

A Visual Analysis of Social Influencers and Influence in the Tourism Domain

Chiara Francalanci and Ajaz Hussain

Abstract Identifying influencers is an important step towards understanding how information spreads within a network. In social media, hub nodes are generally considered as social influencers. Social networks follow a power-law degree distribution of nodes, with a few hub nodes and a long tail of peripheral nodes. While there exist consolidated approaches supporting the identification and characterization of hub nodes, research on the analysis of the multi-layered distribution of peripheral nodes is limited. However, influence seems to spread following multi-hop paths across nodes in peripheral network layers. This paper proposes a visual approach to the graphical representation and exploration of peripheral layers by exploiting the theory of k-shell decomposition analysis. We put forward three hypotheses that allow the graphical identification of peripheral nodes that are more likely to be influential and contribute to the spread of information. Hypotheses are tested on a large sample of tweets from the tourism domain.

Keywords Social media • Influence • Influencers • Power law graphs

1 Introduction

The literature on social media makes a distinction between influencers and influence. The former are social media users with a broad audience, while the latter is instead used to refer to the social impact of the content shared by social media users. In Boyd et al. (2010) and Myers and Leskovec (2014) authors note that a content that has had an impact on a user's mind is shared. Influencers are prominent social media users, but we cannot expect that the content that they share is bound to have high influence, as discussed by Benevenuto et al. (2010) and Messias et al. (2013). Previous research, (Bruni et al. 2013; Klotz et al. 2014) has shown how the content of messages can play a critical role and can be a determinant of the social influence of a message irrespective of the centrality of the message's author. This paper starts from the observation made by Chan et al. (2003) stating that social networks of

C. Francalanci (✉) • A. Hussain

Department of Electronics, Information and Bio-Engineering, Politecnico di Milano, Milan, Italy

e-mail: francala@elet.polimi.it; hussain@elet.polimi.it

influence follow a power-law distribution function, with a few hub nodes and a long tail of peripheral nodes, consistent with the so-called small-world phenomenon as noted by Xu et al. (2007). In social media, hub nodes represent social influencers (Ren et al. 2014), but influential content can be generated by peripheral nodes and spread along possibly multi-hop paths originated in peripheral network layers. In this paper, we exploit a modified power-law based force-directed algorithm (Francalanci and Hussain 2014; Hussain et al. 2014) to highlight the local multi-layered neighbourhood clusters around hub nodes. In our approach, the topology of the periphery is defined by grouping peripheral nodes based on the strength of their link to hub nodes, according to the metaphor of k-shell decomposition analysis (Carmi et al. 2007; Kitsak et al. 2010).

The approach is tested on a large sample of tweets expressing opinions on a selection of Italian locations relevant to the tourism domain. By visually exploring and understanding the multi-layered periphery of nodes, we propose three content related hypotheses exploring the role of peripheral nodes. Empirical and visual results show that peripheral nodes play a role as determinant of the social influence. The main innovative aspect of our approach is that we show our hypotheses visually to understand the practical meaning of our hypotheses.

2 State of the Art

Several research efforts in network visualization have targeted power-law algorithms and their combination with the traditional force-directed techniques, as for example in Andersen et al. (2004, 2007), Boutin et al. (2006), and Chen (2006). Among these approaches, the most notable is the Out-Degree Layout (ODL) for the visualization of large-scale network topologies, presented by Chan et al. (2003) and Perline (2005). The core concept of the algorithm is the segmentation of network nodes into multiple layers based on their out-degree, i.e. the number of outgoing edges of each node. The positioning of network nodes starts from those with the highest out-degree, under the assumption that nodes with a lower out-degree have a lower impact on visual effectiveness. The topology of the network plays an important role such that there are plausible circumstances under which nodes with a higher number of connections or greater betweenness have little effect on the range of a given spreading process (Cha et al. 2010).

Centrality metrics are the most widely used parameters for the structural evaluation of a user's social network. The concept of centrality has been defined as the importance of an individual within a network (Fan and Gordon 2014). A node that is directly connected to a high number of other nodes is obviously central to the network and likely to play an important role (Barbagallo et al. 2012; Sparrowe et al. 2001). The more recent literature has associated the complexity of the concept of influence with the diversity of content. Several research works have addressed the need for considering content-based metrics of influence (Bakshy et al. 2011; Bigonha et al. 2012; Hossain et al. 2006; Li et al. 2014; Naaman et al. 2010).

