CrossMark

An empirical comparison of influence measurements for social network analysis

Khaled Almgren[1] · Jeongkyu Lee[1]

Abstract The studying of social influence can be used to understand and solve many complicated problems in social network analysis such as predicting influential users. This paper focuses on the problem of predicting influential users on social networks. We introduce a three-level hierarchy that classifies the influence measurements. The hierarchy categorizes the influence measurements by three folds, i.e., models, types and algorithms. Using this hierarchy, we classify the existing influence measurements. We further compare them based on an empirical analysis in terms of performance, accuracy and correlation using datasets from two different social networks to investigate the feasibility of influence measurements. Our results show that predicting influential users does not only depend on the influence measurements but also on the nature of social networks. Our goal is to introduce a standardized baseline for the problem of predicting influential users on social networks.

Keywords Social network analysis · Influence measurements · Predicting influential users · Centrality analysis

1 Introduction

Recently, billions of users are using social networks as part of their daily routine. By December 2015, Facebook had approximately 1.04 billion users who are active on a daily

✉ Khaled Almgren
   kalmgren@my.bridgeport.edu

   Jeongkyu Lee
   jelee@bridgeport.edu

[1] University of Bridgeport, 167 Park St, Bridgeport, CT, USA

basis (Facebook 2015). In 2015, Twitter had a total of 288 million active users on a monthly basis who shared 500 million *tweets* per day (Twitter 2015). Flicker had a total of 92 million users in 2014 who shared about 1 million images per day and contributed to 2 million groups (Etherington 2014). This has attracted many people from academia and business to explore various research topics, including measuring influence on social networks. In social sciences, influence is defined as "change in an individual thoughts, feelings, attitudes, or behaviors that results from interaction with another individual or a group (Black 2004)." Webster also defines influence as "the act or power of producing an effect without apparent exertion of force or direct exercise of command (Merriam-Webster 2011)." The changes and effects can be observed as feelings, thought or actions (Li et al. 2015). The reality of these definitions can be found on social networks, through users actions, such as *retweets* on Twitter, *likes* on Facebook and *favorites* on Flicker.

Influence has been studied for a long time by many sociologists because it can affect people's decision-making abilities (Watts and Dodds 2007). Due to that, sociologists have proposed the two-step flow communication model in 1955 (Katz and Lazarsfeld 1955). They state that influence flows from the media to leaders and from leaders to the people (Katz and Lazarsfeld 1955); therefore, to influence people's decisions, their leaders or peers can be targeted. Predicting influential users on social networks can be very useful in viral marketing. Companies can utilize social networks as a marketing tool. They can promote their products and services (Stelzner 2011). However, social networks are not like traditional mass media where the marketing operation is entirely controlled by only the mass media, as in TV and newspapers. In social networks, individuals can generate content that can attract many users

🖄 Springer

who can be targeted by companies (Pempek et al. 2009). Users who have this power can impact other people's decisions. Therefore, companies can directly use them to approach potential customers. Predicting influential users on social networks can also be applied to many other real-world applications, including recommendations systems, and expert search engines (Aggarwal 2011; Cha et al. 2010; Granovetter 1973; Kwak et al. 2010; Lee et al. 2010; Li et al. 2015, 2013a; Probst et al. 2013; Reilly et al. 2014; Sang and Xu 2013; Sun and Ng 2013).

Researchers have looked at the problem of measuring influence from two research perspectives. The first group of researchers addresses influence maximization on social networks. Influence maximization is defined as finding the minimum number of users that can maximize the diffusion of influence on social networks (Kempe et al. 2003; Wang et al. 2012). The second group focuses on finding influence measurements to predict influential users on social networks. This paper focuses on the second group's problem.

Even so many researchers have addressed the problem of predicting influential users on social networks, but the approaches to determine influential users are still in the early stages of discovery and this area encounters multiple challenges (Probst et al. 2013). The major challenge is the absence of ground truth data. Due to the lack of ground truth data, it is very difficult to evaluate the proposed approaches. Therefore, many researchers evaluate the proposed influence measurements on sample social network data without a proper comparison of the influence measurements in terms of accuracy or performance (Bakshy et al. 2011). In addition, previous surveys compared influence measurements based on theoretical analysis instead of empirical analysis (Probst et al. 2013; Rabade et al. 2014; Riquelme 2015; Singh 2013). However, influence measurements are not proved mathematically. Therefore, it is necessary to conduct an empirical analysis of influence measurements.

In order to fill this gap, this paper provides a baseline classification that can describe and categorize influence measurements. Using this classification, we perform a comparison of influence measurements using social network datasets from Digg and Flicker to investigate the feasibility of the influence measurements to different social networks.

The social influence measurements are classified using three main folds:

1. Influence measurement by models,
2. Influence measurement by types and
3. Influence measurement by algorithms.

First, Zafarani et al. (2014) proposed *Influence measurement by models* to categorize influence measurements.

Their classification describes the characteristics of the influence measurements, for example, users' attributes such as activeness (Zafarani et al. 2014). The influence measurement models are categorized into observation-based and prediction-based models. However, this classification does not cover the various points of view of those measurements. Therefore, we add two more classifications that consider different points of view. They are Influence measurement by *types* and *algorithms*. For the second classification, the influence measurement by *types* describes the kinds of structures that are used for measuring influence. This classification includes (i) context, (ii) content and (iii) hybrid. In the context category, researchers consider only network structure, while in the content category, they consider only content from social networks, such as *posted images* on Flicker. In the hybrid category, both content and context are used. The third classification, i.e., the influence measurement by *algorithms*, consists of the techniques that are used to build the measurements. The algorithms used in this classification include (i) social network measures such as centrality analysis, (ii) social network properties such as *number of tweets* on Twitter and (iii) information cascade modeling such as diffusion of content.

The main contributions of our work are briefly summarized below:

1. A comprehensive classification is introduced to categorizes, compares and summarizes the influence measurements in published academic and industrial research.
2. We investigate the adaptability of multiple influence measurements to different social networks by employing several influence measurements to Digg and Flicker.
3. We highlight promising new directions based on an empirical analysis of influence measurements.

The rest of this paper is organized as follows: Sect. 2 defines the problem. Section 3 discusses related works. Section 4 discusses the influence measurement by models. Section 5 discusses the influence measurement by types. We discuss the social network analysis by algorithms in Sect. 6. Section 7 compares the literature survey by the influence measurements by algorithms. We discuss the results in Sect. 8. The conclusion is provided in Sect. 9.

2 Definitions

To clarify the problem, we recall several known definitions of influence measurement in social networks. These definitions will be used throughout the paper:

Definition 1 Social network is a virtual environment that allows people to create profiles, follow and interact with other users and view users' profiles and posts (Ellison 2007).

For example, on Facebook, users can follow pages, people or figures. They can also view their profiles and interact with them in many different ways such as *liking* other users' posts

Definition 2 Social interactions are actions on social networks that represent the users interactions between each other such as *comment*, *favorites*, *retweets* or *likes* (Almgren and Lee 2015).

These interactions can be either directed such as comment or undirected such as friendship.

Definition 3 Influence in social networks is the ability to make other users react to actions of a user by performing further social interactions toward the user's posts (Leavitt et al. 2009).

For example, when a user posts an interesting *tweet*, several users can endorse the *tweet* by liking or retweeting it.

Based on these definitions, we further define important concepts used throughout this paper.

Definition 4 Influential users in social networks are the users who can influence other users.

For example, users who have many followers on Twitter.

Definition 5 Followers in social networks are users that follow or interact with other users.

For example, followers on Twitter.

Definition 6 Influence Models are representations of social networks that can be used to present users and to measure influence.

Influence models can be graph-based, tree-based or linear-based models. These models illustrate how influence flows between users on social networks. Each influence model is used based on the social influence measurement model and technique.

