STIMULATION MODEL – A MULTILAYER SOCIAL NETWORK GENERATOR

Alexander Tarvid

University of Latvia

atarvid@inbox.lv

ABSTRACT
The paper presents the first two stages of building the Stimulation model. The model is aimed at generating a multilayer social network with four layers of size identified in a recent review of empirical studies. The model introduces narrow environments, where all members know everyone, thus, forming cliques in the weighted directed graph representing the social network. The weight of an arc connecting agent v to agent w depends on w's position in the unit square of v and w's layer in v's social network. The unit square is defined by the relative frequency of interactions of v with w and closeness of both agents in quality. At the second stage, stable distribution of layer size and relative time spent interacting with layers is achieved.

Keywords: social network, network generator, multi-layer network

1. INTRODUCTION

A review of empirical studies on the composition of social networks (Zhou et al 2005) identified a stable hierarchical multilayer structure. Four layers were found: support clique (3–5 vertices), sympathy level (9–15 vertices), band level (30–50 vertices) and community level (around 150 vertices). The frequency of contact decreases with hierarchy level, with the support clique being a source of personal advice or help in severe circumstances. To the best of my knowledge, no paper has attempted to re-create this empirical fact yet, and this paper is the first attempt to do it.

Formally, a *multilayer network* is a family of graphs, one graph per layer, with intra-layer edges connecting vertices of the same graph and inter-layer edges connecting vertices in different graphs (hence, layers) (Boccaletti et al 2014). A type of multilayer network is a *multiplex network*, where all graphs from the family contain the same set of vertices (Boccaletti et al 2014). Multilayer networks can, thus, represent both different groups of vertices (the membership of a *vertex* in a particular layer shows its membership in a particular group) and different types of connections between the same two vertices (a membership of an *edge* in a particular layer shows that the connection belongs to a particular type).

The model presented here generates a special type of a multiplex network, where, assuming that vertex v is considered the same vertex irrespective of the layer it appears in, any two vertices may be connected by not more than one arc with same direction (although it can also be viewed as a multilayer network with different vertices on each layer forming cliques). In other words, arc (v, w) can appear in not more than one layer of the network. Referring again to Zhou et al (2016), however, it is easy to see that a layer where a particular vertex lies is not its general position, but it is its subjective position from the point of view of another given vertex. For instance, person c may be very important for person a but not for person b. To complicate it further, person a may actually not be that important for person c. In other words, none of the two representations of a multilayer network mentioned above are applicable in this case.

The literature on multilayer/-level networks has been actively emerging in the last years. For instance, Cantor et al (2015) and Senior et al (2016) developed agent-based models of evolving animal multilayer networks. A multi-layer model of risk in financial markets is studied by Poledna et al (2015). Multi-layer social networks were used in modelling organized crime (Li et al 2015), diffusive processes (e.g., of innovations) (Li, Yan and Jiang 2015; Ramezanian et al 2015) and collaborative learning behaviour in organisations (Różewski et al 2015).

The rest of the paper is structured as follows. The next section introduces some conceptual foundations. Then Sects. 3 and 4 present the first and second stages of the model, respectively. The first stage introduces the main mechanisms but is limited to allow only growth of higher layers in terms of size. At the second stage, a balancing mechanism is introduced to limit the size of each layer around the empirically observed values. The last section concludes.

2. CONCEPTUAL FOUNDATIONS

Assume a population of agents, some of which are connected. Connections are directed, so the network is represented by a digraph. An arc's weight is one of the factors that determines the layer in which the arc exists in an agent's digraph. The higher the weight, the stronger the connection, the higher ('more intimate') the layer.

For the connections to be able to move across layers over time, the mechanism of *stimulation* is introduced, which is also the reason for which the model is called the *Stimulation model*. The stimulation of an arc occurs when its incident vertices (more concretely, the agents represented by these vertices in the social graph) interact. The more frequent the interactions become, the more the arc's weight increases.

