UPCommons

Portal del coneixement obert de la UPC

http://upcommons.upc.edu/e-prints

Aquesta és una còpia de la versió *author's final draft* d'un article publicat a la revista *Physica A.*

URL d'aquest document a UPCommons E-prints:

https://upcommons.upc.edu/handle/2117/169655

Article publicat / *Published paper*:

The neighborhood role in the linear threshold rank on social networks

Fabián Riquelme[a,b,*], Pablo Gonzalez-Cantergiani, Xavier Molinero[c], Maria Serna[d]

[a]*Escuela de Ingeniería Civil Informática, Universidad de Valparaíso, Chile*
[b]*Centro de Investigación y Desarrollo en Ingeniería en Salud, Universidad de Valparaíso, Chile*
[c]*Mathematics Department and Barcelona Graduate School of Mathematics, Universitat Politècnica de Catalunya, Spain*
[d]*Computer Science Department and Barcelona Graduate School of Mathematics, Universitat Politècnica de Catalunya, Spain*

Abstract

Centrality and influence spread are two of the most studied concepts in social network analysis. Several centrality measures, most of them, based on topological criteria, have been proposed and studied. In recent years new centrality measures have been defined inspired by the two main influence spread models, namely, the Independent Cascade Model (IC-model) and the Linear Threshold Model (LT-model). The Linear Threshold Rank (LTR) is defined as the total number of influenced nodes when the initial activation set is formed by a node and its immediate neighbors. It has been shown that LTR allows to rank influential actors in a more distinguishable way than other measures like the PageRank, the Katz centrality, or the Independent Cascade Rank. In this paper we propose a generalized LTR measure that explore the sensitivity of the original LTR, with respect to the distance of the neighbours included in the initial activation set. We appraise the viability of the approach through different case studies. Our results show that by using neighbours at larger distance, we obtain rankings that distinguish better the influential actors. However, the best differentiating ranks correspond to medium distances. Our experiments also show that the rankings obtained for the different levels of neighborhood are not highly correlated, which validates the measure generalization.

Keywords: Spread of influence, Linear threshold model, Social network, Neighborhood, Centrality
2010 MSC: 91D30, 05C22, 68R10

---

[*]I am corresponding author

Email addresses: `fabian.riquelme@uv.cl` (Fabián Riquelme), `gonzalezcantergiani@gmail.com` (Pablo Gonzalez-Cantergiani), `xavier.molinero@upc.edu` (Xavier Molinero), `mjserna@cs.upc.edu` (Maria Serna)

1. Introduction

Influence spread is one of the most studied concepts in social network analysis. The way a group of actors adopting a new tendency can influence others to also adopt it, is an intuitive and well-known phenomenon [1]. Social networks are commonly represented as graphs, whose nodes are the actors of the network, and the edges are interpersonal ties among the actors [2]. Moreover, in order to study the influence spread phenomenon, it is useful to consider influence graphs, i.e., weighted and labeled digraphs, whose edge weights represent the influence power exerted by one actor over another, and the node labels quantify the resistance of each actor to be influenced by its active neighbors [3].

An influence spread model describes the ways in which actors influence each other through their interactions in a social network. The two best known general models for influence spread are the *linear threshold model* (LT-model) [4] and the *independent cascade model* (IC-model) [4]. The first one is based on some ideas of collective behavior [5, 6], while the second one was proposed in marketing contexts [7].

Centrality is another key concept in social network analysis. Centrality measures aim to determine how structurally relevant is an actor within the social network. Identifying the users relevance is particularly useful for many applications, such as viral marketing [8], information propagation [9], search strategies [10], expertise recommendation [11], community systems [12], social customer relationship management [13], percolation theory [14] and education [15]. Centrality measures can also be used to identify the most active, popular, or influential users within a network [16]. Most of the centrality measures used to identify influential actors are based on network topological criteria [16]. However, it is known that the influence spread does not only depends on the network structure, but also on the spreading dynamic process [17].

Recently, the *Linear Threshold Rank* (LTR), a centrality measure based on the LT-model, was proposed in [18]. This measure can be interpreted as how much an actor can spread his influence within a network, investing resources to be able to convince his immediate neighbors. In this measure, an actor with small degree might have a good ranking due to his neighbors. In that research was proved that LTR measure differs from other centrality measures based on influence criteria, such as PageRank [19], Katz centrality [20], and the Independent Cascade Rank (ICR) [21]. In comparison to these three measures, LTR allows to rank the actors in a clearer and more distinguishable way, with a high standard deviation, and a large number of different values [18].

Our aim is to explore the sensitivity of the LTR to the selection of the initial activation neighborhood. In particular we are interested in analyzing LTR when the initial activation set contains all neighbors at distance ℓ, i.e., the nodes contained at neighborhood *level* ℓ. It is clear that increasing the neighborhood level in the initial activation set may result in a higher number of nodes being influenced. However, the increase in the influence spread achievable by a node in terms of its neighborhood level considered in the initial activation has not been studied in depth so far. To this end, in this paper we generalize the LTR measure, by allowing more than one level of neighborhood as initial activation for the influence spread process. In addition, we propose two types of graphics to visualize the behavior of this measure. As

case studies, we consider two large social networks: a Twitter retweets network (directed) and an arXiv network (undirected). Our results show that by using higher neighborhood levels, we obtain rankings that distinguish better the influential actors. However, the best differentiating ranks correspond to medium distances. Furthermore, our experiments show that our generalized LTR measure provides different rankings for the different neighborhood levels considered.

The paper is organized as follows. Section 2 presents a brief literature review about centrality-based studies. Section 3 briefly describes the related work regarding social networks represented as influence graphs. Section 4 is devoted to the LTR measure and its generalization to consider more than one neighborhood level. Section 5 shows our experimental setting and data analysis. Finally, the paper ends by presenting our main conclusions and future work.

2. Literature review

Since 1979 [22], researchers have defined several centrality measures, based on different criteria such as the network structure (e.g., degree, closeness, betweenness, Katz centrality [20], PageRank [19], etc.), the information flow that circulates within the network (e.g., flow betweenness and flow closeness [23]), the content of such spread information (e.g., TRank [24], TS-SRW [25], etc.), the variations in the network actors behavior over time (e.g., IDM-CTMP [13], parameterless mixture [26], etc.), and even on power indices coming from cooperative game theory [27, 28, 29] or some physical principles such as velocity [30]. Furthermore, there also exist several centrality measures created for specific social networks (e.g., retweet impact [31], Acquaintance Score [32], Social Networking Potential [33], etc.). Each centrality measure aims to detect relevant actors for a particular study context. Only for Twitter, there are more than fifty influence measures [16], so in practice there are hundreds of centrality measures to choose. To know more about the centrality measures, the interested reader can see [16, 34].