Definition 7 Influence measurement: By using any social network that can be represented as an influence model, we can predict the influential users who can influence many users to perform social interactions such as *follows* and *retweets* on Twitter, *votes* on Digg, *Favorites* on Flicker or *Likes* on Facebook.

3 Related work

In this section, we review related works that classified the influence measurements. There are four studies that surveyed this problem.

The first paper studies this problem from a non-technical perspective (Probst et al. 2013). It uses Katz's Influence Theory (Probst et al. 2013). Katz states that influence is related to three main concepts (Katz and Lazarsfeld 1955). These concepts are:

1. The location of a person in society,
2. A person's expertise and
3. A person's abilities and traits.

They categorize their literature based on the above three concepts and users' activeness. The survey discusses the approaches that were used to solve the problem without discussing the technical aspects.

The second survey categorizes the survey into social networks, technical approaches and targeting marketing techniques (Singh 2013). They classify social networks using network structures. They categorize them into small graphs, random graphs, networks or power low networks. They categorize the approaches into algorithms that are based on Markov random field, social network topology, random walks, PageRank, time spent by users on social networks, number of follower, SPIN value of nodes, dividing the social networks or models of social networks. They also review the marketing techniques used in viral marketing. Their classification lacks clear structure since many of their classified components can be grouped into the same category. For example, algorithms that are based on random walks and Markov random field should be grouped together since they are both based on Markov Process.

The following survey classified the influence measurements by the techniques that are used to measure influence (Rabade et al. 2014). The measurements are classified into structural measures, diffusion model, community mining, content mining, micro-blog marketing and link polarity. They surveyed 10 papers and categorized them using their classification. This paper only considers the algorithmic aspect of the influence measurement. The surveys categorize papers that try to address influence maximization or prediction of influential users on social networks without distinguishing between the two problems clearly (Probst et al. 2013; Rabade et al. 2014; Singh 2013).

This following survey paper analyzed influence measurements used to identify influential users on Twitter (Riquelme 2015). They only consider influence measures

used for Twitter, such as mention impact and retweet impact. They have performed few experiments that do not reflect the performance of the influence measurements on Twitter.

Previous works have analyzed influence measurements based on theoretical basis neglecting the empirical analysis of the measurements. On the contrary, this paper compares the influence measurements empirically on datasets retrieved from two real-world social networks, Digg and Flicker, to investigate the measurements adaptability to different social networks. The influence measurements are compared in terms of accuracy, correlation and performance. As far as we know, this is the first survey that compares the influence measurements empirically by conducting experiments on different social networks. Our goal is to provide a clear classification that can be used as a baseline structure that covers the various principles of influence measurements.

The classification is shown in Fig. 1. Table 8 in the Appendix contains a complete list of the surveyed papers.

4 Social influence measurement by models

In this section, we discuss the influence measurement models and their subcategories. They are prediction-based and observation-based models (Zafarani et al. 2014).

4.1 Prediction-based model

In order to measure influence, the prediction-based model utilizes the structural location of users in a network, users' attributes or both. This model is classified into location model, attribute model, and location and attribute model (Zafarani et al. 2014).

4.1.1 Location model

In this model, the influence of a user is determined by the user's structural location on social networks. This approach uses network measures such as centrality analysis to measure influence (Zafarani et al. 2014).

Several papers study users' influence on Twitter. Weng et al. (2010) propose TwitterRank to identify influential users. They define influence as the ability to generate content with interesting topics. They predict influential users based on topical similarity between users and link structure. Sun and Ng (2013) predict influential users based on the interactions of posts. They define influence as the ability to share posts that generate many implicit and explicit interactions. They consider two types of interactions: explicit interactions, i.e., *replays* and implicit interactions, i.e., posts that talk about the same topic. Cha et al. (2010) propose several measurements that are based on different models. One of their measurements defines influence as the ability to attract many users to follow influential users. Kwak et al. (2010) propose three influence measurements. Two of their measurements are based on the structural location of users. One measurement defines influence as the ability to attract users to follow other users. The second measurement defines influence as the ability to attract important neighbors. Maharani et al. (2014) use two influence measurements to predict influential users. Their measurements can be defined as users who attract many important users to follow them. They build their influence model using undirected relationships between users.
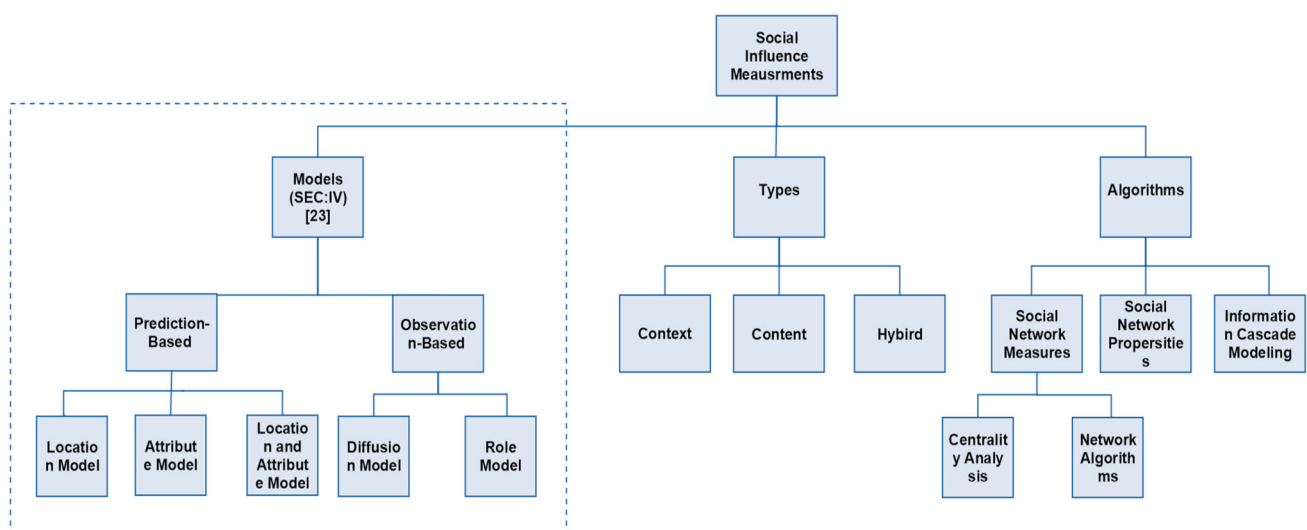

Fig. 1 Social influence measurements classification

Weibo is also one of the social networks that papers have used to measure influence. Li et al. (2013b) propose a new measurement based on the user-to-user influence. The user-to-user influence considers four factors that represent four types of interactions on social networks. They define influence as the ability to generate content that generates high retweeting strength, commenting density, mentioning density and *tweets* that are similar to the influential user's *tweets*. Liao et al. (2013) propose WeiboRank to rank users. They define influence as the ability to attract many important followers based on three processes, i.e., *follow*, *repost* and *comment-only*. They introduce *dependence* to trace the source of influence. Zhang et al. (2012) analyze influence using three social actions, i.e., *following*, *retweeting* and *commenting*. They define influence as the ability to attract many actions from important neighbors.

Other papers have used Digg, Flicker and Delicious to measure influence. Ghosh and Lerman (2010) predict influential users on Digg. They define influence as the average number of votes that each story receives. They state that non-conservative models are the best in predicting influential users. An example of a non-conservative model is information spread. Their methodology shows that users with the most important neighbors are the most influential users. Lu et al. (2011) propose LeaderRank to identify influential users on social networks. They define influence as the ability to attract important neighbors to perform interactions such as voting. Their measurement is based on the users' structural locations on social networks. Almgren and Lee (2015) are one of the first researchers to examine social influence on Flicker. We define influence as the ability to generate content that attracts many people to perform social interactions. We propose an influence measurement based on direct active followers, tie strength and the structural locations of users in a network.

