Proceedings of the 2nd Workshop on Complex Networks and their Applications at SITIS 2013 - The 9th International Conference on Signal Image Technology & Internet based Systems, December 2-5 2013, Kyoto, Japan

IEEE Computer Society, pp. 499-505
DOI: 10.1109/SITIS.2013.85

Convince a Dozen More and Succeed - The Influence in Multi-layered Social Networks

Radosław Michalski, Przemysław Kazienko
Institute of Informatics, Wrocław University of Technology
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
tel. +48 71 320 36 09, fax +48 71 320 34 53
{radoslaw.michalski, kazienko}@pwr.wroc.pl

Jarosław Jankowski
Faculty of Computer Science
West Pomeranian University of Technology
Al. Piastów 17, 70-310 Szczecin, Poland
jjankowski@wi.zut.edu.pl

***Abstract*—Humans utilise multiple communication channels in their social interactions and also information diffusion as well as the spread of influence are practically related with many contexts. Each such context (channel) may represent a different communication method or a different environment of a given person. This facilitates building multiple social networks, that are not independent. They share the same set of nodes connected with many links grounded on different layers – these networks are called multi-layered or multiplex social networks. The influence process may vary in these kinds of social networks depending on the network model, the level of influence for each layer and other factors such as the overlap of nodes and links across layers. In this paper, the influence processes in multi-layered social networks have been analysed showing that for almost all analysed network models, the success in convincing few more individuals may be crucial for the whole influence process. The results revealed that the process is not linear in terms of relation between the number of initially influenced individuals and the total number of influenced nodes. The linear threshold model has been utilized as a base influence model.**

Keywords—social network analysis; multi-layered networks; multiplex networks; spread of influence; linear threshold

I. INTRODUCTION

Nowadays, as the technocratic and dynamic world is offering us a variety of different types of communication methods, it is impossible to use only one type of interaction with our acquaintances or friends. Various communication techniques are used: e-mail services, instant messengers, faxes, landline and mobile phones and, last, but not least, face-to-face communication. By means of these mediums, we exchange day-to-day information, gossip, news and corporate correspondence, hence all these interactions result in a gigantic worldwide social network. However, the simplification of treating all of these interactions equally is not a desirable way of dealing with such a complex environment as it may result in hiding crucial information from researchers.

As the above communication methods have unique properties, such as synchronisation type (synchronous vs. asynchronous), number of simultaneous recipients or reciprocity, the resulting social networks built from various communication channels may be, and mostly are, different. However, what binds these networks together is that one person on one medium (layer) is the same human being on another layer – resulting in a multi-layered social network [1]. In this case, assuming that a person is taking part in a diverse information diffusion process, once somebody receives some information, he or she may spread it over all of the layers he or she is a member of. The dynamics of that process vary from one layer to another, depending on its type, properties and individual preferences. It results in diffusion process stretched among the layers that have different dynamic profiles. And, what is more interesting from this paper's aspect, the multi-layered nature of the network also changes the spread of the influence profile. As the process of diffusion of information differs for multi-layered networks compared to flat ones, the same case applies to the diffusion of innovations process resulting in different adaptation speeds and directions.

This paper studies how one of the models for the spread of influence works in a multi-layered social network composed of different types of layers. The main goal of this manuscript is to answer the question of how the number and types of layers influence the results of the diffusion of innovation following the linear threshold model.

The paper is organized as follows: The next section describes the problem of the linear threshold influence model in multi-layered social networks combined with the related work description. Section III contains the experimental setup, whereas section IV presents the results with the discussion in the following, fifth section. The last section concludes and presents future work directions.

II. RELATED WORK AND PROBLEM DESCRIPTION

A. *Multi-layered Social Networks (MSN)*

A multi-layered social network, referred to also as a multiplex network, is the natural way to represent and analyse the differentiated contacts between users, because it reflects the diverse nature of communication or interaction [1], [2]. This preservation of the communication type or any other type of differences between layers allows us to study the properties of networks separately at each layer, or jointly, by merging the layers.

A typical multi-layered social network is presented in Fig. 1. Different layers may represent different types of communication or different user environments. On each layer the single component network may differ, as well as the number of nodes and edges. What is shared by them all is that the same social entity may exist on a number of layers and it is virtually connected to itself, building a bridge between these layers[1].