Although many centrality measures seek to detect the most influential actors in a network, these measures tend not to be related to the influence spread phenomenon. The first attempts to study centrality measures considering influence spread are quite recent. Most of them are directly inspired by the IC-model [21, 35, 36, 37, 38]. Recently, some classic centrality measures have been applied together with the IC-model to minimize the bad effect of misinformation in social networks [39]. As far as we know, the first centrality measure considering the LT-model was the Linear Threshold Rank (LTR) [18], which is the starting point for the present work.

3. Social networks as influence graphs

A social network can be represented as an *influence graph* [3] (G, f, w), where $G = (V, E)$ is a graph, with the node set V representing the actors of the network, and the edge set E the influence relationships among the actors, so that $(a, b) \in E$ means that actor a exerts a direct influence over b. The *weight function* $w : E \to \mathbb{N}$ determines the influence power

$$F_0(X) = X = \{a\} \qquad F_1(X) = \{a, c\} \qquad F_2(X) = F(X) = \{a, c, d\}$$
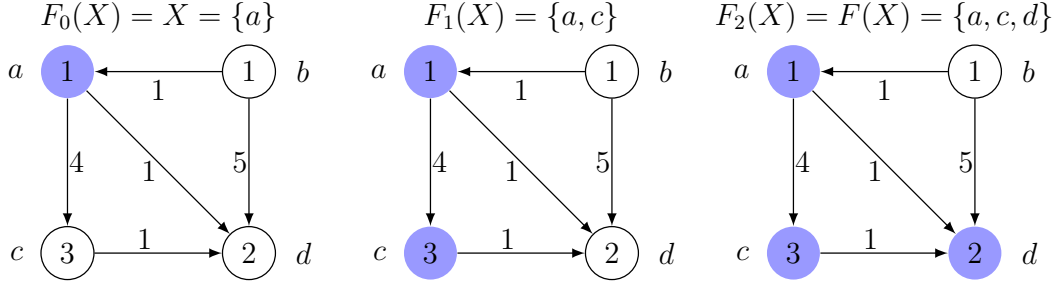
Figure 1: Influence spread on an influence graph, starting from the initial activation $X = \{a\}$.

$w(a, b)$ exerted by a over b, and the *labeled function* $f : V \to \mathbb{N}$ quantifies the resistance of each actor to be influenced. Let be $n = |V|$, and $m = |E|$.

Since the influence exerted by one node over another is represented as a directed edge, the influence graphs are usually digraphs, i.e., directed graphs. However, any undirected graph can be seen as a symmetric digraph, so all what follows also applies for undirected graphs.

Given an influence graph and an initial activation set $X \subseteq V$, the *spread of influence* of X is the set $F(X) \subseteq V$ formed by the nodes activated through an iterative process that runs in polynomial time. Lets $F_k(X)$ denote the set of nodes activated at step k. Initially, at step 0, only the nodes in X are activated, so $F_0(X) = X$. The set of nodes activated at step $i > 0$ consists of all nodes of $F_{i-1}(X)$, plus some amount of nodes that depends on the model considered. For the LT-model, we add all the nodes whose labels are less or equal than the total weight of the edges connecting them from nodes in $F_{i-1}(X)$, i.e.,

$$F_i(X) = F_{i-1}(X) \cup \{v \in V \mid \sum_{\{u \in F_{i-1}(X) | (u,v) \in E\}} w(u, v) \geq f(v)\} \qquad (1)$$

The process stops when no additional activation occurs. The final set of activated nodes is denoted by $F(X) = F_k(X)$, where $k = \min\{i \in \mathbb{N} \mid F_i(X) = F_{i+1}(X)\} \leq n$.

Figure 1 illustrates an example of influence spread $F(X)$ in an influence graph from the initial activation $X = \{a\}$. In the first step is obtained $F_1(X) = \{a, c\}$ and in the second step (the last one), $F_2(X) = \{a, c, d\}$.

4. Linear threshold rank for multiple neighborhood levels

In this section we shall define a generalization of the Linear Threshold Rank, also known as LTR measure, based on a flexibilization at the level of neighborhood considered for the initial activation. The LTR measure is a recent centrality measure used to rank influential actors based on the LT-model. It was originally defined as follows.

Definition 1 ([18]). Let (G, w, f) be an influence graph, with $G = (V, E)$, and $i \in V$ an actor. The *Linear Threshold Rank* of i, denoted by $\text{LTR}(i)$, is given by

$$\text{LTR}(i) = |F(\{i\} \cup \text{neighbors}(i))|$$

where $\text{neighbors}(i) = \{j \in V \mid (i, j) \in E \vee (j, i) \in E\}$.

This measure can be normalized by dividing the result by $n = |V|$. Note that G can be either a directed or an undirected graph. However, $\text{neighbors}(i)$ contains the actors connected to i by edges in any direction. This allows to increase the size of the initial activation set, since the actors with small out-degree would not be able to spread their influence through the network by themselves. Furthermore, as $F(X)$ can be computed in polynomial time, $\text{LTR}(i)$ is polynomial time computable.

The Linear Threshold Rank measures how much an actor i can spread his influence within a network, investing resources outside the formal system (e.g. contacts management, bribe ability, etc.) to be able to convince his immediate neighbors. In this measure, a high rank does not depend so much on the nodes degree in the initial activation, but on the influence capacity of the initial activation as a coalition. Now, if the i's neighbors also possess sufficient resources, they could also convince their own neighbors not yet influenced, and so on. This idea leads us to think about a generalization of the LTR measure, for which we need some previous concepts.

Given a graph, a *path* is a finite sequence of distinct nodes such that two consecutive nodes are joined by an edge. An *undirected path* is a path where the edges direction does not matter. The *length* of an undirected path is given by its number of edges. The *shortest undirected path* among two nodes is the undirected path between that nodes with the minimum length. The *diameter* of a graph is the maximum length among all the shortest undirected paths between any two nodes of the graph.

Definition 2. Let (G, w, f) be an influence graph, with $G = (V, E)$, $i \in V$ an actor, and d the diameter of G. The *Linear Threshold Rank* of i at *level* ℓ, with $0 \leq \ell \leq d$, denoted by $\text{LTR}(i, \ell)$, is given by

$$\text{LTR}(i, \ell) = |F(\{i\} \cup \text{neighbors}(i, \ell))|$$

where $\text{neighbors}(i, \ell) = \{j \in V \mid \text{dist}(i, j) \leq \ell\}$ and $\text{dist}(i, j)$ denotes the length of the shortest undirected path between nodes i and j, i.e., the undirected distance between nodes i and j.