Clearly, this view involves a significant change in perspective, as assessing influence does not provide a static and general ranking of influencers as a result. However, there is a need for effective visualization techniques in social networks, which enable users to visually explore scalable complex social networks to identify the influencers who are responsible for influence spread.

3 The Power-Law Algorithm

An early version of the algorithm has been presented by Francalanci and Hussain (2014) and Hussain et al. (2014). This paper improves the initial algorithm by identifying multiple layers of peripheral nodes around hub nodes as per k-shell decomposition approach. The power-law layout algorithm, shown in following code snippet, belongs to the class of force-directed algorithms, such as the one by Chan et al. (2003) and Fruchterman and Reingold (1991).

```
begin
  NodeCharacterization();
  InitialLayout();
  while Temperature > 0 do
    if Temperature > Th then
      call NodePlacement(Nh, Eh);
    else
      call NodePlacement(Np, Ep);
    end
    call TemperatureCooldown(Temperature);
  end
end
```

We partition the set of nodes N into the set of hub nodes N_h and the set of peripheral nodes N_p , such that $N = N_h \cup N_p$, with $N_h \cap N_p = \emptyset$. As a consequence, the set of edges E is also partitioned in the set of edges E_h for which at least one of the two nodes is a hub node, and the set E_p which contains all the edges connecting only peripheral nodes, with $E = E_h \cup E_p$, with $E_h \cap E_p = \emptyset$. The distinction of a node n as a hub node or peripheral node is based on the evaluation of its degree $\rho(n)$ against the constant ρ_h , which is a threshold defined as the value of degree that identifies the top i th percentile of nodes, sorted by decreasing value of degree. Since the power-law is supposed to hold in the degree distribution, we have assumed $i = 20$ and consequently ρ_h as the 20th percentile, thus considering as hub nodes the 20 % of the nodes with the highest values of degree—the Pareto's 80-20 Rule, as suggested by Koch (1999).

The `NodeCharacterization()` step is a pre-processing phase aimed at distinguishing hub nodes from peripheral nodes, so that in the following steps

it is possible to leverage the power-law distribution of nodes and assigning the level value (l_s) using k-shell decomposition analysis technique. In this paper, this step is performed by pre-identifying hub nodes as N_p , which represent either the predefined 7 brands or the 12 subjects of interest for the community, which contain set of categories of content referring to specific brand drivers of a destination's brand explained in Sect. 5. At first the NodeCharacterization() method builds local neighbourhood multi-clusters by taking placing these predefined hub nodes central to each cluster by using modified force directed algorithm and power-law based degree distribution. Later on, to create multi-layered periphery around each cluster, we apply l -shell decomposition analysis technique. The InitialLayout() step provides the initial placement of nodes (either a random placement or the result of another graph layout algorithm). The NodePlacement (N, E) step performs the placement of nodes based on the computation of forces among nodes; its inputs are a node set N and an edge set E , such that the placement of nodes can be selectively applied to chosen subsets of nodes/edges at each step. The TemperatureCooldown() step is responsible for the control of the overall iteration mechanism.

We tuned this technique by means of the metaphor of k-shell decomposition analysis (Abello and Queyroi 2013; Alvarez-Hamelin et al. 2006; Carmi et al. 2007; Kitsak et al. 2010), in order to define the concept of *level* of each node in the multi-layered periphery of our graphs. This process assigns an integer as level index (l_s) to each node, representing its location according to successive layers (l shells) in the network. The inner-most layer around cluster hub, will have highest l_s value, containing those authors, who tweeted most about that topic (cluster hub). So, by this metaphor, small values of (l_s) define the periphery of the network (outliers), while the innermost network levels correspond to greater values of l_s , as shown in Fig. 1.