This mechanism resembles that of reinforcement learning in that the agent does not decide purely randomly with whom to interact more, but bases this decision on the result of the history of interactions with different agents. More frequent interactions occur with those agents, the interactions with which bring most benefit to the agent.

The ultimate challenge of the model is to simultaneously create a layered structure of a network for every agent and keep the overall structure of the network scale-free (i.e., the distribution of the degrees of vertices should be governed by power law). The exponent of the power law should ideally be one of the parameters of the model, and it should be in the range typical for social networks. In this paper, however, the network is fixed and, by construction, does not have a scale-free structure. That structure should be attained at later stages, when more dynamics are introduced into the model.

The model is built around the following theoretical concepts. Firstly, *homophily* (McPherson, Smith-Lovin and Cook 2001) argues that personal networks are quite homogeneous on a large number of parameters characterising individuals. In other words, individuals tend to connect to similar individuals more frequently than to dissimilar individuals. Secondly, *structural constraint* (Fischer 1982) means that the pool of potential members of a social network is to a large extent determined by the social contexts where the individual participates. These social contexts will appear as narrow environments, where everyone knows everyone.

Besides Zhou et al (2016), the paper was also influenced by the empirical findings of Grossetti (2005) on the initial meeting contexts of network members and their current characterisations. It was shown that the majority of connections was initially met through family or existing friends, but also in other environments, such as at work or in organisations, during education and in the neighbourhood, while only 6% was met by chance. Most of these contexts will be abstracted in the model as narrow environments. It was also shown that the current characterisations typically differed from the initial context, allowing to group contacts into family, friends, work/organisations, neighbours and acquaintances. This (together with the results of Zhou et al (2016)) will be abstracted as moving up and down through the layers of the network.

3. STAGE 1: A GROWTH-ONLY MODEL

A fixed number of N agents, which will be called persons, reside in a system with M narrow environments. Each person is characterized by quality and a list of narrow environments it is a member of.

At this stage, the number of narrow environments is fixed and the membership of a person in narrow environments is immutable. Upon entering the model, a person is assigned to $\mathrm{Pois}(\lambda_e) + 1$ narrow environments.

The quality of person i is a fixed number $q_i \in [0, 1]$ generated from normal distribution $\mathcal{N}(\mu, \sigma)$. It will determine the overall similarity between two given persons.

The social network is represented by a weighted digraph $G = (V, A, w)$, where V is the set of vertices, A is the set of arcs and $w: A \to \mathbb{R}$ is a function mapping arcs to their weights.

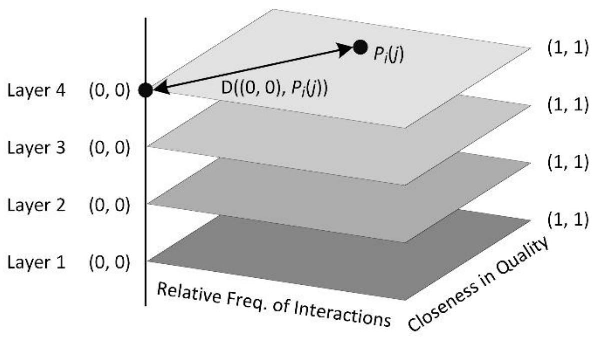

Figure 1: A Stack of Unit Squares Illustrating the Three Components of Arc's (i, j) Weight: Relative Frequency of Interactions, Closeness in Quality and Layer

Note: $P_i(j)$ is the position of j in i's unit square and $D(\cdot,\cdot)$ is the distance between two points.

A *narrow environment* represents a narrow community where everyone knows everyone, which is a typical situation in a small business, a department of a larger business or a group of students in a university. This is also why the environment is called *narrow* – it cannot represent the whole university or a big multinational corporation, where the assumption that everyone knows everyone does not hold. Thus, if a person becomes a member of a narrow environment, it connects with *all other* members of that narrow environment. At this stage, persons cannot leave a narrow environment.