Again, this measure can be normalized by dividing the result by $n = |V|$. Note that $\text{LTR}(i, 0) = |F(\{i\})|$, $\text{LTR}(i, 1) = \text{LTR}(i)$, and $\text{LTR}(i, d) = n$. Furthermore, since $\ell \leq d$, then $\text{LTR}(i, \ell)$ is also polynomial time computable. As example, consider the influence graph illustrated in Figure 2. The Linear Threshold Rank of actor a at all levels is presented in Table 1.
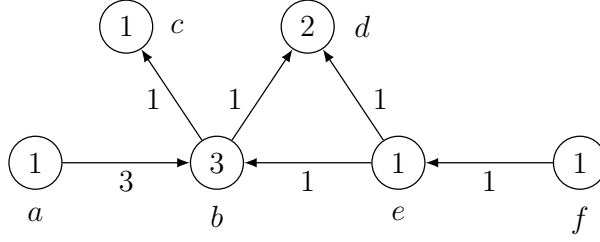
Figure 2: Example of influence graph with $m = 6$, $n = 6$ and $d = 3$.

ℓ	neighbors(a, ℓ)	LTR(a, ℓ)	normalized LTR(a, ℓ)
0	$\{a\}$	3	$3/6 = 0.50$
1	$\{a, b\}$	3	$3/6 = 0.50$
2	$\{a, b, c, d, e\}$	5	$5/6 \approx 0.83$
3	V	6	$6/6 = 1.00$

Table 1: LTR(a, ℓ) computation for the influence graph of Figure 2.

5. Experiments and results

In this section, we analyze the Linear Threshold Rank for each actor and for all levels in two different large networks:

- Retweets network: A directed network.

- arXiv network: An undirected network.

The first network dataset was provided by the Network Repository [40],[1] and the second one by the SNAP's Stanford Large Network Dataset Collection [41].[2] The experiments were programmed with Python 3 programming language. All experiments were run on a machine HP ProLiant DL380p server with two Xeon® E5-2650 CPU.

Every network is represented as an influence graph (G, w, f). Let $G = (V, E)$ be the graph, for each actor $i \in V$ we set the following label functions, usually considered in this kind of experiments [18, 3]:

- Minimum influence: $f(i) = 1$

- Maximum influence: $f(i) = w(i)$

- Simple majority: $f(i) = \lfloor w(i)/2 \rfloor + 1$

- Random: $f(i)$ is a random number such that $w'(i) \leq f(i) \leq w(i) + 1$,

---

where $w'(i) = \min\{w(j,i) \mid (j,i) \in E\}$ but taking $w'(i) = 1$ when the set is empty, and $w(i) = \sum\{w(j,i) \mid (j,i) \in E\}$. The random labels were obtained through the `randint` Python function. For each node i, the label was chosen between $w'(i)$ (i.e., the actor can be influenced by any active neighbor pointing to it) and $w(i) + 1$ (i.e., the actor can not be influenced by anyone).
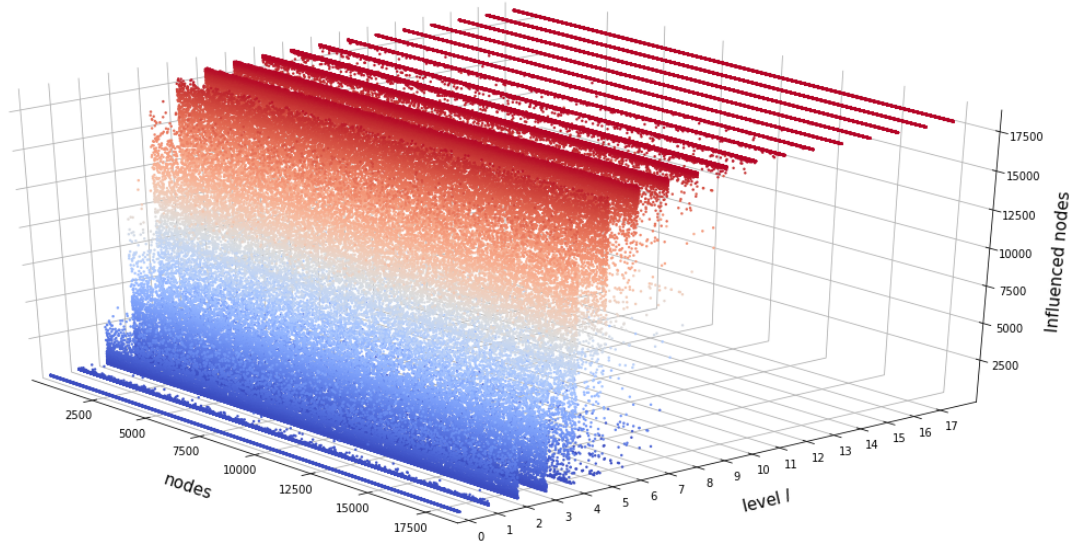
5.1. Retweets network

The first dataset contains the largest connected component of retweets relationships in a Twitter network of political communication. The data were collected during the six weeks prior to the 2010 U.S. midterm elections. It contains $n = |V| = 18,470$ actors (Twitter accounts) and $m = |E| = 48,365$ directed edges (retweets relationships) [42]. A directed edge $(i,j) \in E$ represents an actor j retweeting an actor i, so we say that i exerts a certain influence over j. The weight of the edge, $w(i,j)$, represents how many times actor j retweeted actor i.

Figure 3 analyzes the simple majority case. This figure illustrates the increase in the influence spread for each node i depending on the level ℓ of neighbors that are added to i in the initial activation. The range of colors going from blue to red denotes the number of nodes influenced from the initial activation, that is, the LTR measure $|F(\{i\} \cup \text{neighbors}(i,\ell))|$ (vertical axis of Figure 3(a)). Blue represents a lower capacity for influence spread, while red represents a greater influence capacity. The minimum value is 1 (when $\ell = 0$, and the initial node can not influence to anyone), and the maximum value is n (when the initial activation can spread its influence through the whole network). As we mention in Section 4, this maximum value is always achieved when $\ell = d$, where d is the diameter of the graph (namely, $d = 17$ in this case). However, it can also be achieved for lower neighborhood levels. Figure 3(b) is the top view of the 3D-graphics at Figure 3(a). Hence, the growth curve shown in Figure 3(a) is shown in Figure 3(b) by the "thermometer" on the right.