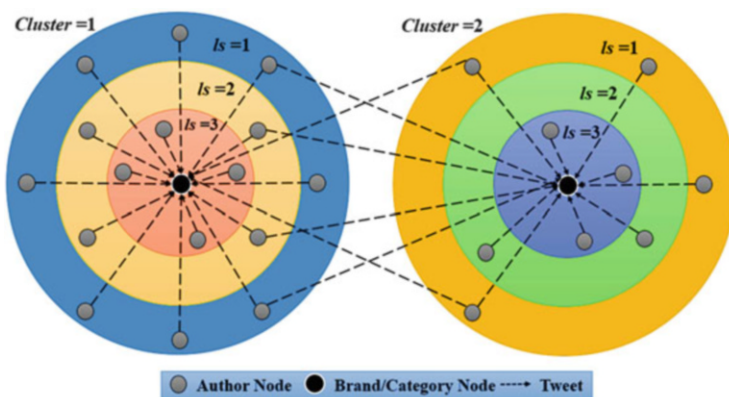


Fig. 1 Metaphor of k-shell decomposition analysis

4 Research Hypotheses

Previous research indicates that social media are associated with a long-tail effect (Meraz 2009; Myers and Leskovec 2014). The long-tail effect suggests that small communities are numerous and their specific interests are virtually boundless (Fan and Gordon 2014). Bruni et al. (2013) and Klotz et al. (2014) have shown how the content of messages can play a critical role and can be a determinant of the social influence of a message irrespective of the centrality of the message's author. Twitter users with a high volume of tweets can be referred to as '*information sources*' or '*generators*' (Hutto et al. 2013). Moreover, social media users intend to post content that is shared frequently by many other users (Asur et al. 2011; Li et al. 2014). Social media users wish to be influential (Myers and Leskovec 2014). Intuitively, since users want to be interesting to many, they post frequently and at the same time they will address the needs of multiple specific communities, multiple topics. Consequently, our first hypothesis posits a positive association among *frequency of tweets* and *content specificity* in multiple topics.

- H1: *Authors tweeting with a high frequency of tweets is positively associated with multiple topics (brands or categories)* (i.e. visually, potential influencers are peripheral authors).

The literature indicates that retweeting is associated with information sharing, commenting or agreeing on other peoples' messages and entertaining followers (Boyd et al. 2010). Kwak et al. (2010) also show that, the most trending topics have an active period of 1 week, while half of retweets of a given tweet occur within 1 h and 75 % within 1 day. The frequency of retweets can be an important criterion since users tend to retweet valuable posts (Myers and Leskovec 2014). Intuitively, if a user tweets about multiple topics, he/she is more likely to be interesting to many specific and active communities and as a consequence, he/she is more likely to obtain more retweets. In the following hypothesis, we posit a positive association between the *content specificity* and *frequency of retweets*.

- H2: *Tweeting about multiple topics (brands or categories) is positively associated with the frequency of retweets* (i.e. visually, peripheral authors, connected to multiple topics, are *actual influencers*).

Traditional media are based on broadcasting rather than communication, while social media are truly interactive (Benevenuto et al. 2010). In traditional media, the influencers intend to target a large audience by broadcasting and talking frequently. Similarly, in social media, e.g. in twitter, influencers intend to be more interactive by participating in the conversation with a variety of mechanisms and, most commonly, by frequently sharing the content that they have liked (Barbagallo et al. 2012; Bruni et al. 2013; Ren et al. 2014). In Leavitt et al. (2009) and Myers and Leskovec (2014), authors show that level of users' activity (number of tweets) depends upon retweets and their in-degree centrality (number of followers). In social media, while sharing content, users may be referred as '*generalist*' or

‘*information sources*’ who talk about multiple topics (Hutto et al. 2013). On the contrary, there exist users who are very specific in sharing content related to specific topics or brands. These authors seem to be potential influence spreaders (Fan and Gordon 2014). They are very likely to be active participants in each community by talking a lot. Our third hypothesis posits that such nodes have a greater probability of being retweeted, and can be both potential and actual influencers.

- H3: *Tweeting more frequently (with a high frequency) about a single topic (brand or category) is positively associated with the frequency of retweets* (i.e. visually, authors, drawn closer to single topic, are both *actual* and *potential* influencers).

We posit the aforementioned three hypotheses that tie *content specificity*, *frequency of tweets* and *frequency of retweets*. Hypothesis H1 can be visualized by observing the peripheral authors positioned in the outer-most layers of each cluster (lowest l-shell value, $l_s = 1$), which are only connected to one cluster hub (brand or category). Such outlier authors can be *potential* influencers, if they further connect to other authors via content sharing and tweeting about multiple topics (brands or categories). Similarly, hypothesis H2 can be visually verified by observing those authors who are placed in between multiple clusters, connected to multiple clusters’ hubs (brands or categories), and accordingly talk about multiple topics. These authors are *actual* influencers as they receive a high number of retweets by tweeting about multiple topics. Moreover, hypothesis H3 can be visualized by observing authors who are positioned in the inner-most periphery of each cluster (highest l_s value), and seem to be placed close to the cluster hub (brand or category). Such authors are both *actual* and *potential* influencers as they are most specific about content sharing.

