological principles. Heider's theory (Heider, 1946) states that two people equivalently perceive their relation as friendly or unfriendly, while triadic relations are balanced or imbalanced, depending on the number of friendly and unfriendly dyadic relations inside them. Each relation between two people is built on the everyday interactions between them, either positive or negative. The model is also inspired by the social behavior of animal species, particularly that of ants in their colony. Ants put pheromone on their path while walking from a food source to the nest. Because the pheromone is evaporable, it slowly disappears. Similarly, the ties between people become weaker if no interactions exist between them.

The model evaluation showed that the IB model has characteristic social-network properties. The IB networks turned out to be sparse, they have a small diameter and an average path length that grows in proportion to the logarithm of the number of vertices. The clustering coefficient is relatively high, and its value stabilizes in larger networks. The degree distributions are slightly right-skewed. In the mature phase of the IB model, i.e., when the number of edges does not change significantly, most of the network properties do not change significantly either.

A further investigation of the model parameters shows how inner changes to the social system cause a change in the network topology. For example, higher probability of positive interactions happening in the network resulted in more clustered networks (higher clustering coefficient), as well as an increasing number of balanced triadic relations. While stronger evaporation also caused higher clustering coefficients, the number of balanced triadic relations decreased.

Two tests where an edge malfunction is simulated showed that networks built as proposed in this paper have high self-repairing abilities. In both cases, i.e., during the deletion of edges and during the unexpected increase of negative interactions between individuals, the LocalInteractions() function in cooperation with the GlobalInteractions() function causes re-stabilization of the network. Good self-repairing abilities are indicated with similar network properties before and after the malfunctions happened and which variations are close to the variation trend explained by the model measured in the mature phase.

The IB model was found to be the best of all the compared models in simulating the e-mail URV network because its properties more closely matched those of the e-mail URV network than the other models.

In the future we plan to analyze additional properties of the IB model (largest component size, average degree, assortativity coefficient) and tune its parameters to various

Table 3. Main properties of the IB model measured in the mature phase, after the simulation of negative interactions (25000) between 50% of all vertices, and after coming again into the stable state. The community structures are compared before and after the simulation and their comparisons are indicated with cc values.

ε	$\lvert E\rvert$			l			c			Q			cc
0.1	24163.56	45623.82	22479.74	2.08	1.94	2.12	0.11	0.16	0.10	0.15	0.12	0.16	0.828
0.2	9822.40	31960.26	9317.88	2.70	2.18	2.73	0.13	0.17	0.13	0.30	0.10	0.31	0.895
0.3	5557.12	27943.32	5336.04	3.30	2.34	3.36	0.20	0.21	0.20	0.45	0.06	0.46	0.931
0.4	3651.12	26130.56	3534.92	4.13	2.48	4.22	0.30	0.26	0.31	0.60	0.04	0.61	0.958
0.5	3272.74	25775.54	3185.40	4.37	2.51	4.45	0.30	0.26	0.30	0.62	0.04	0.63	0.961
0.6	1755.68	24323.56	1727.60	14.65	2.75	10.82	0.71	0.47	0.73	0.96	0.03	0.97	0.985
0.7	1610.24	24187.34	1585.18	9.44	2.77	5.86	0.72	0.47	0.74	0.97	0.02	0.98	0.987
0.8	1523.00	24105.38	1494.84	9.08	2.79	5.62	0.72	0.47	0.73	0.97	0.02	0.98	0.986
0.9	1455.94	24039.40	1429.96	9.18	2.79	6.06	0.71	0.46	0.73	0.97	0.02	0.98	0.984

Table 4. Main properties of the IB model measured in the mature phase, after deletion of edges between 50% of all vertices, and after coming again into stable state. The community structures are compared before and after the simulation and their comparisons are indicated with cc values.

ε	$\lvert E\rvert$			l			c			Q			cc
0.1	24235.02	5829.58	20045.96	2.08	2.31	2.20	0.11	0.05	0.10	0.15	0.20	0.18	0.825
0.2	9818.36	2327.76	8427.86	2.70	3.12	2.80	0.13	0.06	0.13	0.30	0.37	0.34	0.885
0.3	5548.76	1304.60	4920.60	3.31	4.26	3.45	0.20	0.10	0.20	0.45	0.53	0.48	0.924
0.4	3647.98	853.30	3302.10	4.14	6.27	4.36	0.30	0.13	0.31	0.60	0.70	0.62	0.945
0.5	3257.14	756.64	3027.30	4.38	7.08	4.56	0.30	0.12	0.30	0.62	0.73	0.63	0.948
0.6	1746.70	391.84	1628.80	13.36	2.01	10.17	0.72	0.16	0.74	0.96	0.98	0.97	0.970
0.7	1603.42	355.10	1512.16	6.92	1.76	5.74	0.73	0.14	0.75	0.98	0.98	0.98	0.976
0.8	1512.76	343.74	1446.60	6.38	1.75	7.26	0.73	0.14	0.74	0.98	0.98	0.98	0.977
0.9	1444.44	330.52	1387.78	7.36	1.72	7.01	0.72	0.13	0.73	0.98	0.98	0.98	0.977

real-world networks. Additionally, we want to compare the self-repairing mechanism studied in our model with those inside real-world networks. The study of similarities and differences in self-repairing behavior can provide us with new knowledge of network-formation mechanisms. Similarities can explain which part of the mechanism was appropriately designed and the differences can explain which part should be further investigated.