Let persons i and j be connected. Denote the arc connecting i to j as (i, j) and its weight by w_{ij}. The arc's weight represents the importance of the vertex's neighbour to the vertex – in this case, the importance of j to i. It depends on three factors. The first is the closeness of both persons in their quality, $1 - |q_i - q_j|$, which is fixed over time, as quality is immutable.

The second factor is the relative frequency of interactions of person i with person j in the last period of time, I_{ij}/I_i, where $\sum_{j:(i,j)\in A} I_{ij} = I_i$. Thus, the closer the quality and the more time spent interacting, the heavier should be the connection. Because the closeness in quality and relative frequency of interactions are both in $[0, 1]$, the position of j for i on these two factors can be illustrated by a position in the unit square, see Fig. 1: the closer the point to the top right corner $(1, 1)$, the heavier the arc should be.

ISBN 978-88-97999-85-0; Affenzeller, Bruzzone, Jiménez, Longo and Piera Eds.

The third factor is the layer where j resides in i's network. From empirical analysis, it is known that there are four layers. Hence, this is also the number of layers in the model. As shown in Fig. 1, each layer is associated with its own unit square, so that e.g. a point (x, y) in layer 1 will have a substantially lower weight than the point with the same coordinates in layer 4. The weight of the arc (i, j) is given by the following expression:

$w_{ij} = 10^{L_i(j)-1} + D\big((0,0), P_i(j)\big)$,

where $L_i(j)$ is layer where j is located in i's network, numbered from 1 to 4, $D(\cdot,\cdot)$ is the distance between two points and $P_i(j)$ is the point representing j in i's unit square, see Fig. 1 for illustration.

```
Add N persons
Add each person to a random number of narrow
environments ~(Pois(λ_e) + 1)
for t = 1 to T do
  Prepare for interactions
  I ← random number from Pois(λ)
  for i = 1 to I do
    p ← random person with Pr(choose p) ∝ w_pq
    p interacts with random connection q: Pr(choose q) ∝ w_pq
  end for
  Update arc weights
  for all Persons p do
    for all Layers l do
      Promote q: L_p(q) = l ∧ w_pq = max_{v:∃(p,v)∈A} w_pv to next layer with
prob. π_l
    end for
  end for
end for
```

Listing 1: General Algorithm of the Model

The algorithm of the model is shown in Listing 1. It has several particularities. Firstly, every period has a fixed number of interactions I drawn from Poisson distribution with mean λ. During that period, exactly I interactions happen, but although during the model calibration stage, λ will be set with a certain number of interactions per person in mind, the set-up does not guarantee that every person will perform interactions in that period. In fact, persons with higher out-degree (i.e. more connections) will have higher probability to perform any particular interaction. Once the person is chosen, it interacts with someone from its connections, again chosen with probability proportional to the arc's weight. With that, more connected persons interact more often than less connected persons and a person interacts more often with more important connections than with less important ones.

Secondly, arc weights are updated after every period based on the interactions data. Recall that quality does not change, so the only factor that can change the arc's weight at this point is a change in the relative frequency of interaction. Note that it means that as a result of a period, the arc's weight can decrease.

Thirdly, the increase will not be dramatic for two reasons. The first is that connections cannot be moved to lower layers, whatever is the change in the arc's weight. The second is that the higher the weight of the arc, the higher the probability that an interaction happens and, hence, that the relative proportion of interactions is higher and, thus, the arc's weight is around the previous period's value or higher than it.

Fourthly, in the end of every period, every person takes its best-performing connection (the one with the maximum weight) in every layer and promotes it probabilistically. The probability of promotion decreases with layer, so that it is much more difficult to be promoted from layer 3 to layer 4 than from layer 1 to layer 2.