5 Experimental Methodology and Results

This section reports the discussion about the dataset and the network models used in our experiment. The obtained visualization results and proposed hypotheses are empirically evaluated in this section.

5.1 Variable Definition

Each graph $G(A, T)$ has a node set A representing authors and an edge set T representing tweets. We define as $N_T(a)$ the total number of tweets posted by author a . We define as $N_R(a)$ total number of times author a , has been retweeted. Tweets can refer to a brand b or to a category c . We define as $N_B(a)$ the total number of brands mentioned by each author a , in all his/her tweets, i.e. *brand*

specificity. Similarly, $N_C(a)$ represents the total number of categories mentioned by each author a , in all his/her tweets, i.e. *category specificity*.

5.2 Data Sample and Network Models

We collected a sample of tweets over a 2-month period (December 2012–January 2013). For the collection of tweets, we queried the public Twitter APIs by means of an automated collection tool developed ad-hoc. Twitter APIs have been queried with the crawling keywords, representing tourism destinations (i.e. brands). Two languages have been considered, *English* and *Italian*. Collected tweets have been first analysed with a proprietary semantic engine (Barbagallo et al. 2012; Bruni et al. 2013) in order to tag each tweet with information about (a) the location to which it refers, (b) the location’s brand driver (or category) on which authors express an opinion, (c) the number of retweets (if any), and (d) the identifier of the retweeting author. Our data sample is referred to the tourism domain. We have adopted a modified version of the Anholt Nation Brand index model to define a set of categories of content referring to specific brand drivers of a destination’s brand (Anholt 2006). Table 1 refer to the descriptive statistics of the original non-linear variables.

In order to verify the effectiveness of the proposed algorithm with respect to the goal of our research, we have defined two different network models based on the data set.

- Author \rightarrow Brand (N_1). The network is modelled as an undirected affiliation two-mode network, where an author node n_a is connected to a brand node n_b whenever author a has mentioned brand b in at least one of his/her tweets.
- Author \rightarrow Category (N_2). The network is modelled as an undirected affiliation two-mode network, where an author node n_a is connected to a category node n_c whenever author a has mentioned a subject belonging to category c in at least one of his/her tweets.

5.3 Network Visualization

Table 2 provides descriptive statistics on the size of the N_1 and N_2 networks, as discussed in Sect. 5.2. The empirical results and discussions on network visualization will adopt network N_1 network (i.e. Author \rightarrow Brand) as reference example.

Figure 2 provides an enlarged view of network N_1 visualized by means of the proposed power-law layout algorithm. The network visualization depicted in Fig. 2 adopts multicolour nodes to represent authors, and highlighted encircled blue (dark) nodes to represent the tourism destinations (i.e. brands) on which authors have expressed opinions in their tweets.

Table 1 Basic descriptive statistics of our dataset

Variable	Value
Number of tweets	957,632
Number of retweeted tweets	79,691
Number of tweeting authors	52,175
Number of retweets	235,790
Number of retweeting authors	66,227

Table 2 Descriptive statistics on the dimensions of N_1 and N_2 networks

Authors	N_R (a)	N_1	N_2	N_T (a)
		N_B (a)	N_C (a)	
398	92	856	1,913	2,769
1,662	364	2,905	5,959	8,864
10,710	2,907	12,559	18,498	31,057
18,711	5,329	21,140	29,842	50,982
30,310	8,690	33,684	46,120	79,804
37,626	10,529	41,620	56,960	98,580
47,295	12,833	52,208	71,667	1,23,875