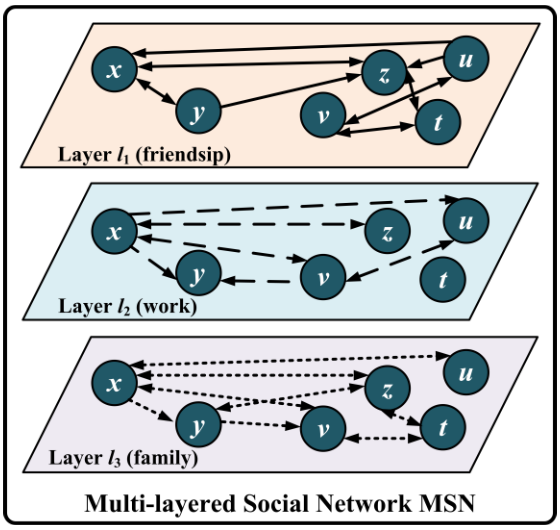

Fig. 1. An example of multi-layered social network [2]

The formalism behind multi-layered social networks is as follows [2]: A multi-layered social network (MSN) is a network extended to multiple edges between pairs of nodes/actors. It is defined as a tuple $<V, E, L>$ where:

V is a not-empty set of actors (social entities),

E is a set of tuples $<x,y,l>$, $x,y\in V$, $l\in L$, $x\neq y$ and for any two tuples $<x,y,l>$, $<x',y',l'>\in E$ if $<x,y,l>\neq<x',y',l'>$, $x=x'$ and $y=y'$ then $l\neq l'$,

L is a fixed set of distinct layers (types of relationships).

One of the reasons why the preservation of layer information is crucial is the dynamic processes in networks such as information diffusion or the spread of influence. In this case, each layer has a different speed for information propagation, and may have different sets of nodes and edges. All these distinctions may strongly affect the influence processes. In fact, some recent research is devoted to the study of the diffusion of information or cooperation within multi-layered networks [3], [4]. However, with this initial work completed in the area of analysing the diffusion of information in MSNs, to the authors' best knowledge, there exist just preliminary research in this area [5] or some research focuses on particular types of layers [18].

[1] It may also happen in the real world that a node on one layer (system, communication channel) may correspond to multiple nodes on another layer, e.g. a single user account in Facebook may be related to many email addresses, see [17] for more in-depth discussion on internet identities.

B. Diffusion of Innovations and the Influence Models

The problem of the adoption of the innovations in society is extensively described in [6] and the results of this research led to the formulation of different influence models [7], [8], [9]. Among others, two of them are described below with particular reference to the linear threshold model.

The first is the Linear Threshold (LT) model which is based on the assumption that the spread of information is the result of the impact of the direct neighbourhood on the node [7]. Infection is related to the ratio of infected neighbouring nodes to all surrounding nodes.

In this model, a node *v* is influenced by each of *v*'s neighbour *w* according to a weight $b_{v,w}$ such that

$$\sum_{w\,neighbour\,of\,v} b_{v,w} \leq 1 \qquad (1)$$

Then each node *v* chooses a threshold θ_v uniformly at random from the interval [0,1]. This interval represents the weighted fraction of *v'*s neighbours that must become influenced in order for *v* to become influenced as well [10]. Given a random choice of threshold or the same threshold for every node in a simplified model [11], an initial set of influenced nodes is chosen. The diffusion process unfolds deterministically in discrete time steps. In step t–1, all influenced nodes remain influenced and every node *v* is activated for which the total weight of its active neighbours is at least θ_v:

$$\sum_{w\,active\,neighbour\,of\,v} b_{v,w} \geq \theta_v \qquad (2)$$

Another model, the Independent Cascade (IC) model is assumed to trigger a cascade of influence on the basis of the probability of influence by the already infected node [10]. In this model, the influenced node has just a single chance to influence each of its neighbours with a certain probability. If the process fails, there is no second chance. So, these two methods LT and IC are different. The former represents the potentially influenced perspective and the latter focusses on the influencer.

The study of diffusion mechanisms on the basis of these models is focussed on maximising the range [10], the selection of nodes that should be selected in the early stages of infection – the seeding strategies [10], [12] and the analysis of the factors and parameters of the social network, which affect the range [13]. But still, none of these works are trying to extend the model to the multi-layered social network environment.