4.1.2 Attributes model

In this model, users influence others using their personal attributes. This approach employs network measurements to quantify the influence of each user. For example, a user can be called active if the user has shared many posts. This can be measured by the weight of each node (Weng et al. 2010).

Leavitt et al. (2009) is one of the first research groups to study the effect of users' attributes on influencing users on Twitter. The measurement utilizes the attribute model based on the users' abilities to make other users engage in conversations. Cha et al. (2010) propose another measurement that is based on the popularity of users on Twitter. They define influence as the ability to make other users engage in conversations. Anger and

Kittl (2011) use several influence measurements on Twitter that are based on different models. One of their measurements is based on the popularity of users. They define influence as the ability to attract many users to follow the influential users. Another measurement is users' activeness. It is based on the users' contribution level. They define influence as the ability to generate many posts, which can show users' activeness. Yi et al. (2013) propose a measurement to evaluate users' attributes on social networks.

4.1.3 Attributes and locations model

In this model, researchers combine personal attributes and structural location to predict influential users. This approach uses network measurements. Almgren and Lee (2016) investigated the relationship between a user's structural location using several centrality analysis algorithms and a user's attributes on Flicker. We represent a user's attributes by the user's activeness, i.e., how many images a user upload to a Flicker group. Influence is defined as the ability to have certain attributes and be connected to many important users (Almgren and Lee 2016).

4.2 Observation-based model

Observation-based models use the amount of influence that each user generates, for example, the number of influenced people, users' ability to spread information and the power of users to increase the value of products. This approach can be classified into three models: role Model, diffusion Model and value Model (Weng et al. 2010). However, in this paper, we use the role model and diffusion Model because none of the measurements is based on the value Model.

4.2.1 Role model

In this model, the influence of each user is based on the power of users; for example, a teacher can influence students because he/she is in a position of power. The teacher's influence can be measured by the number of students (Weng et al. 2010).

Lee et al. (2010) identify influential users on Twitter with the time series of information adoption. They define influence as the ability to generate content that is read by many people. They assume that influence is time-sensitive where users who tweet first have a higher probability of becoming influential. They track *tweets* to measure the spread. The user who has many effective readers is regarded as the role model. Sun et al. (2013) define influence as

the number of effective audience members that users have. The effective audience can be implicit or explicit. The implicit effective audience are users who follow other users and are exposed to their posts. On the other hand, explicit effective audience are users who perform interactions toward the influential users' posts.

4.2.2 Diffusion model

In this model, the influence of users is measured by their ability to spread information. The influence is measured by how much the information has diffused on a social network. For example, *tweets* on Twitter can spread if they are transferred among many users in a short period of time. Influence can be measured using the cascade size (Weng et al. 2010).

Bakshy et al. (2011) propose an influence measurement by tracking the diffusion of URLs on Twitter. Influence is defined as the ability to generate a URL that diffuses massively on Twitter. They define the cascade size as the reposts of URLs from the user's followers and their followers. Several papers define influence as the ability to generate content that spreads on social networks. They use the diffusion of *tweets* to measure influence (Anger and Kittl 2011; Cha et al. 2010; Kwak et al. 2010; Leavitt et al. 2009; Reilly et al. 2014).

5 Social influence measurement by types

In this section, we discuss the types of structure that influence measurements are based on. They are classified into the context, content and hybrid.

5.1 Context

Context measurements measure influence using the structural properties of social networks by considering users on social networks and the relationships between the users. Context measurements use the graph theory to present users as nodes and relationships as edges.

Two papers analyze influence using the followers and friendship networks (Anger and Kittl 2011; Kwak et al. 2010). Lu et al. (2011) measure influence using the friendship network. Cha et al. (2010) propose a measurement using the followers network.

5.2 Content

Content measurement uses content produced by users in measuring influence. In this type, researchers use content in building the influence model such as the diffusion of *tweets* on Twitter.

Several works consider the power of generated content by users as an effective indicator of influence, such as the number of *tweets* in different topics and *tweets'* similarity (Anger and Kittl 2011; Brandes 2001; Cha et al. 2010; Kwak et al. 2010; Lee et al. 2010; Reilly et al. 2014).

5.3 Hybrid

Hybrid measurements integrate network structure and content. In this type, the focus is on the dynamic process that takes place on social networks, for example, *favorite* on Flicker or *retweet* on Twitter. Researchers build the influence model based on the users' posts as nodes and the dynamic processes as edges. Therefore, each user will be represented as a node where the node can be weighted to represent the number of shared posts and a directed edge between two users when they interact through their posts. For example, Almgren and Lee (2015) build the influence model using the users who participate in a Flicker group, where the edges represent *favorite* and *comment* on the images.

Several studies propose influence measurements that use followers network and content (Almgren and Lee 2016; Li et al. 2013b; Liao et al. 2013; Maharani et al. 2014; Sun et al. 2013; Weng et al. 2010; Zhang et al. 2012). Ghosh and Lerman (2010) measure influence on Digg using the fan network and stories. Almgren and Lee (2015; 2016) measure influence using the followers network and images. Sun and Ng (2013) use both content and interactions between users to identify influential users.

6 Influence measurement by algorithms

In this section, we present three major algorithm types used in identifying influential users. They are social network measures, information cascade modeling and social network properties. Social network measures are based on social network theory. Information cascade modeling uses the information diffusion theory. Social network properties use existing social network measurements such as *number of retweets* on Twitter. Network measures fundamentally show the power of users, while information cascade modeling shows the power of content.

6.1 Social network measures

Social networks can be presented as graphs that are comprised of nodes and edges. Nodes can represent actors, where edges represent relationships between actors (Bakshy et al. 2011; Scoot 1992). Since networks are represented as graphs, several network measurements can be

utilized for social networks. There are two main network measurements that have been applied to identify influential users: centrality analysis and network algorithms.

6.1.1 Centrality analysis

Centrality analysis ranks users by their structural locations on social networks. Centrality represents the importance of users on networks (Weng et al. 2010; Tang and Liu 2010). Freeman (1978) and Newman (2001) state that centrality is an important attribute of social networks. Different centrality analysis techniques are used to reflect the importance of standing positions of the users. For example, in-degree centrality ranks users using their direct neighbors. In this paper, we will discuss the centrality measures that are used in the influence measurements covered in this paper. The equations used throughout this paper are the same, but the notations are changed to maintain the consistency of the paper.

Almgren and Lee (2015) use weighted in-degree, i.e., $C_{wd}(v_i)$ to identify influential users. $d_i^{w(in)}$ is calculated as follows:

$$C_{wd}(v_i) = d_i^{w(in)}, \tag{1}$$

where $d_i^{w(in)}$ represent the total weight of the incoming edges to node v_i. This measurement considers the tie strength between nodes with the directed relationships between them.

Lee et al. (2010) apply PageRank on Twitter. The PageRank, i.e., C_{pr} is computed as below (Zafarani et al. 2014).

$$C_{pr}(v_i) = \alpha \sum_{j=1}^{n} A_{j,i} \frac{C_p(v_j)}{d_j^{out}} + \beta, \tag{2}$$

where α is used to avoid zero and β is an attenuation factor. $A_{j,i}$ is the adjacency matrix and d_j^{out} is the out-degree of j. There are many proposed algorithms that are based on PageRank. For example, WeiboRank applies PageRank on Weibo (Liao et al. 2013). Zhang et al. (2012) use weighted PageRank that combines several interaction types such as *follow* and *retweet*. Yi et al. (2013) combine interactions and connections. Li et al. (2013b) consider four types of interactions. TwitterRank uses the topical similarity between users (Weng et al. 2010).