Acknowledgment

This paper is partially supported by the Slovenian Research Agency research programs P1-0383 and P2-0098.

References

Antal, T., Krapivsky, P.L. and Redner, S. (2005). Dynamics of social balance on networks, *Physical Review E* **72**(3): 036121.

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks, *Science* **286**(5439): 509–512.

Camazine, S., Deneubourg, J.-L., Franks, N.R., Sneyd, J., Théraulaz, G. and Bonabeau, E. (2001). *Self-Organization in Biological Systems*, Princeton University Press, Princeton, NJ.

Davidsen, J., Ebel, H. and Bornholdt, S. (2002). Emergence of a small world from local interactions: Modeling acquaintance networks, *Physical Review Letters* **88**(12): 128701.

de Sola Pool, I. and Kochen, M. (1978). Contacts and influence, *Social Networks* **1**(1): 5–58.

Dorigo, M., Bonabeau, E. and Théraulaz, G. (2000). Ant algorithms and stigmergy, *Future Generation Computer Systems* **16**(9): 851–871.

Dorigo, M., Maniezzo, V. and Colorni, A. (1996). The ant system: Optimization by a colony of cooperating agents, *IEEE Transactions on Systems, Man, and Cybernetics B* **26**(1): 29–41.

Erdös, P. and Rényi, A. (1960). On the evolution of random graphs, *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* **5**(1–2): 17–61.

Freeman, L.C. (2004). *The Development of Social Network Analysis: A Study in the Sociology of Science*, Empirical Press, Vancouver, BC.

Granovetter, M.S. (1973). The strength of weak ties, *American Journal of Sociology* **78**(6): 1360–1380.

Table 5. Main properties of the IB model with parameters $t = 150$, $M_{\mathrm{L}} = 5$, $M_{\mathrm{G}} = 1$, $\varepsilon = 0.4\delta$, $p_{\mathrm{L}} = 0.6$, and $p_{\mathrm{G}} = 0.9$, the real e-mail network, the BA model and the ER model.

network	N	$\lvert E\rvert$	d	l	c	Δ	Q
email	1133	5451	8	3.61	0.220	5345	0.531
BA	1133	5650	5.0 ± 0.0	3.09 ± 0.01	0.028 ± 0.002	916.6 ± 50.0	0.180 ± 0.004
ER	1133	5450.7 ± 79.6	5.84 ± 0.37	3.36 ± 0.01	0.008 ± 0.001	147.4 ± 13.0	0.234 ± 0.005
IB	1133	4856.3 ± 71.0	6.36 ± 0.48	3.84 ± 0.03	0.265 ± 0.004	3059.4 ± 98.4	0.537 ± 0.009

Grassé, P.-P. (1959). La reconstruction du nid et les coordinations inter-individuelles chez *Bellicositermes natalensis* et *Cubitermes* sp. La théorie de la stigmergie: Essai d'interprétation du comportement des termites constructeurs, *Insectes Sociaux* **6**(1): 41–81.

Grossetti, M. (2005). Where do social relations come from? A study of personal networks in the Toulouse area of France, *Social Networks* **27**(4): 289–301.

Guimerà, R., Danon, L., Diaz-Guilera, A., Giralt, F. and Arenas, A. (2003). Self-similar community structure in a network of human interactions, *Physical Review E* **68**(6): 065103(R).

Handl, J., Knowles, J. and Dorigo, M. (2006). Ant based clustering and topographic mapping, *Artificial Life* **12**(1): 35–61.

Heider, F. (1946). Attitudes and cognitive organization, *Journal of Psychology* **21**(1): 107–112.

Jiang, J., Wang, R. and Wang, Q.A. (2011). Network model of deviation from power-law distribution in complex network, *The European Physical Journal B* **79**(1): 29–33.

Korošec P. (2006). *Stigmergy as an Approach to Metaheuristic Optimization*, Ph.D. thesis, Jožef Stefan International Postgraduate School, Ljubljana.

Korošec, P., Šilc, J. and Filipič, B. (2012). The differential ant-stigmergy algorithm, *Information Sciences* **192**(1): 82–97.

Kumpula, J.M., Onnela, J., Saramäki, J., Kaski, K. and Kertesz, J. (2007). Emergence of communities in weighted networks, *Physical Review Letters* **99**(22): 228701.