C. Problem Description

The main problem discussed in this paper is how the spread of influence depends on the structure of individual layers in a multi-layered social network and whether the reach of the influence grows linearly with the number of initial seeds. To answer that question the authors focused on only one of the influence models – linear threshold (LT) – and generated a number of multi-layered networks where layers represented various network models – Watts–Strogatz small-world model [14], Barabási-Albert preferential attachment model [15], Erdős–Rényi random graph [16], including square lattice as well. As it is a purely artificial setup, some assumptions were made and they are presented in section IV.

Symbolically, the process of influence in a multi-layered social network is presented in Fig. 2. It presents a three-layered network, where non-influenced node $x_{1,l}$ exists in layer with the second subscript l indicating the layer number with some probability $p_{1,l}$. The influence propagation process occurs in *Layer 1* and *Layer 2*, assuming that in Layer 3 there are currently no influenced nodes in the neighbourhood of x_1. Due to the fact that the ratio of influenced neighbours of that node to all the neighbours reached the threshold value, this node may also influence the neighbours at the third layer in the next time step, changing the dynamics and reaching of the spread of influence in the whole MSN.

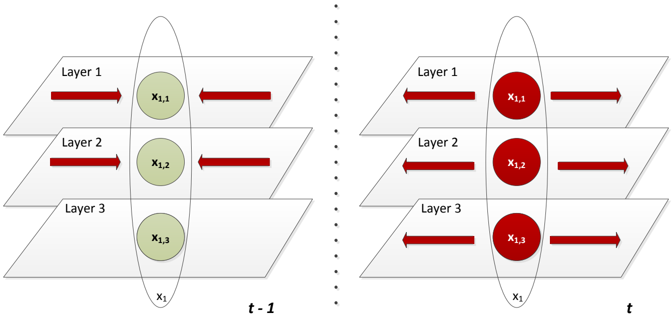

Fig. 2. The spread of influence in a multi-layered social network

III. EXPERIMENTAL SETUP

A. Analysed Scenarios

In this work, multi-layered networks of different type and size were evaluated to see how the properties of the networks at each layer influence the linear threshold influence process. Four kinds of networks were analysed:

- Watts–Strogatz small-world model (WS) [14],
- Barabási–Albert preferential attachment model (BA) [15],
- Erdős–Rényi random graph (ER) [16],
- square lattice (LA).

It was decided to evaluate how the spread of influence behaves in two- and three-layered networks combined of the same or different network models. To limit the parameterisation, just two global parameters of the networks and a small set of assumptions were introduced, which are described in the next subsection.

The most important analysed properties of the whole process were: the probability that the cascade of influence succeeds, i.e. all the nodes will adapt the idea; and the number of influenced nodes in time representing how the topology affected the dynamics of the process.

All the networks were assumed to be undirected and all the experiment passes were repeated hundred times to average the results. For the layer size N, i.e. the number of nodes, the same for all layers has been chosen – 10 000 unique nodes. Due to the fact that bigger networks behaved similarly, these results may be generalized for bigger datasets.

B. Assumptions and Parameters

Multi-layered social networks do not assume that each node is required to exist at each layer, because it is unlikely to find exactly the same set of users among different layers in different than artificially created environments.

Authors of this work assumed that to reach the goal of pure topological analysis, two other parameters should also be simplified: the θ parameter, which is the threshold level is constant for each node and in the multi-layered network setup, the sum of weights of neighbours is calculated across all of the layers. This implies that there is no distinction in influence probability at each layer. So current experiments were conducted with the assumption of the same importance at each layer. Additionally, it is assumed that the influenced person influences at all the layers it belongs to.

For some researchers, especially those familiar with computer networks, the term *layer* may provide some confusion by suggesting that the order of layers matters. In multi-layered social networks this is not true and the term *layer* corresponds to different types of activity or relationship, not the level in the hierarchy – the layers may be swapped around without any consequences.

The relaxation of the above assumptions is planned as a future work direction, as mentioned in the last section.

Two global parameters affecting the spread of the influence were finally used:

- the threshold level θ, the same for all nodes,
- the probability of node existence on a layer - P, the same for all nodes,
- the number of initial seeds S.