Ghosh and Lerman (2010) propose the normalized α-centrality. Normalized α-centrality utilizes α-centrality proposed by Bonacich and Lloyd (Bonacich and Lloyd 2001); α-centrality considers the importance of the incoming neighbors as well as external factors. Bonacich and Lloyd (2001) state that not only the centrality of users depends on their connections, but it can also depend on some external factors. Therefore, he proposes α-centrality

where α represents the importance of endogenous versus exogenous factors. Endogenous factors represent the importance of incoming connections where exogenous factors represent the external factors. α-centrality is calculated using the equation below.

$$C_{alpha,\alpha} = v \left(\sum_{t=0}^{k \to \infty} \alpha^t A^t \right), \tag{3}$$

where v represents the vector of exogenous factors and A^t is the adjacency matrix. Ghosh and Lerman (2010) further normalized this measure by the total $C_{alpha} \ \forall \ i$ neighbors. Ghosh and Lerman (2010) further prove that their measurement converges.

$$C_{N_\alpha,\alpha} = \frac{C_{alpha,\alpha}}{\sum_{i,j}^{n}(\alpha^t A^t)_{i,j}}, \tag{4}$$

where $\sum_{i,j}^{n}(\alpha^t A^t)_{i,j}$ represents the centrality value between j and i.

This measurement is based on centrality analysis (Cha et al. 2010). The measurement is in-degree centrality C_{din}, which reflects the number of nodes that point to each node on social networks. The measurement is computed as below:

$$C_{din}(v_i) = d_i^{(in)}, \tag{5}$$

where $d_i^{(in)}$ is the number of nodes that point to v_i.

Lu et al. (2011) propose LeaderRank. It is similar to PageRank. The difference between them is LeaderRank is parameter-free. LeaderRank adds a node to the graph called *ground node*. This node makes the graph well connected to make the algorithm parameter-free. LeaderRank computes the influence score s_i for each node at time t. However, it neglects the score for *ground node*. LeaderRank is computed as below:

$$C_{ld}(t+1) = \sum_{j=1}^{N+1} \frac{a_{ji}}{k_j^{out}} s_j(t), \tag{6}$$

where $\frac{a_{ji}}{k_j^{out}}$ represents the random walk of nodes. a_{ji} represents the directed edge from j to i. Therefore, if the edge exists $a_{ji} = 1$; otherwise $a_{ji} = 0$. k_j^{out} is the out-degree of j, i.e., number of nodes j point to.

Sun and Ng (2013) propose a measurement to identify starter posts. They identify starter posts using degree centrality. Starter posts have many followers and follow very few. This is measured by computing the difference between in-degree and out-degree for all nodes. It is computed using the following equation:

$$C_d = d_v^{in} - d_v^{out}, \tag{7}$$

where d_v^{in} represents the nodes that point to v and d_v^{out} represents the nodes that v points to.

Their third measurement is graph entropy measure that is based on centrality analysis, where the centrality of v_i is computed using its decedents, i.e., nodes that point to v_i and nodes that point to nodes that point to v_i and so on, i.e., $Des(v_i)$. The centrality is computed using the following equation:

$$C_{Inf_e}(v_i) = \frac{En(i)}{log(EN(i)/E(i))}, \tag{8}$$

where $En(i)$ is the entropy of each v_i and it is computed using the distance between v_i and its ts descendants.

Maharani et al. (2014) propose using complex degree centrality, i.e., C_c and eigenvector centrality, i.e., C_{eg}, on Twitter. C_c computes the centrality of node A as follows:

$$C_c = (DC_A \times TR_A)^{.5}, \tag{9}$$

where DC_A is degree of node A, and TR_A is the weighted degree of A.

C_{eg} is based on the largest eigenvalue. It is computed as follows:

$$C_{eg} = \frac{1}{\lambda} \times \sum_{j=1}^{n} a_{ij} x_j, \tag{10}$$

where λ is a constant, a_{ij} represents the adjacency matrix, and x_j is the eigenvector of the nodes.

Almgren and Lee (2016) propose a hybrid measurement using centrality analysis and user's activeness, i.e., $C_{HEG}(v_i)$, to identify influential users. We found eigenvector centrality to be the most stable centrality analysis algorithm. It is calculated as follows:

$$C_{HEG}(v_i) = T \times C_{eg} + (1 - T) \times UActive, \tag{11}$$

where T is a variable used to control the relative importance between eigenvector centrality and a user's activeness. We recall C_{eg} from Eq. 10. UActive is computed using the total number of uploaded images for each user. Based on our experiments, we adopt $T = 0.5$ because it gives the optimal results.

6.1.2 Network algorithms

In this section, we discuss the network algorithms that are used to measure influence on social networks.

Sun and Ng (2013) propose another algorithm to identify starters. The measurement uses the shortest path cost algorithm. The basic idea behind this algorithm is to measure the influence of a node by observing how many other nodes will be affected if that node is removed. They use two sets to represent the relationship between nodes. The first set is $Des(v_i)$ of v_i, where v_i belongs to $Anc(v_d)$ because of its ancestor v_d, $\forall v_i \in Anc(v_d)$ and v_d

$\in Dec(v_i)$. This algorithm computes the influence of every node $Inf_c(v_i)$ as below:

$$Inf_c(v_i) = \sum_{v_d \in Des(v_i)} C(v_d, G) - C(v_d, G, v_i), \tag{12}$$

where $C(v_d, G)$ is the average shortest path cost between node v_d and its ancestors when removing v_i, and $C(v_d, G)$ is used when nodes do not have decedents, and the value will be set to 0. The average shortest path cost is used to reflect the influence effect $c(v_i)$ on nodes when removing v_i as computed below:

$$C(v_d) = \frac{1}{Anc(v_d)} \sum_{v_d \in Anc(v_d)} W_{(v_d.v_a)}, \tag{13}$$

where $Anc(v_d)$ is the ancestor of v_d, $W_{(v_d.v_a)}$ represents the relationship strength from v_d to v_a, and it is calculated as below:

$$W(v_d, v_n) = \prod_{i=d}^{n-1} W_{i,i+1}, \tag{14}$$

where $W_{i,i+1}$ represents the strength from the accumulative weight from v_d to v_n.

6.2 Social network properties

To measure influence, several researchers use existing social network properties from social networks, such as the *number of retweets* or the *number of tweets* on Twitter. These measurements can reflect many characteristics such as the popularity of users or diffusion of posts.

Kwak et al. (2010) rank users using the total number of *retweets* on Twitter. This measurement can reflect the popularity of *tweets*. Cha et al. (2010) use the total number of *mentions*, which can show the ability of users to make other users engage in conversations. Anger and Kitll (2011) combine several Twitter properties in two measurements. The first measurement is the average number of *followings* over the total number of *followers*. Their other measurement computes the total number of *mentions* and *retweets*. Leavitt et al. (2009) also propose two measurements that use Twitter properties. The first measurement reflects the spread of content where the second measurement reflects the conversational activities that the *tweet* generates. The first measurement is based on the total sum of *retweets* and *attribution* over *tweets*, while the second measurement is the total sum of *replays* and *mentions* over the number of *tweets*. Reilly et al. (2014) use the number of *tweets* and the number of *retweets* in identifying influential users. Their measurement considers the diffusion of *tweets*. They consider the percentage of the diffused *tweets* over the users' *tweets*.

6.3 Information cascade modeling

The theory of information cascade is defined as how information is transferred to users' followers and so on (Zafarani et al. 2014). A simple example of information cascade is *retweet* on Twitter. However, *retweet* is considered a social network property technique since it uses the *retweet* property. In information cascade modeling, researchers use the information cascade theory to model and measure influence to identify influential users.