In this simple majority case, the temperature increases abruptly between levels 3 and 6. Up to level 2, the effort expended to extend the initial activation seems not to make a big difference in the increase of the influence spread. Similarly, starting at level 7, for most actors it no longer makes much sense to continue increasing their initial activation. Remarkably, level 1, used by the original LTR measure [18], does not seem to distinguish so much the actors according to their ability to influence, as it does seem to occur for levels 4 and 5. Indeed, note that the average growth corresponds to a sigmoid function, where levels 4 and 5 are the tipping points [43].

Figure 4 shows the top views for the remaining cases: minimum, random, and maximum influence. From level 11, all the graphics have the same behavior, because the initial activation size is such that the influence spread process ends immediately, i.e., $F_0(X) = F_1(X)$, with X being the initial activation. This does not mean that from level 11 all the initial activation sets already contain all the network nodes, but from that level the initial activation sets are so big that they can influence all the other actors. Table 2 shows, for each level ℓ, the number of nodes having neighbors at that depth level. Observe that all the nodes have neighbors until a depth level $\ell = 9$, while there are just two nodes having neighbors until a depth level $\ell = 17$.

(a) 3D view
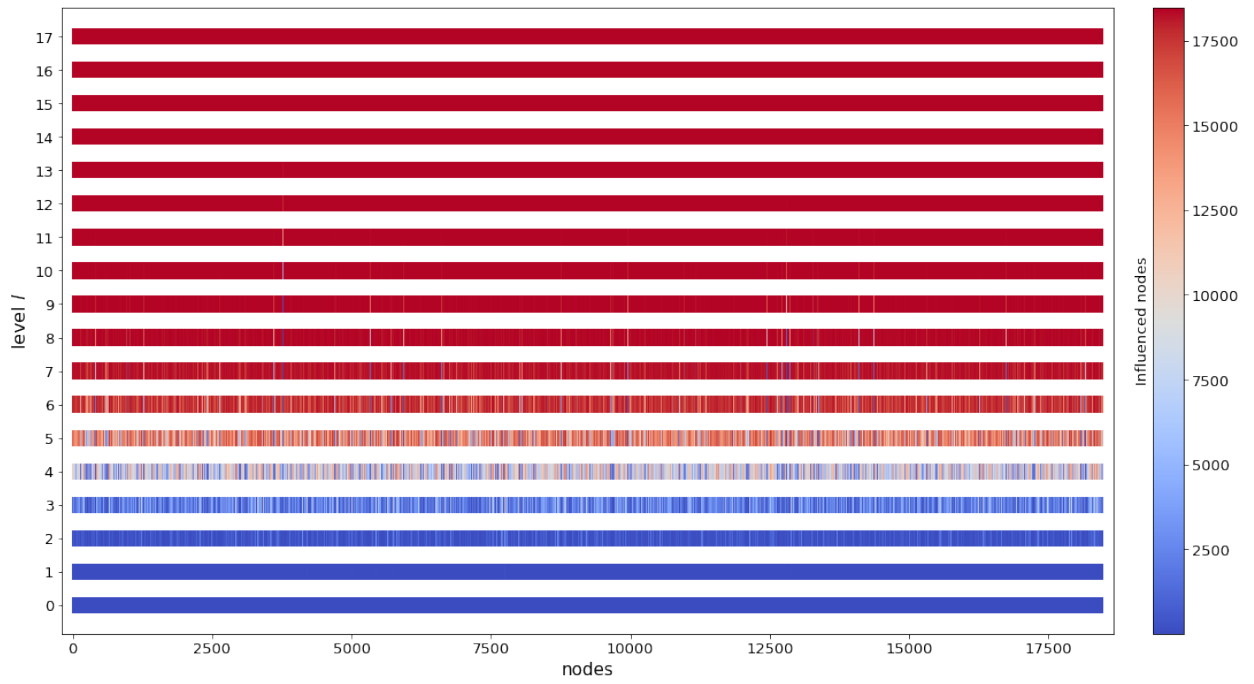
(b) Top view

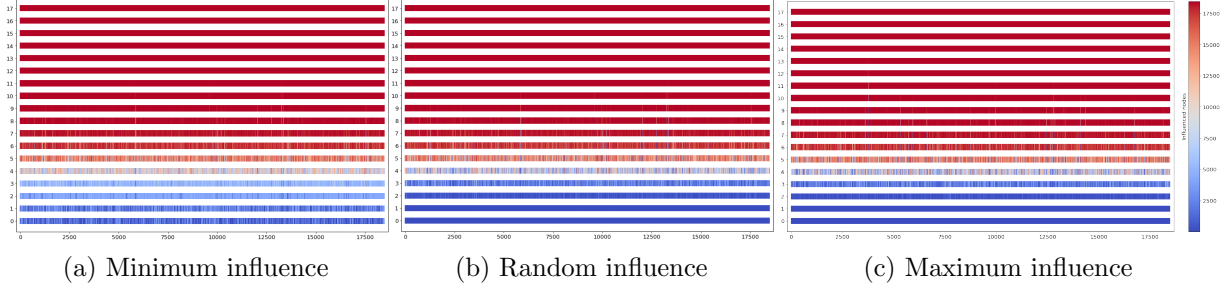Figure 3: Data analysis of retweets network for simple majority.

(a) Minimum influence (b) Random influence (c) Maximum influence

Figure 4: Data analysis of retweets network.

level ℓ	0...9	10	11	12	13	14	15	16	17
#actors	18470	18459	17506	8786	2113	408	105	13	2

Table 2: Number of actors having neighbors at each level in the retweets network.

By using the Kendall correlation coefficient, we corroborate that the initial activation sizes and the LTR results obtained for each actor are highly correlated. This is quite clear, since the larger initial activation size, the greater the influence spread and hence the LTR value. However, this does not mean that the rankings obtained with LTR for a given level, are maintained for the next level. Again, using the Kendall correlation coefficient, we correlate the rankings of actors obtained by the LTR measure for all the levels $0 \leq \ell \leq 17$. The correlation matrices at Figure 5 summarize these latter results for the different labeling functions. Each axis represents the different neighborhood levels ℓ of the LTR measure ($1 \leq \ell \leq 17$). The intensity of each position of the matrix represents the degree of correlation between the rankings obtained with the measure, considering such levels. It is observed that although relatively high correlations are obtained, when considering different levels, $\textsc{ltr}(i, \ell)$ will produce different rankings. Therefore, the generalization of the measure makes sense.