Some explanation has to be provided on how the probability of node's existence on a layer was used. Initially, the first layer the node will belong to was drawn with the same probability for all layers. Later, with the probability P the node could exist on other layer(s) in the MSN setup. However, the node existing on multiple layers was counted as just one node.

Network-specific parameters of individual layers are presented in Table I.

TABLE I. NETWORK-SPECIFIC PARAMETERS

Network type	***Parameter***	***Value***
Watts-Strogatz	Rewiring probability	0.05
Barabási-Albert	Power	1
Erdős–Rényi	Wiring probability	0.02

IV. Results

A. Two-layered Social Network

Three levels of the threshold parameter θ were evaluated to see how the percentage of seed influences the process: 0.25, 0.33 and 0.50. The percentage of seeds varied from 1% to 50%, but the results revealed that for maximum of 35% of initial seeds, all other nodes became influenced. The results for two-layered networks are presented in Fig. 3 and Fig. 4, with the values of P=1 and 0.5, respectively.

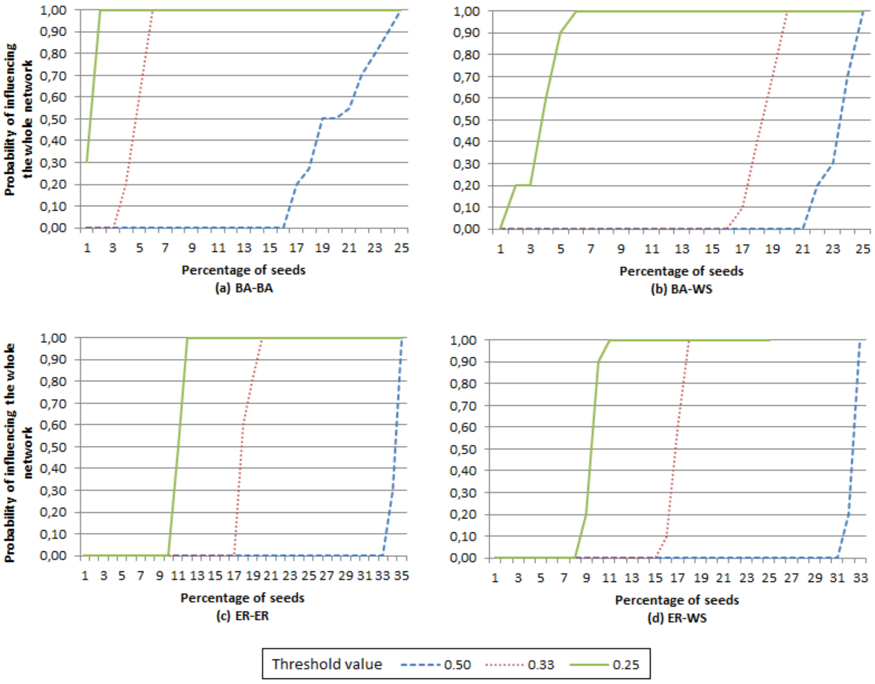

Fig. 3. The probability of influencing the whole network for two-layered networks of different type with P=1 (WS – Watts–Strogatz, BA – Barabási–Albert, ER – Erdős–Rényi)