Table 1: Parameter Values

Parameter Name	Notation	Value
Length of simulation, periods	T	100
Number of persons	N	1000
Number of narrow environments	M	35
Mean narrow environments of a person	λ_e	1.5
Quality distribution parameters	(μ, σ)	(0.5, 0.3)
Weight of a new arc		1.0
Mean global number of interactions	λ	10^5
$\Pr(\text{promotion}\mid\text{layer} = 1)$	π_1	0.60
$\Pr(\text{promotion}\mid\text{layer} = 2)$	π_2	0.20
$\Pr(\text{promotion}\mid\text{layer} = 3)$	π_3	0.05

Table 2: Out-Degree Statistics by Number of Narrow Environments (M)

M	Median Out-Degree	
	Mean	Range
20	249	$[233, 264]$
25	203	$[190, 218]$
30	172	$[161, 184]$
35	149	$[139, 161]$
40	132	$[123, 141]$
45	118	$[111, 129]$
50	107	$[98, 116]$

Note: Statistics over 100 runs reported.

The model was implemented in Repast Simphony. The values of model parameters are shown in Table 1. The parameters were set so that there are on average 100 interactions per person in every period. Every person is a member of 2.5 $(= \lambda_e + 1)$ narrow environments on average, so that most of them participate in two or three narrow environments, which is close to reality. The number of narrow environments, given that the number of persons is 1000, is set so that the median out-degree is around 150 (Dunbar 1993), as determined empirically by running the graph-construction part of the model (see Table 2). The probabilities of promotion to the next layer, given the current layer, $\pi_1, \dots, \pi_3$, are set to values where by $t = 100$, the layered structure in terms of layer sizes on average reproduces the values reported in Zhou et al (2016).

After conducting 30 runs, the average sizes of the four layers at t = 100 are close to the target sizes reported in Zhou et al (2016), see Fig. 2.

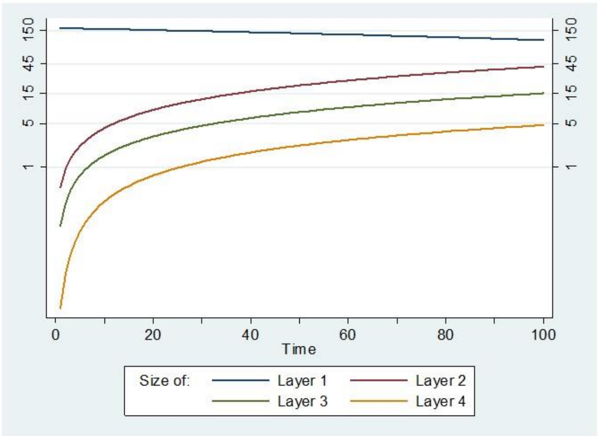

Figure 2: Size Dynamics of Layers over Time, Stage 1

4. STAGE 2: A BALANCING MECHANISM

In the previous stage, the model operated in the unrestricted layer size growth mode, which is an obvious disadvantage. After all, while the parameters were set to values where the expected average sizes of layers are attained at $t = 100$, the lack of a balancing force means that at large values of t (and with no changes in the membership in narrow environments), all connections will be in the highest layer of every person. In the second stage, a mechanism restricting the size of every layer starting from the second layer is introduced (the first layer cannot be restricted, as it depends on the number and sizes of narrow environments with which a person is affiliated).

There are several ways to introduce layer size balancing. In this model, the mechanism of attractors is applied. Experience-weighted attraction is used by Gemkow and Neugart (2011) to learn the best number of connections, but it is a machine-learning algorithm allowing to choose from several existing strategies, while in this model, attractors are built in the person. In summary, every person has a random attractor for every layer (starting from layer 2). The attractor is set to a value that keeps the respective layer size in the needed boundaries. After a new promotion to a layer, if that action resulted in a substantial upward deviation from the attractor of that layer, the person tries to demote the least worthy persons from the layer to the previous layer.

Attractors are a mathematical formalism applied in the analysis of dynamical systems. The attractors used in this model belong to a variety of attractors called *fixed point attractors*, because in every layer of a person's social network, there is exactly one point where the size of the network's layer will finally rest. Any deviation from that point leads to a pressure to change the network to return to it. See Vallacher and Nowak (2007) for details and examples of applying attractors in social psychology.