Finally, as it was expected, for the case of minimal influence, the number of influenced nodes increases faster than in the rest of the cases. This can be seen at Figure 4(a), where the lower bars have more "temperature". For these experiments, the case of random influence (see Figure 4(b)) presents an average growth more similar to both the majority and maximum
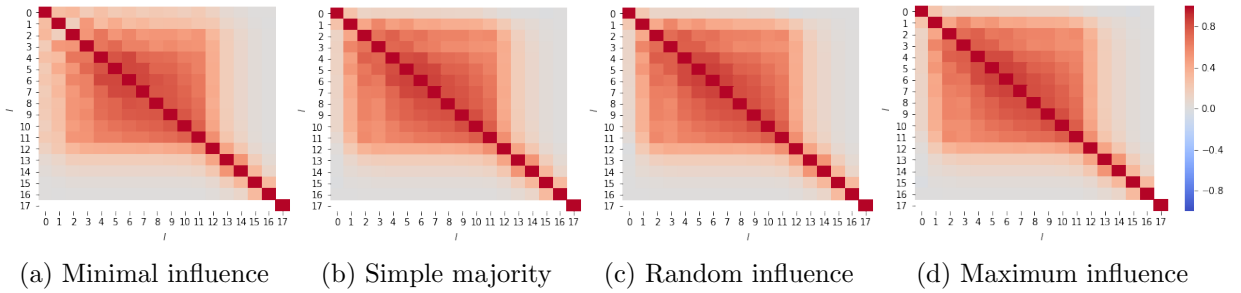

(a) Minimal influence (b) Simple majority (c) Random influence (d) Maximum influence

Figure 5: Kendall correlation matrices for the retweets network.

level ℓ	0...9	10	11	12	13	14	15	16	17
#actors	4158	4145	3683	2207	778	207	66	29	8

Table 3: Number of actors having neighbors at each level in the arXiv network.

cases than to the minimum influence case. In addition, the simple majority case is quite close to the maximum influence case (see Figure 4(c)). The maximum case is the one that presents the smallest increase in the influence spread between one level and the next one. Remarkably, for all the cases, the growth is quite abrupt between levels 3 and 5, thus being levels 4 and 5 the tipping points for all the cases.
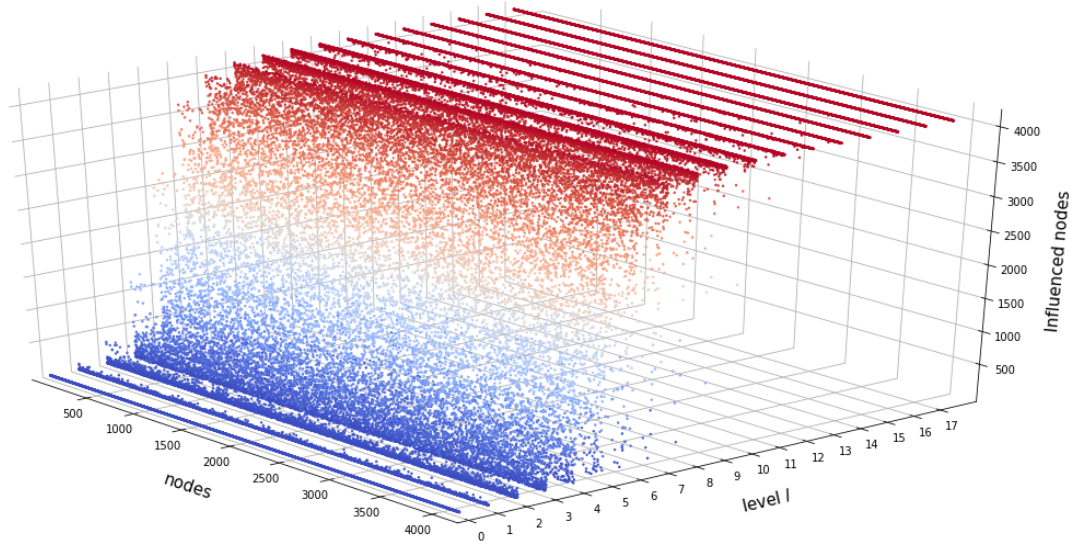
5.2. arXiv network

The second dataset contains scientific collaborations between authors papers submitted to arXiv's General Relativity and Quantum Cosmology category.[3] The collection was initially used to study graph evolution [44]. It contains 5242 nodes and $14,496$ undirected edges. We consider the largest connected component, obtaining an influence graph with $n = 4158$ actors and $m = 13,428$ undirected edges. An edge (i, j) represents that author i co-authored at least one paper with author j, in such a way that, for each edge $(i, j) \in E(G)$, $w(i, j) = 1$.

Analogously to Figure 3, Figure 6 illustrates the increase in the influence spread for each node i based on the level ℓ, for the simple majority case. As the previous network, here all the actors have neighbors until a depth level $\ell = 9$. Furthermore, the maximum neighborhood level is also $\ell = 17$, which is fulfilled for eight actors, instead of the two in the previous network. Analogously to Table 2, Table 3 shows the number of nodes having neighbors at each depth level. Furthermore, analogously to Figure 5, the correlation matrices at Figure 7 shows that despite of the relatively high correlations obtained among the different levels, $\text{LTR}(i, \ell)$ produces different rankings for each level ℓ. Recall that this is an undirected network. Therefore, it is quite natural that for the case of minimal influence, the spread of influence increases to the maximum just by starting from the isolated nodes ($\ell = 0$). The latter implies that for the minimal influence case, all the actors share the same place in the rankings, in such a way that the Kendall correlation coefficient can not distinguish differences between the actors and thus it returns null values, except for the actors compared with themselves (see Figure 7(a)).

In this study case the growth is quite abrupt between levels 4 and 6. Up to level 2, the effort expended to extend the initial activation does not make a big difference in the influence spread increase. Similarly, starting at level 9, for most actors it no longer makes much sense to continue increasing their initial activation. Unlike the previous network, in this case the growth is less abrupt between one level and the next one. Note that level 1, used by the original LTR measure, here it also does not help to properly hierarchize the nodes, as it does seem to occur for levels 4, 5, or 6. In this case, these three levels are the tipping points.