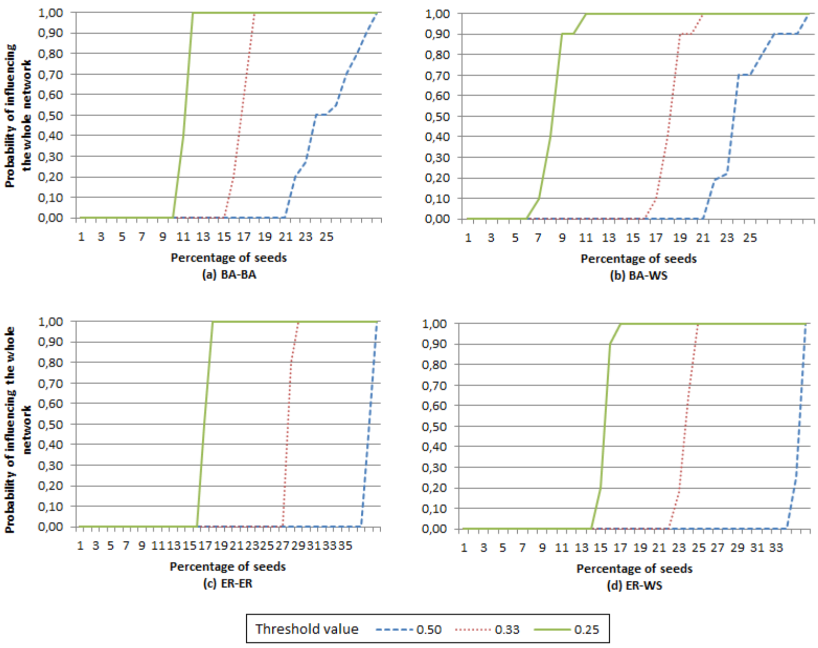

Fig. 4. The probability of influencing the whole network for two-layered networks of different type with P=0.5 (WS – Watts–Strogatz, BA – Barabási–Albert, ER – Erdős–Rényi)

B. *Three-layered Social Network*

Similar experiments were conducted for three-layered social networks consisting of layers of different type.

The most interesting results containing six of all evaluated networks are presented in Fig. 5.

Since the influence of *P* parameter provided similar results than for two-layered networks, the authors decided not to present them here.