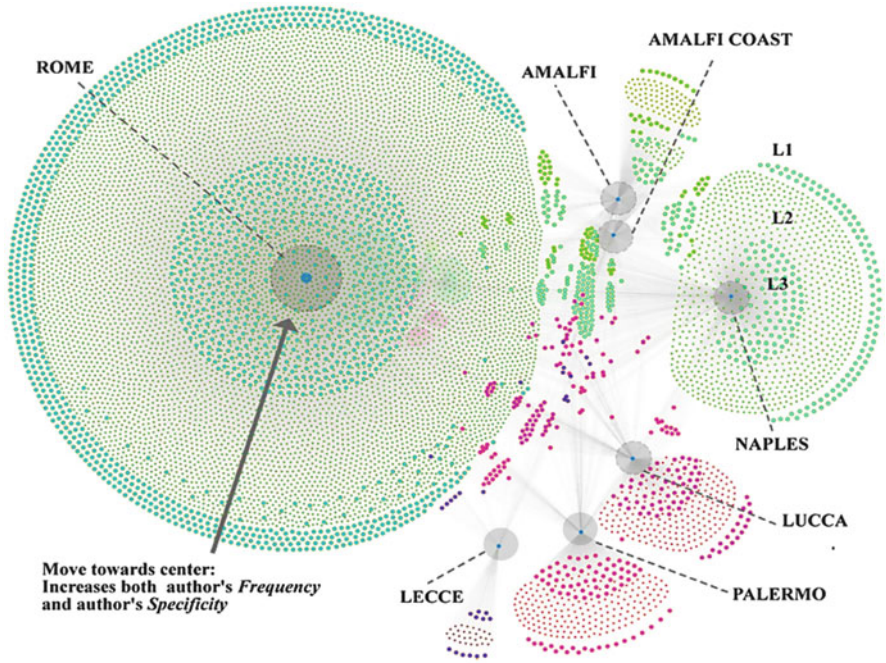
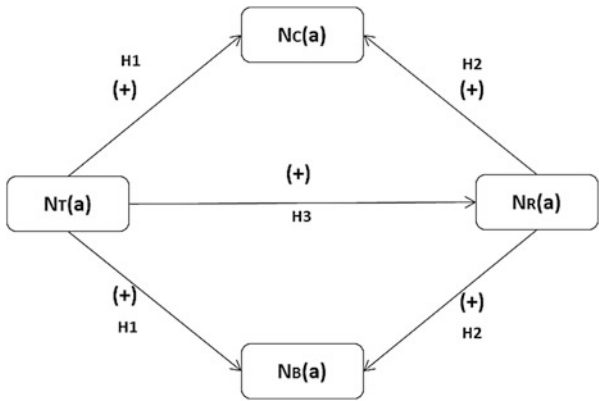


Fig. 2 Network N_1 : Author \rightarrow Brand (enlarged view)

Fig. 3 Research model



5.4 Empirical Results

AMOS 20 (Arbuckle 2011) has been used to analyse the research model by means of structural equation modelling (SEM) (Bagozzi and Fornell 1982). All statistical analyses have been performed with SPSS 20 (Pallant 2010). The research model used for estimation analysis is shown in Fig. 3.

Table 3 presents the correlation matrix of our data variables. Table 4 follows that correlation is significant at 0.01 level (2-tailed). All persistence variables are positively correlated with each other, and thus have a significant impact upon each other.

The regression estimation results of the research model are shown in Table 4. All relationships between persistence metrics are significant, with $p < 0.001$.

Hypothesis H1 Hypothesis H1 has been tested through correlation. By Table 3, both $N_c(a)$ and $N_b(a)$ have a positive correlation of 0.898 and 0.590, respectively with $N_t(a)$, at 0.01 level of significance, supporting the hypothesis H1. Authors by having greater probability of sharing contents, can be *potential influencers* in network. Similarly, through visualization results, Fig. 2 highlights clusters that group all the authors who tweeted about 7 distinct brands, in which ‘ROME’ and ‘NAPLES’ are seem to be mostly tweeted by authors i.e. they possess ‘*high specificity*’, and the *peripheral* authors (visually drawn in outmost peripheries, lowest *l*-shell value), can be *potential influencers* in social network, if they further connect to other clusters through tweets (i.e. to talk about multiple topics).