(a) 3D view
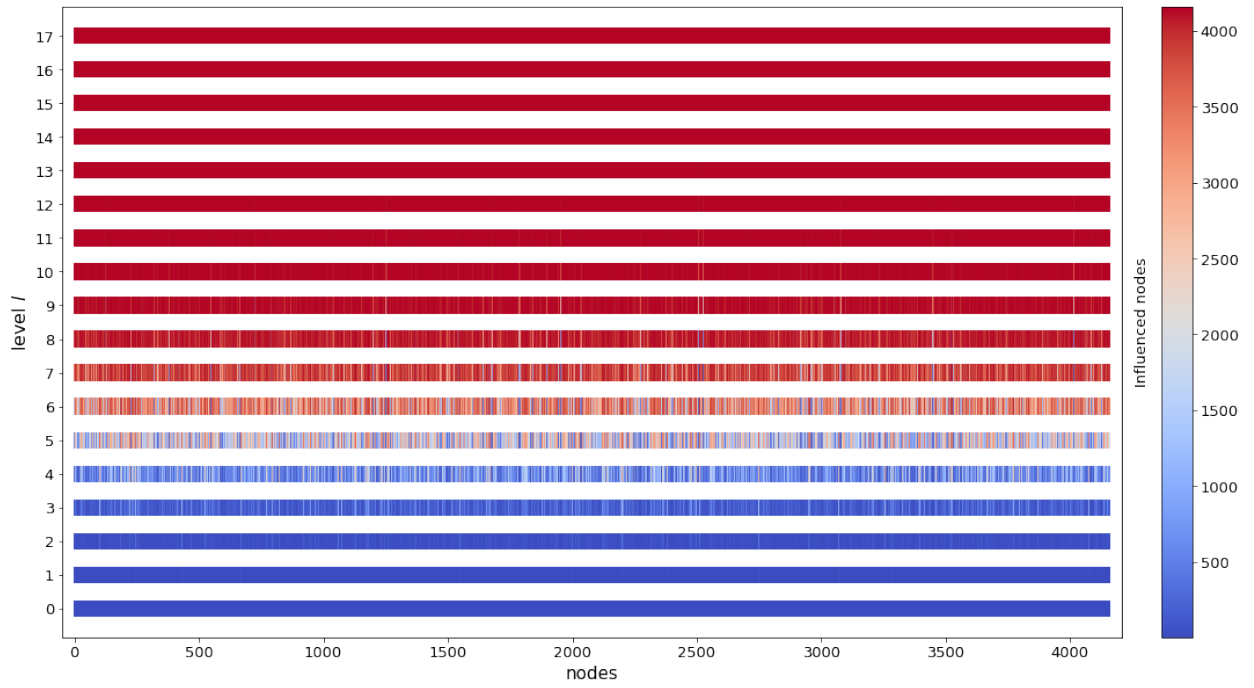
(b) Top view

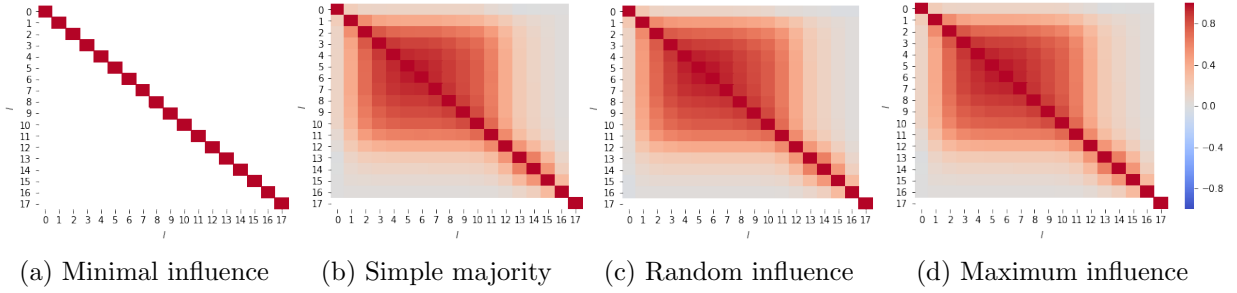Figure 6: Data analysis of arxiv network for simple majority.

(a) Minimal influence   (b) Simple majority   (c) Random influence   (d) Maximum influence

Figure 7: Kendall correlation matrices for the arXiv network.

(a) Minimum influence   (b) Random influence   (c) Maximum influence
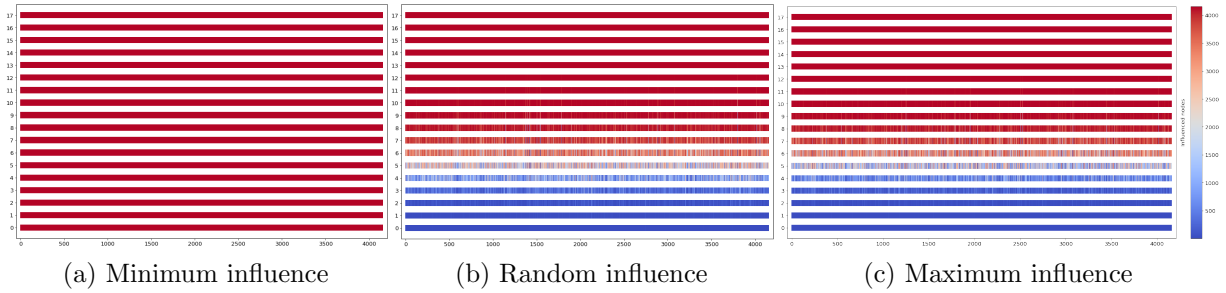
Figure 8: Data analysis of arxiv network.

Analogously to Figure 4, Figure 8 shows the top views for the remaining cases: minimum, random, and maximum influence. Here, from level 12, all the graphics have the same behavior, in the sense that $F_0(X) = F_1(X)$, with X being the initial activation. Again, for the case of minimal influence, as the spread of influence increases to the maximum just by starting from the isolated nodes ($\ell = 0$), then all the levels have the maximum "temperature" (see Figure 8(a)). Here the case of random influence (Figure 8(b)) also presents a relatively similar growth to the cases of simple majority and maximum influence. Furthermore, the simple majority case is also quite close to the maximum influence case (see Figure 8(c)). The case of minimal influence does not have tipping points, because all the levels have the same temperature.

6. Conclusions and future work

In this paper we have proposed a family of centrality measures based on a extension of the Linear Threshold Rank (LTR) [18], a novel measure used to rank influential actors within a social network. Recall that in a social network represented as an influence graph, the original LTR measure only considers one depth level of neighbors in the initial activation set. In our generalization of the LTR measure, the depth level of the neighbors in the initial activation set is as a key parameter. We have been able to compare the impact of the initial neighborhood in the ability of the LTR to distinguish influential actors, providing an answer to the future works raised in previous research [18].