In detail, assume that person j is promoted to layer l in person's i social network. Denote the quality of correspondence between i and j as q_{ij} and i's attractor of layer l as $\alpha_i(l)$. Person i tries to keep the sum of quality correspondence with each connection at layer l, $Q_i(l)$, above but maximally close to $\alpha_i(l)$. Thus, when j is promoted to l, if $Q_i(l) > \alpha_i(l)$, person i moves all persons k from layer l with the lowest q_{ik} to layer $(l - 1)$ until removing any other person from that layer would move $Q_i(l)$ below $\alpha_i(l)$. Listing 2 formalises these actions.

$Q_i(l) \leftarrow \sum_{k:L_i(k)=l} q_{ik}$
$\Delta_\alpha \leftarrow Q_i(l) - \alpha_i(l)$
flag $\leftarrow$ **false**
while $\Delta_\alpha > 0 \wedge$ **not** flag **do**
  $q \leftarrow \min_{k:L_i(k)=l} q_{ik}$
  $p \leftarrow k: L_i(k) = l \wedge q_{ik} = q$
  if $\Delta_\alpha - q > 0$ **then**
    $\Delta_\alpha \leftarrow \Delta_\alpha - q$
    Demote p to layer $L_i(p) - 1$
  else
    flag $\leftarrow$ **true**
  end if
end while

Listing 2: Application of Attractors After Promoting j in i's Network

Assume for example that a certain layer of i's network contains persons with quality correspondence of 0.4, 0.5 and 1.0, which sum to 1.9, and that a person with quality correspondence of 0.8 is promoted to this layer. If the attractor is $\alpha_i(l) > 2.7$, no demotion occurs. If it is $2.3 < \alpha_i(l) \leq 2.7$, again no demotion occurs, because the worst correspondence is 0.4 but removing it from this layer would result in $Q_i(l)$ falling below $\alpha_i(l) > 2.3$. Finally, if $\alpha_i(l) \leq 2.3$, e.g., it is $\alpha_i(l) = 2.0$, the person with quality correspondence of 0.4 is moved to a lower layer, resulting in $Q_i(l) = 2.7 - 0.4 = 2.3$. Because the worst correspondence is now 0.5, and removing it from this layer would result in $Q_i(l) = 1.8 < 2.0 = \alpha_i(l)$, the demotion procedure stops.

For every person, attractors are chosen from uniform distributions: $[3, 5]$ for $\alpha(4)$, $[9, 15]$ for $\alpha(3)$ and $[30, 45]$ for $\alpha(2)$. These distribution boundaries were set according to empirical data in Zhou et al (2016).

The use of attractors means that a person has a built-in need for a certain overall quality of interactions at each level of its network. Overall quality in this case is approximated by the sum of closeness in quality of the person with its connections on a given level.

Too low overall quality of interactions is considered unsatisfactory, but the person is not ready to attain the attractor of a given layer immediately (by promoting a sufficient number of connections from lower layers). That means that persons are in a sense risk-averse, reluctant to quickly let others have close relationships with them.

Too high overall quality of interactions is also considered unsatisfactory, because it puts too much pressure on the person. It is, thus, assumed that interactions not only give satisfaction, but also require certain effort. Above the attractor, the person feels discomfort and becomes willing to decrease the overall quality of interactions.

At the second stage, the model was run 30 times with the same parameter values shown in Table 1 except for T, which was increased to 200 to check how fast layer size stabilisation occurs.