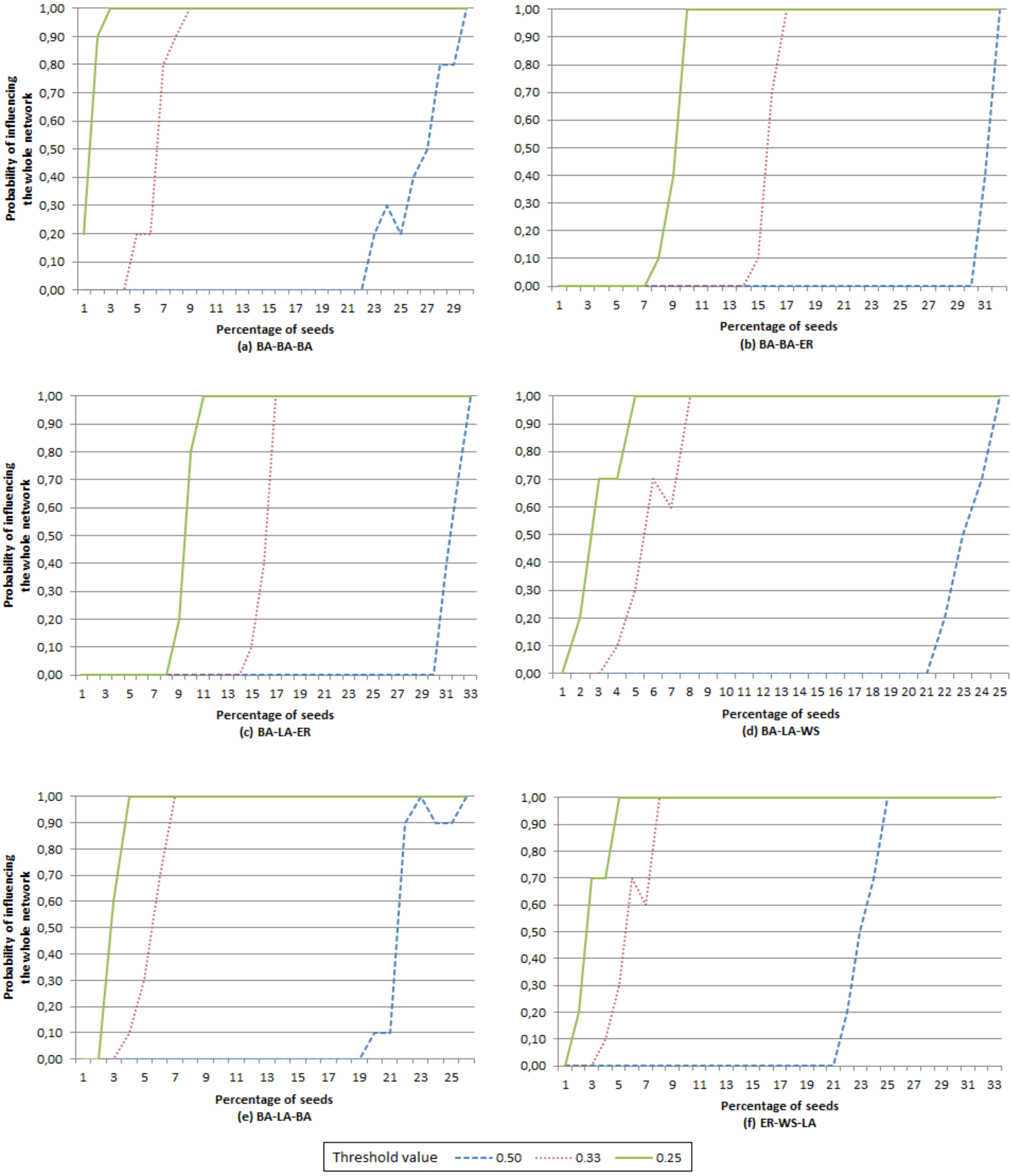

Fig. 5. The probability of influencing the whole network for three-layered networks of different type (WS – Watts–Strogatz, BA – Barabási–Albert, ER – Erdős–Rényi, LA – square lattice), P = 1

Moreover, to compare particular models in more detail for the highest evaluated threshold level, four of the network types were compared each other by showing the average number of seeds infected for different seed percentage. The results are presented in Fig. 6.

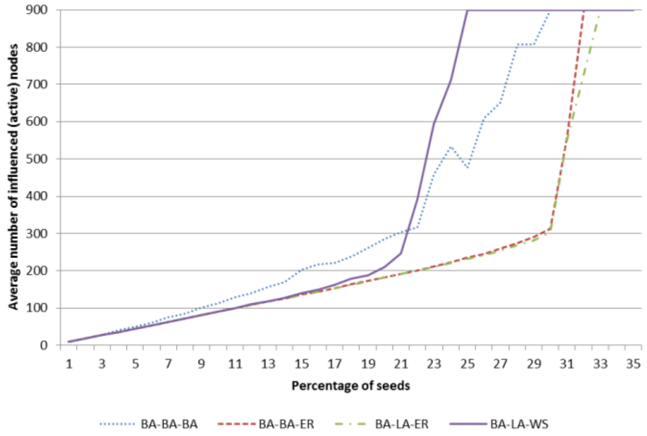

Fig. 6. Average number of influenced nodes for different percentage of seeds for θ=0.5, P=1

To observe how infections occur in time, Fig. 7 shows the percentage of influenced nodes at each step of the whole process for four chosen networks.

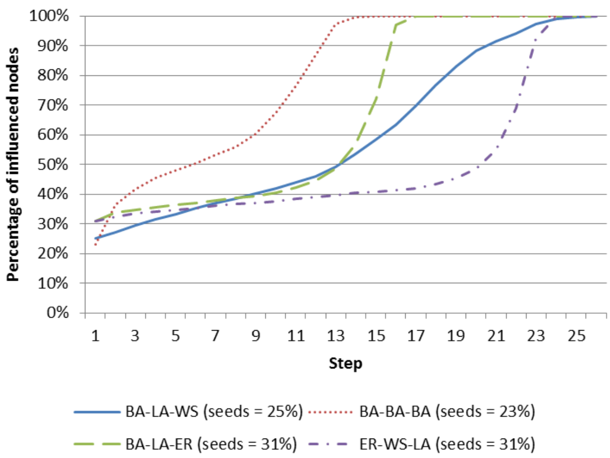

Fig. 7. Percentage of influenced nodes in particular process steps, parameters: θ=0.5, P=1

V. DISCUSSION

For the smallest analysed threshold at the level of 0.25 the BA-BA and BA-LA networks were getting highest probability of infection with just 3% of seeds assuming that all the nodes exist on both layers. Multi-layer structures based on other networks required 9% of seeds with up to 11% for ER-ER networks. For the 0.33 threshold the diffusion at BA-BA networks was performing well with only 5% of seeds. For other networks combinations this threshold required 13% up to 19% of seeds. For highest analysed threshold value at the level of 0.50 at BA-BA networks there the increase of probability from 15% to 25% of seeds was observed. But connecting BA network with ER model resulted the need of seeding about 30% of network to get high probability of infecting the whole network. ER graphs were generally most difficult to infect not only with BA but with WS and ER graphs together as well with minimal seeding at the 31% level up to 35% for ER-ER layers. For two-layered network changing the threshold from 0.25 to 0.33 with 0.08 increase was resulting the need of increased about two times more seeds, while changing from 0.33 to 0.50 by 0.17 required about to increase seeding from the range of 7% up to 25% or more to get highest probability of whole network infection. Results showed that for successful infections of entire network in multilayer approach relatively high percentage of seeded nodes was required. In most cases probabilities of reaching whole network was limited with seeding less than 5% of nodes.