Hypothesis H2 Similarly hypothesis H2, has been tested through correlation. By Table 3, both $N_c(a)$ and $N_b(a)$ have positive correlation of 0.254 and 0.235, respectively with $N_r(a)$, at 0.01 level of significance, supporting the hypothesis H2. It means that, authors, who have large number of retweets, can also be ‘*information sources*’ or ‘*generators*’. Such authors can be *actual influencers* in spreading the influence among networks, as they receive large number of retweets by tweeting about multiple topics. Through visualization standpoint, if we explore the produced

Table 3 Correlation matrix of persistence variables (Pearson Index)

	$N_T(a)$	$N_R(a)$	$N_B(a)$	$N_C(a)$
$N_T(a)$	1	0.326	0.590	0.898
$N_R(a)$	0.326	1	0.254	0.235
$N_B(a)$	0.590	0.254	1	0.392
$N_C(a)$	0.898	0.235	0.392	1

Table 4 Estimates of regression weights

$V_{\text{Dependent}}$	$V_{\text{Independent}}$	R_w	S.E	p -value
$N_R(a)$	$N_T(a)$	0.082	0.000	<0.001
$N_B(a)$	$N_T(a)$	0.303	0.002	<0.001
$N_B(a)$	$N_R(a)$	0.000	0.000	<0.001
$N_C(a)$	$N_T(a)$	0.000	0.000	<0.001
$N_C(a)$	$N_R(a)$	0.648	0.009	<0.001

graph (e.g. Fig. 2), authors who seems to be big sized nodes (visually drawn in-between multiple cluster peripheries) talking about multiple topics (brands or categories), also have the high number of retweets as well. As these authors can be referred as ‘*information sources*’, it is evident to receive high number of retweets upon tweets about multiple topics (brands or categories).

Hypothesis H3 Similarly hypothesis H3, has been tested through correlation. From Table 3, $N_T(a)$ and $N_R(a)$ have positive correlation of 0.326 at 0.01 level of significance. Although the correlation coefficient is not high, the p -value (<0.001) in Table 4 showing significance and seems to support a positive (though weak) correlation between $N_T(a)$ and $N_R(a)$. Through visual standpoint, as shown in Fig. 2, we know that the nodes (which are drawn closer to single brand in innermost periphery of distinct clusters) are those authors who tweet most frequent about specific brand in its cluster. Such author nodes may be referred as *most specific* authors and can be both *potential and actual influencers* in social network, as they are frequent in tweeting and as well as in retweeting.

6 Discussions

Authors belonging to different clusters are in fact those who are more *generalist* in their content sharing, since they tweet about multiple different brands. On the contrary, authors belonging to the innermost clusters are those who are very *specific* in sharing content related to one selected brand. Since the *specificity* (generality) and *frequency of tweets* and *retweets* of authors was not an explicit variable in our dataset, it is possible to posit that the proposed network layout help to unveil specific (implicit) properties of the represented networks. We also noticed that, as the graph sizes increases, more peripheral layers seems to be formed surrounding hub nodes, which increases the influence spread across newly formed peripheral layers in multi-layered form. Thus authors tweeting about multiple topics among

multiple peripheries can be potential influence spreaders. An enlarged version of the network layouts for both networks N_1 and N_2 can be accessed online.¹ The clustering of nodes provides a distinct multi-layering of those authors who have tweeted about the same destination. The layering of nodes around brands is instead related to the intensity of tweeting about a given destination.

The emerging semantic of network visualization is related to the *brand fidelity* of authors, as shown in Fig. 2. Moreover, it is possible to point out which authors are tweeting about a brand as well as a competing brands to support the definition of specific marketing campaigns and for categories as well. Similarly, tourism practitioners can also point out the highly discussed touristic destination, and they can also identify the less popular destinations, upon which they can perform some strategic advertising campaigns.

7 Conclusion and Future Work

This paper proposes a novel visual approach for the analysis and exploration of social networks in order to identify and visually highlight influencers (i.e., hub nodes) and influence (i.e., spread of information across multi-layer peripheral nodes), represented by the opinions expressed by social media users on a given set of topics. Results show that our approach produces aesthetically pleasant graph layouts, by highlighting multi-layered clusters of nodes surrounding hub nodes (the main topics). These multi-layered peripheral node clusters represent a visual aid to understand influence. Empirical testing and evaluation results show that the proposed three hypothesis that tie *content specificity*, *frequency of tweets* and *retweets* are valid. Moreover, the parameters like *specificity*, *frequency*, and *retweets* are also mutually correlated, and have a significant impact on an author's influence and encourage us to further explore social network's intrinsic characteristics.

Such outcomes can be further utilized by tourism practitioners, marketing departments or social media community. For example, one can analyse the most competitive locations, events or initiatives in the market. Social media marketing managers can also visually identify major key players in the network, like *information spreaders* and *information sources*. In social media communities, users like *information seekers*, would be able to visually identify the actual and potential influencers and can further follow them.