Lee et al. (2015) identify influential users on Twitter using the adoption of *tweets* based on adoption times. Their model basically considers the users who are first exposed to the users' *tweets*, i.e., *effective readers*. Since many people follow more than one user, each user can *tweet* the same information. They consider the first person who posted the *tweet* as influential. Therefore, early users are more influential. Their model includes two phases: initialization and information diffusion. The model consists of two states. The first state is 0, which means user has not yet read the *tweet* and 1 that means user has read the *tweet*. In the initialization phase, all users start with 0 for all $u \in U$, $S(u) = 0$, where U is all the users and $S(u)$ represents the state of the users. The information diffusion phase accounts for the users whose states have changed to 1. It is shown as below:

$$ER_0(w) = \{v | v \in follower(u) \, and \, s(v) = 0\}, \quad (15)$$

where ER is the effective reader. $follower(u)$ is u followers. The influence of user $IF_0(u)$ is then computed as below:

$$IF_0(u) = \sum_{w \in T(u)} ||ER_0(w)||, \quad (16)$$

where $\sum_{w \in T(u)} ||ER_0(w)||$ is the total number of effective readers of u for all of his/her *tweets* $T(u)$.

Bakshy et al. (2011) propose an information diffusion model to identify influential users on Twitter. Their model uses the repost of users' posts that contain the URLs based on time. Their model uses an influence tree that represent the influential user's post as the seed node and the users who repost the URLs as leaves. They do not use *retweet*; they track the actual posts that contain URLs. They measure the cascade size from the total number of users in the influence tree. Their model has two cases when it comes to assigning influence scores to users. The first case gives full credit to the first person who posts the URL in the influence tree. The second case occurs when one user follows two people who posts the same URL. In this case, the influence score can be given to the last user who post the URL or it can be divided equally to the users who posted the URL. They are computed as below:

$$IF_0(u) = \sum_{i=0}^{n} u, \quad (17)$$

where $\sum_{i=0}^{n} u$ represents the total number of users in the influence tree. This is used in the first case. For the second case, they consider u from the seed who posts lastly or divide the influence score equally to the number of seeds.

Sun et al. (2013) propose six measurements to measure influence. The first three measurements use the total number of effective audience, i.e., *UI*. The measurement considers both implicit and explicit effective audience. It is calculated as follows:

$$UI(u) = U_{w \in T(u)} ||EA(u, w)||, \quad (18)$$

where $EA(u, w)$ represents the total number of effective audience members for all tweets T posted by user u. It is computed as follows:

$$EA(u, w) = EEA(u, w) \cup IEA(u, w), \quad (19)$$

where $EEA(u, w)$ represents the total number of explicit effective audience and $IEA(u, w)$ is the total number of implicit effective audience. They are computed as follows:

$$EEA(u, w) = \sum_{i=0, i \neq j}^{i < n} IN(i, w), \quad \forall w \in T(u), \quad (20)$$

$$IEA(u, w) = u \in U, \quad v \in fol(u), S(v, w) = 0 \quad (21)$$

where $\sum_{i=0, i \neq j}^{i < n} IN(i, w)$ represents the total number of users who reply or retweet to u tweets T. $IEA(u, w)$ represent the total number of followers v that read user u tweets and do not perform an interaction. This measurement is based on the probability of followers reading the tweets, i.e., $p = \frac{\alpha}{||followee(i)||}$. $followee(i)$ is the number of i followers and α is scaling parameter.

7 Empirical analysis

In order to evaluate and compare the influence measurements introduced in the previous sections, we have conducted three experiments on two different datasets. The first experiment evaluates the *accuracy* of the measurements. The second experiment performs *correlation analysis* for the ranking of influential users produced by different measurements and the ground truth. The third experiment shows the performance analysis of measurements in terms of *computational complexity*.

7.1 Dataset

We have used two datasets to assess the measurements discussed in this paper: Flicker[1] and Digg[2] datasets to investigate the adaptability of influence measurements to different social networks. All source codes and datasets can be downloaded from http://www1bpt.bridgeport.edu/~jelee/sna/sna.html.

Flicker is a social network that is based on images. The dataset is downloaded from one of the groups in Flicker that includes users, images, interactions and other meta data such as photo tags. A total of 30,759 users have participated in the group where some of them can be popular by posting images, while others only interact with other users. For example, 1559 users have uploaded 4991 images. There are 46,059 interactions between users representing comments and favorites.

On the other hand, Digg is a social network that allows users to share news stories. Users can interact with each other by voting on stories. The dataset contains 139,409 users and 1,534,314 edges represent voting. Among these users, 474 users have shared 3553 stories. This dataset is provided by Hogg and Lerman (2012).

These two datasets are different in nature. Digg is focused on news, while Flicker is focused on images and photographers. By conducting our analysis on these two datasets, we are considering different behaviors on social networks.

Table 1 shows the characteristics of both datasets including the number of nodes, which represents the number of users, the number of edges that represents interactions, number of images, number of users who contribute to social networks by posting images or stories, the number of users who are selected as influential users candidate, and the final number of influential users.

The ground truth is created using the statistical influence measure proposed by Ghosh and Lerman (2010). They show that the average number of fan votes can effectively measure users' influence using the *URN model* (Ghosh and Lerman 2010). The fan votes basically represent the directed interactions between users.

Therefore, we apply this measurement to both datasets to evaluate and assess the influence measurements. We consider only users who contributed at least ten posts and has at least ten followers because these users are considered active. The Flicker's dataset has only 72 users who met this requirement, while 74 users met this requirement in the Digg's dataset. These empirical rankings are referred to as *EM*.

---

[1] http://www.Flicker.com.

[2] http://www.digg.com.

7.2 Measurements selection

In this subsection, we show the influence measurements that are used in this experiment. We have selected the influence measurements based on their feasibility to various social networks including our datasets Flicker and Digg. For example, we have neglected the measurement from Li et al. (2013b) since it is based on functions that are only adopted on Weibo. As shown in Table 8 in the Appendix, we have surveyed a total of 19 research papers that proposed 37 measurements. Out of these 37 measurements, 9 measurements can be applied to both datasets and can be adopted to any other social network. Table 2 shows the selected influence measurements.

7.3 Experiment 1: accuracy analysis

We predict the rankings of influential users using each influence measurement. These rankings are referred to as *PRD*. Then, the results, i.e., *PRD*, are compared with the ground truth mentioned in the previous Sect. 7.1 (i.e., *EM*).

To evaluate the accuracy of influence measurement, we draw recall–precision graphs and compute the area under the curve using the trapezoidal rule implemented in Pedregosa et al. (2011). In order to draw this graph, we need to compute the recall (R) and precision (P) adopted from information retrieval theory (Robertson 2000). We perform the recall and precision over several number of returned candidates of influential users referred to as *K*. They are calculated using the following equations below:

$$R = \frac{(Number\ of\ relevant\ items\ retrieved)}{(Total\ number\ of\ relevent\ items)}, \qquad (22)$$

$$P = \frac{(Number\ of\ relevant\ items\ retrieved)}{(Total\ number\ of\ retrieved\ items)}, \qquad (23)$$

where the number of relevant items is the number of *PRD*, while the number of relevant items retrieved is *PRD* $\cap$ *EM*. (Note that the top *size(EM)*-ranked users according to *PRD* are considered to be retrieved).

Tables 3 and 4 contain the recall and precision rates for both datasets. These rates are used to compute the area under the curve to compare the performance of the accuracy between different influence measurements.

To better illustrate the accuracy results, we grouped the results into different subgroups: AccGr1, i.e., weak accuracy, AccGr2, i.e., medium accuracy, and AccGr3, i.e., strong accuracy. The area rates that are less than 0.30 are assigned to AccGr1, rates between 0.30 and 0.69 are assigned to AccGr2, and rates greater than or equal to 0.70 are assigned to AccGr3.