In previous work, it was shown that the LTR measure can allow to obtain more distinguishable ranking values than other known influential measures like PageRank, Katz centrality, or Independent Cascade Rank [18]. The proposed generalization of the LTR measure was analyzed on two different case studies: a directed retweets network, and an undirected arXiv bibliometric network. For both case studies, we show that we can obtain even more distinguishable ranking values by using higher levels of neighbors. For the first case study (retweets network), tipping points was given for levels 4 and 5, regardless of the labeling function. For the second case study (arXiv network), tipping points were detected for almost all the cases in levels 4, 5, and 6.

These tipping points correspond to the levels for which a greater differentiation on the influence spread process is obtained, and therefore in the value of the LTR measure. Moreover, the average growth of the influence spread in terms of the level of neighbors fits largely to a sigmoid function. Tipping points have been studied in depth in this family of functions [43].

In addition, the tipping points seem to have to do with the network centralization. It is left as future work to extract from this behavior a centralization measure. Note that centralization measures return a unique value for the whole network, instead of the centrality measures, which are local measures that return one value for each actor. This would allow to expand the studies related to centralization in influence graphs, a subject still little studied. In particular, it would be interesting to analyze how the Linear Threshold Centralization (LTC) measure proposed in [18] can be generalized using this approach. In particular whether extending the main core by additional levels of neighbors, we can provide a more sensitive centralization measure to quantify the global influence spread.

Finally, it is very interesting to note the similarities between the increase in the generalized LTR measure and the Bass diffusion model. The Bass diffusion model [45, 46], coming from marketing science, is a differential equation defined to describe the process of how new products get adopted in a population. The increase in "adopters" (influenced actors) over time corresponds to a sigmoid function, very similar to those we have obtained in this study. The Bass diffusion model considers two parameters: a coefficient of innovation, that represents the external influence or advertising effect, and the coefficient of imitation, which is the internal influence or word-of-mouth effect. The main difference between both models is that the Bass diffusion model analyzes the increase of adopters over time, while in this article we have studied the total increase of adopters in terms of the initial number of active neighbors. Indeed, in the original idea of the LTR measure [18], in order to activate the initial neighbors, an external effort to the formal network is required. This effort can be related with the one considered by the Bass diffusion model's coefficient of innovation. Thus, the relationship between both models becomes clearer, insofar as we can consider, within this initial effort, the time necessary to convince the neighbors to support the initial node. This relationship between both models could be studied in more detail, as another line of future work.

Acknowledgments

References

[1] D. A. Easley, J. M. Kleinberg, Networks, crowds, and markets - Reasoning about a highly connected world, Cambridge University Press, 2010.

[2] J. Sun, J. Tang, A survey of models and algorithms for social influence analysis, in: C. C. Aggarwal (Ed.), Social Network Data Analytics, Springer, 2011, pp. 177–214.

[3] X. Molinero, F. Riquelme, M. J. Serna, Cooperation through social influence, European Journal of Operational Research 242 (3) (2015) 960–974.

[4] D. Kempe, J. M. Kleinberg, É. Tardos, Maximizing the spread of influence through a social network, Theory of Computing 11 (2015) 105–147.

[5] M. Granovetter, Threshold models of collective behavior, American Journal of Sociology 83 (6) (1978) 1420–1443.

[6] T. Schelling, Micromotives and macrobehavior, Norton, 1978.

[7] J. Goldenberg, B. Libai, E. Muller, Using complex systems analysis to advance marketing theory development, Tech. rep., Academy of Marketing Science Review (2001).

[8] P. M. Domingos, M. Richardson, Mining the network value of customers, in: D. Lee, M. Schkolnick, F. J. Provost, R. Srikant (Eds.), Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, San Francisco, CA, USA, August 26-29, 2001, ACM, 2001, pp. 57–66.

[9] D. Gruhl, D. Liben-Nowell, R. V. Guha, A. Tomkins, Information diffusion through blogspace, SIGKDD Explorations 6 (2) (2004) 43–52.

[10] L. A. Adamic, E. Adar, How to search a social network, Social Networks 27 (3) (2005) 187–203.

[11] X. Song, B. L. Tseng, C. Lin, M. Sun, Personalized recommendation driven by information flow, in: E. N. Efthimiadis, S. T. Dumais, D. Hawking, K. Järvelin (Eds.), SIGIR 2006: Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Seattle, Washington, USA, August 6-11, 2006, ACM, 2006, pp. 509–516.

[12] X. Zhang, J. Zhu, Q. Wang, H. Zhao, Identifying influential nodes in complex networks with community structure, Knowl.-Based Syst. 42 (2013) 74–84.

[13] J. Li, W. Peng, T. Li, T. Sun, Q. Li, J. Xu, Social network user influence sense-making and dynamics prediction, Expert Syst. Appl. 41 (11) (2014) 5115–5124.

[14] F. Morone, H. A. Makse, Influence maximization in complex networks through optimal percolation, Nature 524 (2015) 64–68.

[15] R. Noël, F. Riquelme, R. Mac Lean, E. Merino, C. Cechinel, T. S. Barcelos, R. Villarroel, R. Muñoz, Exploring collaborative writing of user stories with multimodal learning analytics: A case study on a software engineering course, IEEE Access 6 (2018) 67783–67798.

[16] F. Riquelme, P. Gonzalez-Cantergiani, Measuring user influence on Twitter: A survey, Inf. Process. Manage. 52 (5) (2016) 949–975.

[17] Z.-Y. Wang, J.-T. Han, J. Zhao, Identifying node spreading influence for tunable clustering coefficient networks, Physica A: Statistical Mechanics and its Applications 486 (2017) 242 – 250.

[18] F. Riquelme, P. G. Cantergiani, X. Molinero, M. J. Serna, Centrality measure in social networks based on linear threshold model, Knowledge-Based Systems 140 (2018) 92–102.

[19] L. Page, S. Brin, R. Motwani, T. Winograd, The PageRank citation ranking: Bringing order to the web, Tech. rep., Stanford Digital Library (1999).

[20] L. Katz, A new status index derived from sociometric analysis, Psychometrika 18 (1) (1953) 39–43.

[21] D. Kempe, J. M. Kleinberg, É. Tardos, Influential nodes in a diffusion model for social networks, in: L. Caires, G. F. Italiano, L. Monteiro, C. Palamidessi, M. Yung (Eds.), Automata, Languages and Programming, 32nd International Colloquium, ICALP 2005, Lisbon, Portugal, July 11-15, 2005, Proceedings, Vol. 3580 of Lecture Notes in Computer Science, Springer, 2005, pp. 1127–1138.