Although our experiment can be repeated with data from domains different from tourism, additional empirical work is needed to extend testing to multiple datasets and domains. Future work will consider measures of influence with additional parameters (e.g. number of followers, lists, mentions, URLs, etc.). In our current work, we are studying a measure of influence through the proposed visualization

¹ Further visualizations can be accessed online from: <http://goo.gl/FmyWTq>

approach, which can be used to rank influential nodes in social networks (Metra 2014) and help the practical use of our research results.

References

- Abello, J., & Queyroi, F. (2013). *Fixed points of graph peeling*. Paper presented at the Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining.
- Alvarez-Hamelin, J. I., Dall'Asta, L., Barrat, A., & Vespignani, A. (2006). Large scale networks fingerprinting and visualization using the k-core decomposition. *Advances in Neural Information Processing Systems*, 18, 41.
- Andersen, R., Chung, F., & Lu, L. (2007). Drawing power law graphs using a local global decomposition. *Algorithmica*, 47(4), 397.
- Andersen, R., Chung, F., & Lu, L. (2004). Drawing power law graphs using local/global decomposition. In *Twelfth annual symposium on graph drawing*.
- Anholt, S. (2006). *Competitive identity: The new brand management for nations, cities and region*. Basingstoke: Palgrave Macmillan.
- Arbuckle, J. L. (2011). *IBM SPSS Amos 20 user's guide*. Amos Development Corporation, SPSS Inc.
- Asur, S., Huberman, B. A., Szabo, G., & Wang, C. (2011). *Trends in social media: Persistence and decay*. Paper presented at the ICWSM.
- Bagozzi, R. P., & Fornell, C. (1982). Theoretical concepts, measurements, and meaning. *A Second Generation of Multivariate Analysis*, 2(2), 5–23.
- Bakshy, E., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011). *Everyone's an influencer: Quantifying influence on Twitter*. Paper presented at the Proceedings of the fourth ACM International Conference on Web Search and Data Mining.
- Barbagallo, D., Bruni, L., Francalanci, C., & Giacomazzi, P. (2012). An empirical study on the relationship between twitter sentiment and influence in the tourism domain. In M. Fuchs, F. Ricci, & L. Cantoni (Eds.), *Information and Communication Technologies in Tourism 2012* (pp. 506–516). Helsingborg: Springer.
- Benevenuto, F., Cha, M., Gummadi, K. P., & Haddadi, H. (2010). *Measuring user influence in twitter: The million follower fallacy*. Paper presented at the International AAAI Conference on Weblogs and Social (ICWSM10).
- Bigonha, C., Cardoso, T. N. C., Moro, M. M., Gonçalves, M. A., & Almeida, V. A. F. (2012). Sentiment-based influence detection on Twitter. *Journal of the Brazilian Computer Society*, 18(3), 169–183.
- Boutin, F., Thievre, J., & Hascoët, M. (2006). Focus-based filtering + clustering technique for power-law networks with small world phenomenon. In: R. F. Erbacher, J. C. Roberts, M. T. Gröhn, K. Börner (Eds.), *Electronic imaging 2006* (pp. 60600Q–60600Q). International Society for Optics and Photonics.
- Boyd, D., Golde, S., & Lotan, G. (2010). Tweet, tweet, retweet: Conversational aspects of retweeting on Twitter. *IEEE*, 1–10.
- Bruni, L., Francalanci, C., Giacomazzi, P., Merlo, F., & Poli, A. (2013). *The relationship among volumes, specificity, and influence of social media information*. Paper presented at the Proceedings of International Conference on Information Systems.
- Carmi, S., Havlin, S., Kirkpatrick, S., Shavitt, Y., & Shir, E. (2007). A model of Internet topology using k-shell decomposition. *Proceedings of the National Academy of Sciences of the United States of America*, 104(27), 11150–11154.