For Flicker's dataset, all measurements except hybrid eigenvector have weak accuracy rates, and therefore, they

Table 1 Flicker's dataset Characteristics

Dataset	Number of nodes	Number of edges	Number of images	Number of contributing nodes	Number of candidate influential users
Flicker	30,759	46,213	4838	1420	72
Digg	139,409	1,534,314	3553	474	74

Table 2 Selected influence measurements

Index	Influence measurement	Mathematical definition	Equation number	References
1	Complex degree	$C_c = (DC_A \times TR_A)^{.5}$	13	Maharani et al. (2014)
2	Degree	$C_d = d_v^{in} - d_v^{out}$	7	Sun and Ng (2013)
3	Eigenvector	$C_{eg} = \frac{1}{\lambda} \times \sum_{j=1}^{n} a_{ij} x_j$	14	Maharani et al. (2014)
4	Hybrid eigenvector	$C_{HEG}(v_i) = T \times C_{eg} + (1-T) \times U\,Active$	11	Almgren and Lee (2016)
5	In-degree	$C_{d_{in}} = d_i^{(in)}$	5	Cha et al. (2010)
6	LeaderRank	$C_{ld} = \sum_{j=1}^{N+1} \frac{a_{ji}}{k_j^{out}} s_j(t)$	6	Lü et al. (2011)
7	Normalized alpha	$C_{N_\alpha} = \frac{C_{alpha,\alpha}}{\sum_{i,j}^{n}(\alpha^t A^t)_{i,j}}$	4	Ghosh and Lerman (2010)
8	PageRank	$C_{pr} = \alpha \sum_{j=1}^{n} A_{j,i} \frac{C_p(v_j)}{d_j^{out}} + \beta$	2	Kwak et al. (2010)
9	Weighted in-degree	$C_{d_{win}} = d_i^{w(in)}$	1	Almgren and Lee (2015)

Table 3 Recall and precision results of measurements with the empirical measurement for Flicker's dataset

Influence measurements	Recall precision	The number of returned candidates of influential users (K)								
		1	10	20	30	40	50	60	70	72
Hybrid eigenvector	Eigenvector	0.01	0.04	0.13	0.27	0.38	0.55	0.76	0.93	0.95
	Precision	1	0.3	0.5	0.66	0.7	0.8	0.9	0.95	0.95
Weighted in-degree	Recall	0.01	0.1	0.19	0.22	0.25	0.27	0.31	0.34	0.34
	Precision	1	0.9	0.7	0.53	0.45	0.4	0.38	0.35	0.34
In-degree	Recall	0.01	0.125	0.180	0.22	0.25	0.25	0.29	0.305	0.33
	Precision	1	0.9	0.65	0.53	0.45	0.36	0.35	0.31	0.33
Normalized alpha	Alpha	0.01	0.1	0.18	0.22	0.23	0.25	0.27	0.30	0.31
	Precision	1	0.8	0.65	0.53	0.42	0.36	0.33	0.31	0.31
LeaderRank	Recall	0.01	0.1	0.18	0.22	0.23	0.25	0.27	0.30	0.31
	Precision	1	0.8	0.65	0.53	0.42	0.36	0.33	0.31	0.31
Complex	Recall	0.01	0.06	0.12	0.18	0.22	0.27	0.39	0.31	0.31
	Precision	1	0.5	0.45	0.43	0.4	0.4	0.38	0.32	0.31
Degree	Recall	0.01	0.1	0.180	0.22	0.23	0.25	0.27	0.30	0.31
	Precision	1	0.8	0.65	0.53	0.42	0.36	0.33	0.31	0.31
Eigenvector	Recall	0.01	0.069	0.12	0.18	0.22	0.27	0.39	0.319	0.31
	Precision	1	0.5	0.45	0.43	0.4	0.4	0.38	0.32	0.31
PageRank	Recall	0	0.11	0.18	0.22	0.23	0.25	0.29	0.31	0.31
	Precision	0	0.8	0.65	0.53	0.42	0.36	0.35	0.32	0.31

Table 4 Recall and precision results of measurements with the empirical measurement for Digg's dataset

Influence measurements	Recall Precision	The number of returned candidates of influential users (K)								
		1	10	20	30	40	50	60	70	74
Hybrid eigenvector	Recall	0	0.04	0.10	0.17	0.29	0.44	0.66	0.86	0.98
	Precision	0	0.3	0.4	0.43	0.55	0.66	0.81	0.91	0.98
Weighted in-degree	Recall	0.013	0.135	0.27	0.40	0.54	0.60	0.74	0.78	0.82
	Precision	1	1	1	1	1	0.9	0.91	0.82	0.82
Complex	Recall	0.013	0.135	0.27	0.40	0.51	0.60	0.72	0.78	0.81
	Precision	1	1	1	1	0.95	0.9	0.9	0.82	0.81
PageRank	Recall	Alpha	0.015	0.27	0.40	0.51	0.60	0.73	0.78	0.81
	Precision	1	1	1	1	0.95	0.9	0.9	0.82	0.81
In-degree	Recall	0.013	0.135	0.27	0.39	0.51	0.59	0.70	0.75	0.79
	Precision	1	1	1	0.96	0.95	0.88	0.86	0.8	0.79
Normalized alpha	Recall	0.013	0.135	0.27	0.39	0.51	0.60	0.70	0.75	0.79
	Precision	1	1	1	0.96	0.95	0.9	0.86	0.8	0.79
LeaderRank	Recall	0.013	0.135	0.27	0.39	0.51	0.60	0.70	0.75	0.79
	Precision	1	1	1	0.96	0.95	0.9	0.86	0.8	0.79
Degree	Recall	0.013	0.135	0.27	0.39	0.51	0.59	0.70	0.75	0.79
	Precision	1	1	1	0.96	0.95	0.88	0.86	0.8	0.79
Eigenvector	Recall	0	0.06	0.18	0.29	0.45	0.52	0.67	0.75	0.79
	Precision	0	0.5	0.7	0.73	0.85	0.78	0.83	0.8	0.79

are assigned to AccGr2. On the other hand, hybrid eigenvector improves the accuracy rate significantly and achieves a medium accuracy rate of 0.69. We hypothesize that on Flicker, users' characteristics, such as activeness, are important for users to become influential. Therefore, a user needs to be more active, and hybrid eigenvector is the only measurement that considers the users' characteristics.

For Digg's dataset, hybrid eigenvector and eigenvector have medium accuracy rates, while other measurements have strong accuracy rates. This shows that on Digg, the user's structural location is more important than the user's characteristics.

We hypothesize that these two observations reflect the nature of Flicker and Digg. Digg acts as a news medium where the structural location of the story channel, i.e., user, is very important in spreading news. For example, a news channel that is widely known and has a huge audience has a better chance of spreading the news more widely. On the other hand, Flicker is focused more on social activities between users; therefore, it is more emotional and that is why a user's characteristic is more important than the user's structural location. See Table 5, for the accuracy results. Also, to illustrate the performance of the ranking, we draw the precision and recall curve for each measurement (see Figs. 2, 3).

7.4 Experiment 2: correlation analysis

In order to measure the statistical significance between the rankings produced by the influence measurements and the ground truth, we use Pearson's correlation coefficient. Pearson's correlation coefficient measures how much two variables are related to each other by measuring the linear dependence between them. The possible outcomes from Pearson's correlation are a value, i.e., r, between $[-1, 1]$ representing the negative and positive relationship strength, respectively. A strong correlation exists if $-0.5 \leq r \geq +0.5$ where medium correlations exist when $\pm 0.3 \leq r \geq \pm 0.5$; otherwise, there is weak relationship or no relationship when $r = 0$ (Benesty et al. 2009). PCC is calculated as follows:

$$r_{EM,PRD} = \frac{\sum_{i=0}^{n}(EM_i \ PRD_i) - n \overline{EM} \ \overline{PRD}}{n \ \sigma_{EM} \ \sigma_{PRD}} \tag{24}$$

The results shown in Table 6 are grouped into three groups based on the correlation results for each dataset, i.e., Gr1, Gr2 and Gr3. These groups contain weak, medium and strong correlations, respectively, according to Benesty et al. (2009).