[22] L. Freeman, Centrality in social networks: Conceptual clarification, Social Networks 1 (3) (1979) 215–239.

[23] D. Gómez, J. R. Figueira, A. Eusébio, Modeling centrality measures in social network analysis using bi-criteria network flow optimization problems, European Journal of Operational Research 226 (2) (2013) 354–365.

[24] M. Montangero, M. Furini, Trank: Ranking twitter users according to specific topics, in: 12th Annual IEEE Consumer Communications and Networking Conference, CCNC 2015, Las Vegas, NV, USA, January 9-12, 2015, IEEE, 2015, pp. 767–772.

[25] G. Katsimpras, D. Vogiatzis, G. Paliouras, Determining influential users with supervised random walks, in: A. Gangemi, S. Leonardi, A. Panconesi (Eds.), Proceedings of the 24th International Conference on World Wide Web Companion, WWW 2015, Florence, Italy, May 18-22, 2015 - Companion Volume, ACM, 2015, pp. 787–792.

[26] M. Bouguessa, L. B. Romdhane, Identifying authorities in online communities, ACM TIST 6 (3) (2015) 30:1–30:23.

[27] M. del Pozo, C. Manuel, E. González-Arangüena, G. Owen, Centrality in directed social networks. A game theoretic approach, Social Networks 33 (3) (2011) 191–200.

[28] M. T. Irfan, L. E. Ortiz, On influence, stable behavior, and the most influential individuals in networks: A game-theoretic approach, Artif. Intell. 215 (2014) 79–119.

[29] X. Molinero, F. Riquelme, M. J. Serna, Power indices of influence games and new centrality measures for agent societies and social networks, in: C. Ramos, P. Novais, C. E. Nihan, J. M. C. Rodríguez (Eds.), Ambient Intelligence - Software and Applications - 5th International Symposium on Ambient Intelligence, ISAmI 2014, Salamanca, Spain, June 4-6, 2014, Vol. 291 of Advances in Intelligent Systems and Computing, Springer, 2014, pp. 23–30.

[30] D. Gayo-Avello, D. J. Brenes, D. Fernández-Fernández, M. E. Fernández-Fernández, R. García-Suárez, De retibus socialibus et legibus momenti, EPL (Europhysics Letters) 94 (3) (2011) 38001.

[31] A. Pal, S. Counts, Identifying topical authorities in microblogs, in: I. King, W. Nejdl, H. Li (Eds.), Proceedings of the Forth International Conference on Web Search and Web Data Mining, WSDM 2011, Hong Kong, China, February 9-12, 2011, ACM, 2011, pp. 45–54.

[32] M. S. Srinivasan, S. Srinivasa, S. Thulasidasan, Exploring celebrity dynamics on twitter, in: Proceedings of the 5th IBM Collaborative Academia Research Exchange Workshop, I-CARE '13, ACM, New York, NY, USA, 2013, pp. 13:1–13:4. doi:10.1145/2528228.2528242.

[33] I. Anger, C. Kittl, Measuring influence on twitter, in: S. N. Lindstaedt, M. Granitzer (Eds.), I-KNOW 2011, 11th International Conference on Knowledge Management and Knowledge Technologies, Graz, Austria, September 7-9, 2011, ACM, 2011, p. 31.

[34] K. Das, S. Samanta, M. Pal, Study on centrality measures in social networks: a survey, Social Netw. Analys. Mining 8 (1) (2018) 13.

[35] S. K. Kundu, C. A. Murthy, S. K. Pal, A new centrality measure for influence maximization in social networks, in: S. O. Kuznetsov, D. P. Mandal, M. K. Kundu, S. K. Pal (Eds.), Pattern Recognition and Machine Intelligence - 4th International Conference, PReMI 2011, Moscow, Russia, June 27 - July 1, 2011. Proceedings, Vol. 6744 of Lecture Notes in Computer Science, Springer, 2011, pp. 242–247.

[36] S. K. Pal, S. K. Kundu, C. A. Murthy, Centrality measures, upper bound, and influence maximization in large scale directed social networks, Fundam. Inform. 130 (3) (2014) 317–342.

[37] I. Gaye, G. Mendy, S. Ouya, D. Seck, New centrality measure in social networks based on independent cascade (IC) model, in: I. Awan, M. Younas, M. Mecella (Eds.), 3rd International Conference on Future Internet of Things and Cloud, FiCloud 2015, Rome, Italy, August 24-26, 2015, IEEE Computer Society, 2015, pp. 675–680.

[38] H. Wu, J. Shang, S. Zhou, J. Zhong, Y. Feng, B. Qiang, IMPC: Influence maximization based on multi-neighbor potential in community networks, Physica A: Statistical Mechanics and its Applications 512 (2018) 1085–1103.

[39] W. Zhu, W. Yang, S. Xuan, D. Man, W. Wang, X. Du, M. Guizani, Location-based seeds selection for influence blocking maximization in social networks, IEEE Access 7 (2019) 27272–27287.

[40] R. A. Rossi, N. K. Ahmed, The network data repository with interactive graph analytics and visualization, in: B. Bonet, S. Koenig (Eds.), Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA., AAAI Press, 2015, pp. 4292–4293.

[41] J. Leskovec, A. Krevl, SNAP Datasets: Stanford large network dataset collection, `http://snap.stanford.edu/data` (Jan. 2017).

[42] M. Conover, J. Ratkiewicz, M. R. Francisco, B. Gonçalves, F. Menczer, A. Flammini, Political polarization on Twitter, in: L. A. Adamic, R. A. Baeza-Yates, S. Counts (Eds.), Proceedings of the Fifth International Conference on Weblogs and Social Media, Barcelona, Catalonia, Spain, July 17-21, 2011, The AAAI Press, 2011, pp. 89–96.

[43] F. Phillips, On s-curves and tipping points, Technological Forecasting and Social Change 74 (6) (2007) 715–730.

[44] J. Leskovec, J. M. Kleinberg, C. Faloutsos, Graph evolution: Densification and shrinking diameters, TKDD 1 (1) (2007) 2.

[45] F. M. Bass, A new product growth for model consumer durables, Management Science 15 (5) (1969) 215–227.

[46] F. M. Bass, Comments on "A new product growth for model consumer durables the Bass model", Management Science 50 (12-Supplement) (2004) 1833–1840.