One of the most important results is that the probability of nodes existence does not change the process in other way than just more seeds are needed and this is one of the most interesting results of this work. Authors expected that with decreasing probability of the node's existence on other layers, the process will become more linear. But, in fact, it just moved towards higher percentage of seeds.

The three layer based experiments showed that in most cases with the changes of the threshold from 0.25 to 0.33 the increase of seeding from 3% to 8% in most analysed networks was required, apart from BA-BA-ER and BA-LA-ER where this values were 10% and 17% respectively. But changing the threshold value from 0.33 to 0.50 resulted in much higher increase in the needed seeding - up to 25% for BA-LA-WS or 33% for BA-LA-ER structure.

Changes of percentage of seeds was affecting the increase of infected nodes with similar linear characteristics up to 21% of seeds then for BA-LA-WS and BA-BA-BA was observed higher dynamics with the increase of infected nodes. For BA-BA-ER and BA-LA-ER similar was observed after 31% percentage of seeds.

Results showed that increase of percentage of infected nodes for most networks up to 13-th step goes with stable speed and for each network can be observed critical point when dynamics changes within few simulation steps. For the BA-LA-WS and 25% of seeds and for BA-LA-ER with 31% of seeds it was observed after 14-th step of simulation. For BA-BA-BA network with 23% of seeds it happened earlier after 3-rd step and at the latest stage for the layers based on BA-LA-ER models with 31% of seeds.

To conclude the discussion part, it is observed that each layer type introduces specific properties. For instance, the most hindering layer was the Erdős–Rényi graph, despite the fact that its properties were set up to build single cluster. The limitations introduced by this model were hardly to overcome by adding different layer types, but the best layer types in that case were the Watts–Strogatz and the Barabási–Albert models. But, surprisingly, the two-layered combination of these two models was not as good as one could think. Moreover, an interesting result was the way how the nodes become influenced in time steps, because some layer combinations were introducing a very smooth increase in number of influenced nodes while some others were waiting for a particular number of active nodes to start rapid influence process in next few steps. The most interesting result is that the changing percentage of *P* parameter, i.e. the probability of existence of nodes on other layers does not change the type of an influence curve – it just moves it to the right. As authors are far from defining this process as phase transition, it is observed that in multi-layered social network it is worth to invest more money or time in convincing slightly bigger set of initial spreading users. As the relationship between number of seeds and the chances for influencing the whole network is not linear, the success of the campaign may be the case of dozens more influenced users.

VI. Conclusions and Future Work

The goal of this work was to analyse how the Linear Threshold influence model behaves in a multi-layered social network (MSN). By combining the same and different types of the networks, the authors built two- and three-layered MSNs to face the question what role the topology plays in the spread of influence within the mentioned model.

The presented results have shown that it is possible to control and adjust the share of seeds to obtain desired results based on the knowledge about multi-layered network topology. Marketers want to minimize number of seeds and the experimental outcome revealed that matching networks to one of the tested models can give some hints on the seeding strategy. Simulations proved that results of random seeding are dependent on types of networks in each layer. The analysis of the influence of the threshold level on the whole process effects can be helpful in determining the potential of viral action in the multi-layer networks. As the results have shown that it makes sense to lower down the threshold level, for instance by using more influential content. Networks based on BA models can be infected with relatively small seeds sets, while ER networks are most difficult to infect the whole network. Nevertheless, most of the real-life multi-layered social networks are consisting of layers of the same type, but built for different contexts, such as private life or work and these are rarely random. In that case it seems that the number of layers is not the limitation in the spread of influence, what may be considered either negative or positive, depending on the intentions of influencer.

Still, one of the most unexpected results was that the varying the probability of node’s existence on other layers does not change the influence probability curve type. It appears that the LT process behaves similarly with smaller probability, but it just needs more seeds to succeed.

This work may be considered as the introduction for an in-depth analysis of influence models in a multi-layered environment. Some future work directions are to be seen to weaken the assumptions used in this work. It is worth analysing how this influence model will behave in the case of directed networks with different reciprocity probability at each layer. The most interesting future work direction is the extended analysis of the probability of node’s existence on each layers in relation to probability of infecting the whole network.