Kleinberg, J.M. (2000). The small-world phenomenon: An algorithmic perspective, *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing, Portland, OR, USA*, pp. 163–173.

Leskovec, J. (2010). Kronecker graphs: An approach to modeling networks, *Journal of Machine Learning Research* **11**: 985–1042.

Ludwig, M. and Abell, P. (2007). An evolutionary model of social networks, *European Physical Journal B* **58**(1): 97–105.

Marsili, M., Vega-Redondo, F. and Slanina, F. (2004). The rise and fall of a networked society: A formal model, *Proceedings of the National Academy of Sciences of the United States of America* **101**(6): 1439–1442.

Martens, D., De Backer, M., Haesen, R., Vanthienen, J., Snoeck, M. and Baesens, B. (2007). Classification with ant colony optimization, *IEEE Transactions on Evolutionary Computation* **11**(5): 651–665.

Marvel, S.A., Kleinberg, J., Kleinberg, R.D., and Strogatz, S.H. (2011). Continuous-time model of structural balance, *Proceedings of the National Academy of Sciences of the United States of America* **108**(5):1771–1776.

Milgram, S. (1967). The small world problem, *Psychology Today* **1**(1): 61–67.

Newman, M.E.J. (2001). The structure of scientific collaboration networks, *Proceedings of the National Academy of Sciences of the United States of America* **98**(2): 404–409.

Newman, M.E.J. (2008). *Mathematics of Networks*, Palgrave Macmillan, Basingstoke.

Péter, T. (2012). Modeling nonlinear road traffic networks for junction control, *International Journal of Applied Mathematics and Computer Science* **22**(3): 723–732, DOI: 10.2478/v10006-012-0054-1.

Pons, P. and Latapy, M. (2005). Computing communities in large networks using random walks, *arXiv:physics/0512106v1*.

Rand, W.M. (1971). Objective criteria for the evaluation of clustering methods, *Journal of the American Statistical Association* **66**(336): 846–850.

Rapoport, A. (1957). Contribution to the theory of random and biased nets, *Bulletin of Mathematical Biology* **19**(4): 257–277.

Srinivasan, A. (2011). Local balancing influences global structure in social networks, *Proceedings of the National Academy of Sciences of the United States of America* **108**(5): 1751–1752.

Toivonen, R., Kovanen, L., Kivelä, M., Onnela, J., Saramäki, J. and Kaski, K. (2009). A comparative study of social network models: Network evolution models and nodal attribute models, *Social Networks* **31**(4): 240–254.

Vukašinović, V., Šilc, J. and Škrekovski, R. (2012a). Towards social networks model, *Proceedings of the 5th International Conference on Bioinspired Optimization Methods and Their Applications, Bohinj, Slovenia*, pp. 49–60.

Vukašinović, V., Šilc, J. and Škrekovski, R. (2012b). Swarm-inspired social network model and its properties, *Proceedings of the 4th International Conference on Information Technologies and Information Society, Novo Mesto, Slovenia*.

Wang, W., Hu, B., Zhou, T., Wang, B. and Xie, Y. (2005). Mutual selection model for weighted networks, *Physical Review E* **72**(4): 046140.

Watts, D.J. (1999). *Small Worlds: The Dynamics of Networks Between Order and Randomness*, Princeton University Press, Princeton, NJ.

Watts, D.J. and Strogatz, S.H. (1998). Collective dynamics of 'small-world' networks, *Nature* **393**(6684): 440–442.

White, D.R., Kejzar, N., Tsallis, C., Farmer, D. and White, S. (2006). A generative model for feedback networks, *Physical Review E* **73**(1): 016119.

Wynne, C.D.L. (2001). *Animal Cognition: The Mental Lives of Animals*, Palgrave Macmillan, Basingstoke.

Xiong, F., Liu, Y., Zhu, J., Zhang, Z.J., Zhang, Y.C. and Zhang, Y. (2011). A dissipative network model with neighboring activation, *The European Physical Journal B* **84**(1): 115–120.

Vida Vukašinović graduated from the Faculty of Mathematics and Physics at the University of Ljubljana and received her Ph.D. from the Jožef Stefan International Postgraduate School, Ljubljana, Slovenia, in 2013. Her research interests include graph theory, optimization techniques and meta-heuristic algorithms.

Jurij Šilc received his Ph.D. in electrical engineering from the University of Ljubljana, Slovenia, in 1992. He is now a senior researcher at the Jožef Stefan Institute, Ljubljana, and an assistant professor at the Jožef Stefan International Postgraduate School, Ljubljana. His research interests include parallel processing, computer architectures, and bio-inspired optimization algorithms.

Riste Škrekovski is an associate professor at the Department of Mathematics of the University of Ljubljana, Slovenia, where he also obtained his Ph.D. degree in graph theory. His research interests are various classical topics in graph theory as well as graph applications in chemistry and networks.

Received: 21 June 2013
Revised: 27 November 2013
Re-revised: 29 January 2014