- Cha, M., Haddadi, H., Benevenuto, F., & Gummadi, P. K. (2010). Measuring user influence in twitter: The million follower fallacy. *International Conference on Web and Social Media*, 10, 10–17.
- Chan, D. M., Chua, K. S., Leckie, C., & Parhar, A. (2003). *Visualisation of power-law network topologies*. Paper presented at the Networks, 2003. ICON2003. The 11th IEEE International Conference on (pp. 69–74). IEEE.
- Chen, C. (2006). *Information visualization: Beyond the horizon*. New York, NY: Springer.
- Fan, W., & Gordon, M. D. (2014). The power of social media analytics. *Communications of the ACM*, 57(6), 74–81.
- Francalanci, C., & Hussain, A. (2014). *A visual approach to the empirical analysis of social influence*. Paper presented at the DATA 2014—Proceedings of 3rd International Conference on Data Management Technologies and Applications. doi:[10.5220/0004992803190330](https://doi.org/10.5220/0004992803190330)
- Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, 21(11), 1129–1164.
- Hossain, L., Wu, A., & Chung, K. K. (2006). *Actor centrality correlates to project based coordination*. Paper presented at the Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work (pp. 363–372). New York, NY: ACM.
- Hussain, A., Latif, K., Rextin, A., Hayat, A., & Alam, M. (2014). Scalable visualization of semantic nets using power-law graphs. *Applied Mathematics & Information Sciences*, 8(1), 355–367.
- Hutto, C. J., Yardi, S., & Gilbert, E. (2013). *A longitudinal study of follow predictors on twitter*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.
- Kitsak, M., Gallos, L. K., Havlin, S., et al. (2010). Identification of influential spreaders in complex networks. *Nature Physics*, 6(11), 888–893.
- Klotz, C., Ross, A., Clark, E., & Martell, C. (2014). Tweet!—And I can tell how many followers you have. In *Recent Advances in Information and Communication Technology* (pp. 245–253). Berlin: Springer.
- Koch, R. (1999). *The 80/20 principle: The secret to achieving more with less*. New York, NY: Crown Business.
- Kwak, H., Lee, C., Park, H., & Moon, S. (2010). *What is Twitter, a social network or a news media?* Paper presented at the Proceedings of the 19th International Conference on World Wide Web.
- Leavitt, A., Burchard, E., Fisher, D., & Gilbert, S. (2009). The influentials: New approaches for analyzing influence on Twitter. *Web Ecology Project*, 4(2), 1–18.
- Li, J., Peng, W., Li, T., Sun, T., Li, Q., & Xu, J. (2014). Social network user influence sense-making and dynamics prediction. *Expert Systems with Applications*, 41(11), 5115–5124.
- Meraz, S. (2009). Is there an elite hold? Traditional media to social media agenda setting influence in blog networks. *Journal of Computer-Mediated Communication*, 14(3), 682–707.
- Messias, J., Schmidt, L., Oliveira, R., & Benevenuto, F. (2013). You followed my bot! Transforming robots into influential users in Twitter. *First Monday*, 18(7).
- Metra, I. (2014). *Influence based exploration of Twitter social network*. Milan, Italy: Politecnico di Milano.
- Myers, S. A., & Leskovec, J. (2014). *The bursty dynamics of the Twitter information network*. Paper presented at the Proceedings of the 23rd International Conference on World Wide Web.
- Naaman, M., Boase, J., & Lai, C.-H. (2010). *Is it really about me?: Message content in social awareness streams*. Paper presented at the Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work, Savannah, GA.
- Pallant, J. (2010). *SPSS survival manual: A step by step guide to data analysis using SPSS*. New York, NY: McGraw-Hill.
- Perline, R. (2005). Strong, weak and false inverse power laws. *Statistical Science*, 20(1), 68–88.
- Ren, Z.-M., Zeng, A., Chen, D.-B., Liao, H., & Liu, J.-G. (2014). Iterative resource allocation for ranking spreaders in complex networks. *Europhysics Letters*, 106(4), 48005.

- Sparrowe, R. T., Liden, R. C., Wayne, S. J., & Kraimer, M. L. (2001). Social networks and the performance of individuals and groups. *Academy of Management Journal*, 44(2), 316–325.
- Xu, X., Yuruk, N., Feng, Z., & Schweiger, T. A. J. (2007). *Scan: A structural clustering algorithm for networks*. Paper presented at the Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

Information and Communication Technologies in
Tourism 2015

Proceedings of the International Conference in Lugano,
Switzerland, February 3 - 6, 2015

Tussyadiah, I.; Inversini, A. (Eds.)

2015, XVI, 883 p. 139 illus., Softcover

ISBN: 978-3-319-14342-2