Acknowledgment

This work was partially supported by the European Union as part of the European Social Fund, the European Commission under the 7th Framework Programme, Coordination and Support Action, Grant Agreement Number 316097, ENGINE - European research centre of Network intelliGence for INnovation Enhancement (http://engine.pwr.wroc.pl/) and the Polish National Centre of Science - the research project 2010-2013. The calculations were carried out in Wroclaw Centre for Networking and Supercomputing (http://www.wcss.wroc.pl), grant No 177.

References

[1] M. Magnani and L. Rossi, “The ML-Model for Multi-layer Social Networks,” *ASONAM 2011 International Conference on Advances in Social Networks Analysis and Mining*, IEEE, pp. 5–12, Jul. 2011.

[2] P. Bródka, P. Kazienko, K. Musiał, and K. Skibicki, “Analysis of Neighbourhoods in Multi-layered Dynamic Social Networks,” *International Journal of Computational Intelligence Systems*, vol. 5, no. 3, pp. 582–596, 2012.

[3] J. Gómez-Gardeñes, I. Reinares, A. Arenas, and L. M. Floría, “Evolution of cooperation in multiplex networks,” *Scientific reports*, vol. 2, 2012.

[4] S. Gómez, a. Díaz-Guilera, J. Gómez-Gardeñes, C. J. Pérez-Vicente, Y. Moreno, and a. Arenas, “Diffusion Dynamics on Multiplex Networks,” *Physical Review Letters*, vol. 110, no. 2, p. 028701, Jan. 2013.

[5] J. Jankowski, M. Kozielski, W. Filipowski, R. Michalski, “The Diffusion of Viral Content in Multi-layered Social Networks”, *ICCCI 2013 – 5th International Conference on Computational Collective Intelligence Technologies and Applications. Lecture Notes in Artificial Intelligence LNAI* vol. 8083, pp. 30-39, Springer, Berlin 2013.

[6] E. M. Rogers, *Diffusion of innovations*, vol. 65, no. 5. Free Press, 1995.

[7] M. Granovetter, “Threshold Models of Collective Behavior,” American Journal of Sociology, vol. 83, no. 6, pp. 1420–1443, 1978.

[8] M. W. Macy and R. Willer, “From Factors to Actors : Computational Sociology and Agent-Based Modeling,” *Annual Review of Sociology*, vol. 28, no. 1, pp. 143–166, Aug. 2002.

[9] T. W. Valente, “Network Models of the Diffusion of Innovations,” *Journal of Marketing*, vol. 2, no. 3, p. 134, 1996.

[10] D. Kempe, J. Kleinberg, and É. Tardos, “Maximizing the spread of influence through a social network,” *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD’03*, p. 137, 2003.

[11] E. Berger, “Dynamic monopolies of constant size,” *Communication*, vol. 200, no. 2, p. 10, 1999.

[12] M. Bampo, M. T. Ewing, D. R. Mather, D. Stewart, and M. Wallace, “The Effects of the Social Structure of Digital Networks on Viral Marketing Performance,” *Information Systems Research*, vol. 19, no. 3, pp. 273–290, 2008.

[13] H. Ma, H. Yang, M. R. Lyu, and I. King, “Mining social networks using heat diffusion processes for marketing candidates selection,” *Proceeding of the 17th ACM CIKM 08*, p. 233, 2008.

[14] D. J. Watts and S. H. Strogatz, “Collective dynamics of ‘small-world’ networks,” *Nature*, vol. 393, no. 6684, pp. 440–2, 1998.

[15] A.-L. Barabasi and R. Albert, “Emergence of scaling in random networks,” *Science*, vol. 286, no. 5439, p. 11, 1999.

[16] P. Erdős and A. Rényi, “On the evolution of random graphs,” *Evolution*, vol. 5, no. 1, pp. 17–61, 1960.

[17] K. Musiał, P. Kazienko, “Social Networks on the Internet,” *World Wide Web*, 16(1), pp. 31-72, 2013.

[18] C.D. Brummitt, K.M. Lee & K.I. Goh, “Multiplexity-facilitated cascades in networks”. Physical Review E, 85(4), 045102, 2012.