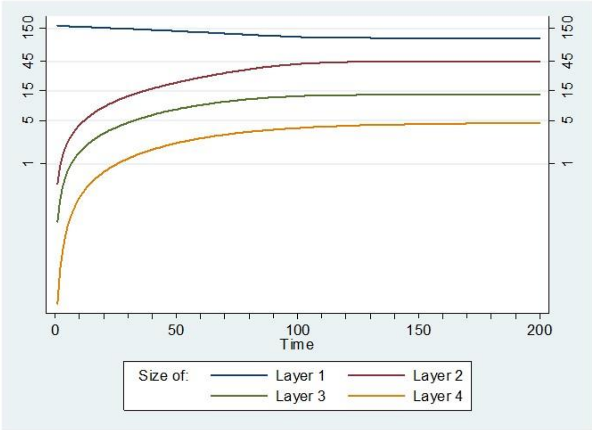

Figure 3: Size Dynamics of Layers over Time, Stage 2

As shown in Fig. 3, the use of attractors allows to stabilise the sizes of layers by period 150. By construction, the stable size levels mimic the empirically observed results.

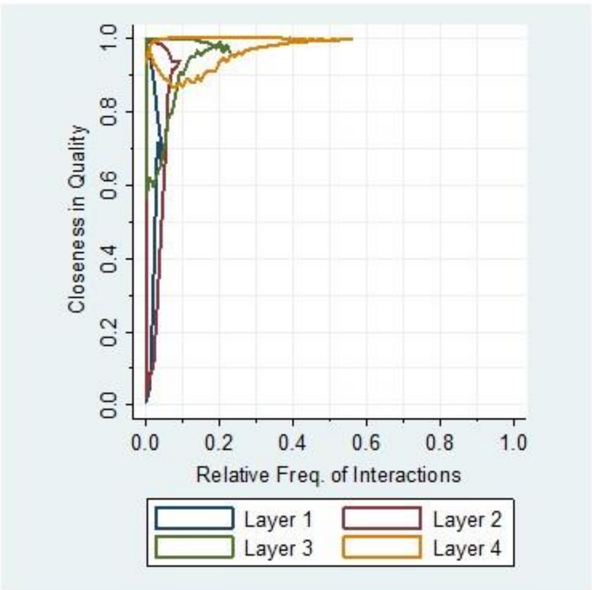

Figure 4: Distribution of Connections in Unit Squares by Layer at $t = 200$

Fig. 4 shows the regions of the unit squares where connections from each of the four layers are located. It is generally visible that, again by construction, more time is spent interacting with those connections which are closer in quality to the person. However, at higher layers, increasingly more time is spent with those with closest quality. In fact, it is clearly shown that the closeness in quality is not very important to be selected for promotion to the second layer. At the same time, no agent has on average lower quality correspondence than around 0.6 in the third layer and around 0.9 in the fourth layer. Neither closeness in quality nor membership in higher layers guarantees high relative frequency of interactions, as shown by all four layers intersecting around the top left corner of the unit square.

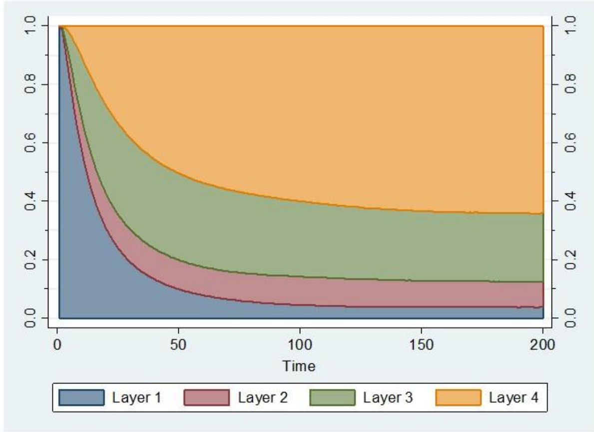

Figure 5: Average Time Share Spent on Interactions by Layer

The relative amount of time spent interacting with the four layers also stabilises by period 150, see Fig. 5. By the end of the simulation, more than 60% of time is spent on interactions with the highest layer, around 20% with the third layer, around 10% with the second layer, while only less than 5% of time is spent with the first layer.