For Flicker's social network, degree centrality is assigned to Gr1, where hybrid Eigenvector is assigned to

Table 5 The area under the recall–precision graph for each influence measurement

Influence measurements	Area	
	Flicker dataset	Digg dataset
Weighted in-degree	0.22	0.77
Hybrid eigenvector	0.69	0.65
Complex	0.14	0.76
LeaderRank	0.20	0.73
Eigenvector	0.14	0.56
Normalized alpha	0.20	0.73
PageRank	0.15	0.76
Degree	0.19	0.73
In-degree	0.21	0.73

Table 6 Correlations of measurements with the empirical measurement

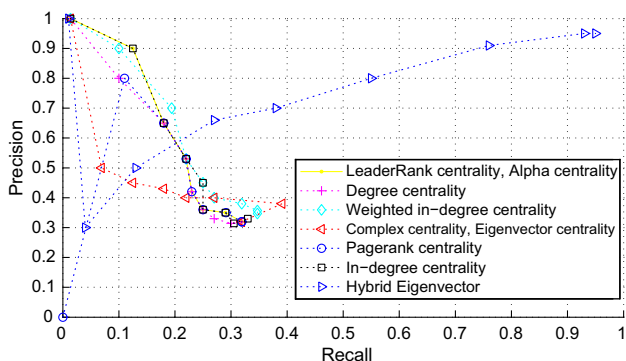| Influence measurements	r	
	Flicker dataset	Digg dataset
Hybrid eigenvector	0.50	0.90
Weighted in-degree	0.42	0.74
Complex	0.38	0.70
LeaderRank	0.327	0.69
Eigenvector	0.323	0.66
Normalized alpha	0.322	0.67
PageRank	0.316	0.70
In-degree	0.312	0.66
Degree	0.27	0.67

Fig. 2 The precision–recall graphs for the influence measurements applied to Flicker's dataset. The *axis* represents the recall and precision for each K
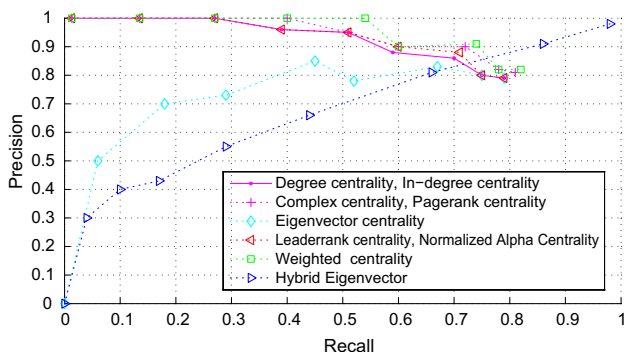

Fig. 3 The precision–recall graphs for the influence measurements applied to Digg's dataset. The *axis* represent the recall and precision for each K

Gr3. The rest of the measurements are assigned to Gr2. This can show that users' characteristics and number of followers are important factors for determining influential users since hybrid eigenvector is the only measurement that

considers these characteristics. This supports the results from the accuracy analysis. On the other hand, degree centrality has a weak correlation because reciprocal relationships may not be important on Flicker. Degree is the only measurement that considers reciprocal relationships. Other measurements show medium correlations with similar rates. The only common thing between these measurements is that they all consider the in-degree of nodes, which shows that the number of followers each user has is an indicator of influence.

On Digg's dataset, all measurements are classified to Gr3. However, one measurement shows a very strong correlation where the other measurements have similar correlations. Therefore, we further divide Gr3 into Gr3.1 and Gr3.2 for very strong and strong correlations, respectively. Hybrid eigenvector is classified to Gr3.1 where other measurements are classified to Gr.3.2. Results for Digg's dataset have a similar trend to the results on Flicker. However, all correlation rates have significantly increased. Also, measurements that were classified to Gr.1 and Gr.2 have jumped to Gr.3.

7.5 Experiment 3: performance analysis

To assess the performance of the influence measurements, we use a sorted adjacency list to build the social network because the two datasets exhibit sparse graphs (McConnell and Spinrad 1994). The sparse graph is a graph that has less edges than a normal graph (Black 2004). Directed graphs can have a maximum of $n(n - 1)$ edges where n is the number of nodes (Black 2004). The *Flicker Dataset* can possibly have 946,085,322 edges but it only has 46,059. In addition, the *Digg dataset* has only 1,534,314, but it can have up to 19,434,729,872 edges. Therefore, we use a sorted adjacency list for efficiency.

Table 7 Performance of measurements, n = nodes, m = edges

Influence Measurements	Runtime complexity	Space complexity	I/O cost	Runtime in milliseconds (Flicker dataset)	Runtime in milliseconds (Digg dataset)
Weighted in-degree	$O(m)$	$O(m + n)$	$O(m + n)$	10	40
In-degree	$O(m)$	$O(m + n)$	$O(m + n)$	10	40
Degree	$O(m)$	$O(m + n)$	$O(m + n)$	10	45
Complex	$O(m)$	$O(m + n)$	$O(m + n)$	30	225
PageRank	$O(m)$ per iteration	$O(m + n)$ per iteration	$O(m + n)$	100 (iterations = 55)	6045 (iterations = 55)
LeaderRank	$O(m)$ per iteration	$O(m + n)$ per iteration	$O(m + n)$	345 (iterations = 55)	5435 (iterations = 55)
Eigenvector	$O(m)$ per iteration	$O(m + n)$ per iteration	$O(m + n)$	170 (iterations = 55)	3349 (iterations = 55)
Hybrid eigenvector	$O(m)$ per iteration	$O(m + n)$ per iteration	$O(m + n)$	178 (iterations = 55)	3396 (iterations = 55)
Normalized alpha	$O(m)$ per iteration	$O(m + n)$ per iteration	$O(m + n)$	26,876 ($\alpha < 0.1$)	40,056 ($\alpha < 0.05$)

The measurements are grouped into iterative and non-iterative. C_{din}, C_d, C_c and C_{dwin} are non-iterative algorithms, where the rest of the measurements are iterative algorithms. As shown in Table 7, non-iterative measurements have a linear runtime complexity of $O(m)$ since they only need to compute the degree of each node once it is given m edges. They have a space complexity of $O(m + n)$ because they store the ranking of each node n based on its adjacent edges m. Iterative measurements have the same exponential runtime and space complexity of $O(m)$ and $O(m + n)$ per iteration, respectively, since they need to compute and store the ranking list for each iteration. All of the measurements have I/O cost of $O(m + n)$. To obtain sense of their complexities in the implementation, we have computed the runtime of each measurement in the two datasets. The results confirmed the performance analysis in terms of O notation.

In Flicker Dataset, C_{dwin}, C_{din} and C_d take 10 ms to complete using wall clock time, which makes them the fastest measurements. C_c took slightly longer runtime because it computes the exponent for each node. These influence measurements are well suited for large-scale social networks.

For C_{eg}, C_{pr} and C_{ld}, we limited the number of iteration to 55 since they are already proven to converge. C_{pr} takes 100 ms, where C_{eg} is executed in 170 ms. $C_{HEG}(v_i)$ takes slightly more time than C_{eg} since they are both based on computing the eigenvector. C_{ld} completes in 345 ms. $C_{N\alpha < 0.1}$ takes 26,876 ms since it iterates much more than the previous measurements. There was not a lot of variation in the performance of each measurement in the Digg dataset. C_d took more time than C_{dwin} and C_{din}. C_{eg} was the fastest iterative measurement where C_{ld} was the second

fastest iterative measurement. $C_{N\alpha < 0.05}$ again took the longest running time of 40,056 ms. These results reflect the complexities of influence measurements. Table 7 shows the summary of the performance evaluation for the selected influence measurements in term of complexity.