5. DISCUSSION AND CONCLUSIONS

The paper presented only the first two stages of building the Stimulation model, but it already allows to reproduce sizes of four connection layers: support clique (3–5 vertices), sympathy level (9–15 vertices), band level (30–50 vertices) and community level (around 150 vertices) (Zhou et al 2016).

However, the contribution of this model was not only in reproducing that structure, but also in moving the focus from the ‘skeleton’ of the network – whether or not two given vertices are connected – to its ‘meat’ – the dynamics of the weights of arcs and the promotion of vertices to higher levels in the opinion of their connected vertices. The ‘skeleton’ part is focused on narrow environments, where everyone knows everyone, instead of purely random connections, as modelled usually. The ‘meat’ part is focused on the correspondence in quality and relative frequency of interactions as determinants of the arc’s weight and becoming a candidate for a promotion to a higher layer.

Further work on this model should start with allowing persons to probabilistically enter and exit narrow environments instead of assigning them to a (randomly) defined number of narrow environments. It is expected that this change will allow to generate a vertex degree distribution that follows power law. As a result, the model will allow to generate a multilayer scale-free social network.

REFERENCES

Boccaletti, S., Bianconi, G., Criado, R., del Genio, C., Gómez-Gardeñes, J., Romance, M., Sendiña-Nadal, I., Wang, Z., Zanin, M., 2014. The structure

and dynamics of multilayer networks. Physics Reports 544(1), 1–122
Cantor, M., Shoemaker, L.G., Cabral, R.B., Flores, C.O., Varga, M., Whitehead, H., 2015. Multilevel animal societies can emerge from cultural transmission. Nature Communications 6(8091)
Dunbar, R.I., 1993. Coevolution of neocortical size, group size and language in humans. Behavioral and Brain Sciences 16(4), 681–694
Fischer, C.S., 1982. To dwell among friends: Personal networks in town and city. The University of Chicago Press, Chicago
Gemkow, S., Neugart, M., 2011. Referral hiring, endogenous social networks, and inequality: an agent-based analysis. Journal of Evolutionary Economics 21(4), 703–719
Grossetti, M., 2005. Where do social relations come from?: A study of personal networks in the Toulouse area of France. Social Networks 27(4), 289–300
Li, B., Sun, D., Li, Z., Ou, C., Zhu, R., 2015. Understanding the dynamic of covert organization with multilayer networks model. In: European Intelligence and Security Informatics Conference (EISIC) 2015. pp. 37–44
Li, Z., Yan, F., Jiang, Y., 2015. Cross-layers cascade in multiplex networks. Autonomous Agents and Multi-Agent Systems 29(6), 1186–1215
McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: Homophily in social networks. Annual Review of Sociology 27, 415–444
Poledna, S., Molina-Borboa, J.L., Mart´ınez-Jaramillo, S., van der Leij, M., Thurner, S., 2015. The multilayer network nature of systemic risk and its implications for the costs of financial crises. Journal of Financial Stability 20, 70–81
Ramezanian, R., Magnani, M., Salehi, M., Montesi, D., 2015. Diffusion of innovations over multiplex social networks. In: International Symposium on Artificial Intelligence and Signal Processing (AISP) 2015. pp. 300–304
Różewski, P., Jankowski, J., Bródka, P., Michalski, R., 2015. Knowledge workers' collaborative learning behavior modeling in an organizational social network. Computers in Human Behavior 51(B), 1248–1260
Senior, A.M., Lihoreau, M., Buhl, J., Raubenheimer, D., Simpson, S.J., 2016. Social network analyses and nutritional behavior: An integrated modeling approach. Frontiers in Psychology 7(18)
Vallacher, R.R., Nowak, A., 2007. Dynamical social psychology: Finding order in the flow of human experience. In: Social Psychology: Handbook of Basic Principles, pp. 734–758. The Guilford Press, New York
Zhou, W.X., Sornette, D., Hill, R.A., Dunbar, R.I.M., 2005. Discrete hierarchical organization of social group sizes. Proceedings of the Royal Society of London B: Biological Sciences 272(1561), 439–444