Table 7 shows the runtime complexity that represents the number of steps to run the algorithm, and the space complexity shows the computational resources needed by the algorithms. I/O is the cost of input and output management. The runtime in both datasets represents the actual time spent by each algorithm.

8 Discussion

In this section, we will discuss further the experimental results and their limitations.

8.1 Accuracy analysis

The accuracy of the measurements has significantly increased in Digg's dataset compared to Flicker's dataset. We hypothesize that this is because of the social network characteristics and nature; for example, Flicker is a social network that supports images, while Digg acts like a new medium. This means that feasibility of influence measurements depends on their adaptability to social networks. Therefore, when measuring influence on social networks, the nature of the social network must be considered.

C_{HEG} performs very well in both datasets where other measurements achieve very low accuracy rates on the Flicker dataset. This shows that C_{HEG} is stable in different social networks. This can also show that further

investigation should be done to find the critical features that optimize the results for other measurements such as user's experience. However, C_{HEG} accuracy rate has decreased when employed in the Digg's dataset. This supports our hypothesis that news social networks are more focused on users' structural locations. We suggest that when predicting influential users on news social networks, there is a limited need to consider users' characteristics.

C_{dwin} is the most accurate measurement in Digg. We hypothesize that this is because C_{dwin} and the empirical measurement both consider users' votes. C_{din} came in the third place when applied to Flicker's dataset. This is not surprising because C_{din} is a generalization of the C_{dwin}. However, C_{din} is surprisingly ranked as the third accurate measurement in Digg's dataset. We hypothesize that this is because there are many users with many followers that do not vote in the Digg dataset.

C_c and C_{pr} are the fourth most accurate measurements in the Flicker's dataset, where they are the second most accurate measurements in Digg's dataset. C_c considers both users' followers and followings, and their number of votes, where the empirical measurements also consider the followers' votes. C_{pr} considers users with important followers.

As we see, the order of the most accurate measurements slightly changes in the two datasets. Only C_c, C_{pr}, C_{HEG} and C_{din} positions have changed. C_{din} dropped one position in Digg's dataset, while C_c and C_{pr} jumped one position in Digg's dataset. Therefore, we can say that C_{dwin}, $C_{N\alpha}$, C_{eg}, C_{ld} and C_d are more robust than other measurements in terms of social networks since their order does not change. Figures 2 and 3 support our results and claims. The results from both datasets show a similar trend.

8.2 Correlation analysis

From Table 6, we notice that C_{HEG} and C_{dwin} are the most correlated measurements with the empirical measurement in both datasets. First as discussed earlier, we hypothesized that users' characteristics are important indicators of influence, which is one of the bases of C_{HEG}. On the other hand, C_{dwin} considers the number of followers and their votes, which represents the tie strength.

C_c is the second most correlated measurement in both datasets. It is due to the fact that it considers the tie strength as well. C_{ld} is the third top correlated measurement with the empirical measurement in both datasets. C_{eg} is the fourth most correlated measurements with the empirical measurement in Flicker Dataset, while it is the fifth most correlated measurement in the Digg's dataset. $C_{N\alpha}$ is the fifth most correlated measurement in Flicker's dataset, but its correlation increases to the top fourth correlated

measurement in Digg's dataset. C_{pr} is the sixth most correlated measurement in Flicker's dataset, while it is the second most correlated measurement in Digg's dataset. The previous four measurements have a similar correlation rate since they all consider the depth of the social network. However, they perform differently in both datasets, which is not surprising because of their accuracy results. C_{din} is the seventh most correlated measurement in Flickr's dataset, while it is the second most correlated measurement in Digg's dataset. C_d is the least correlated measurement in the Flicker's dataset, while it is the fourth top correlated measurement in Digg's dataset.

8.3 Performance analysis

The iterative measurements are more costly than the non-iterative measurements in terms of performance. However, but iterative measurements consider the graph depth where iterative measurements do not consider that. In both datasets, users who receive many interactions based on weighted in-degree are high ranked by the iterative measurements as well. This can show that users with many followers can also be dominant in terms of depth. The accuracy and correlation results show slight changes for iterative and non-iterative measurements. Therefore, future research can adopt non-iterative measurements since the size of social networks is increasing, and also several non-iterative measurements such as weighted in-degree measurements perform well.

8.4 Limitations

In our experiment, the datasets considers the dynamic process of activities such as voting, favoring and commenting. That means if a user never interacts with other users, the user will be ignored. In addition, we assume that all the selected measurements are hybrid measurements as mentioned in Sect. 5, i.e., content as well as context. However, some of them only consider context in their algorithms such as C_{din}, C_{ld} and C_{pr}. We argue that the accuracy of these measurements will be increased if we consider both content and context.

In addition, we employ the selected measurements to the entire social networks that contain disconnected and connected components. However, the giant component, which means the largest connected component in a graph, for both datasets contains more than 80 % of the total number of nodes. For example, in the Flicker dataset, the giant component contains 99.36 % of the total number of nodes. Therefore, recursive measurements such as normalized alpha may perform better if they are employed to connected components only.

9 Conclusion and future directions

In this paper, we highlight the recent research on social influence measurements. The social influence measurements are standardized using our classification. These classifications consider three main aspects on which the measurements are based on. The first category is based on the subset of characteristics that these measurements use, for example, users' attributes or structural location on social networks. The second category explains the type of structures that the measurements consider in measuring influence such as content. The third category is the algorithms that are used in developing these measurements. This classification can be used as a baseline to explain the influence measurements and can help future research in proposing influence measurements. We further evaluate the existing measurements in terms of accuracy, performance and correlations using two different datasets from real-world social networks to evaluate the feasibility of influence measurements to different social networks.

Our results show that the structural location of users is more important than users' attributes in predicting influential users on Digg, a news medium social network, while users' attributes are more important in predicting influential users on Flicker, an entertainment social network.

Moreover, integrating both user's structural location and characteristics shows stable performance in different social networks. Influence measurements are affected by the nature of social networks. Therefore, it is necessary to consider the nature of social networks when measuring influence. We observe that many of the existing measurements consider only the users' structural location in social networks. However, there are other characteristics that should be considered including users' attributes. Moreover, weighted in-degree, a non-iterative influence measurement, is very effective in predicting influential users and is well suited for large-scale networks.

For future work, it would be interesting to use more social networks to investigate the adaptability of influence measurements. It is also important to consider the nature of social networks as a feature for predicting influential users on social networks. By making available the datasets and code, we hope to engage other researchers into the challenging quest of predicting influential users on social networks.

Appendix

See Table 8.

Table 8 Classification of the state of art by their definitions, models, number of measurements, types, algorithms and datasets

Index	Influence definition(s)	Influence model(s)	Number of measurements	Measurement Model(s)	Measurement type(s)	Measurement algorithms(s)	Dataset	References
1	Average number of votes	Graph based	1	Location model	Hybrid	Centrality analysis	Digg	Ghosh and Lerman (2010)
2	Number of posts that generate many implicit and explicit interactions	Graph based	3	Location model	Hybrid	Centrality analysis, network algorithms	Twitter	Sun and Ng (2013)
3	Amount of interactions that posts generate	Graph based	1	Location model	Hybrid	Centrality analysis	Flicker	Almgren and Lee (2015)
4	Number of topics posted in tweets that further generate many other tweets	Graph based	1	Location model	Hybrid	Centrality analysis	Twitter	Weng et al. (2010)
5	Number of followers, number of users that engage in conversations, the diffusion size of tweets	Graph based	3	Location model, attribute model, diffusion model	Context, hybrid, content	Centrality analysis, social network properties	Twitter	Cha et al. (2010)
6	Number of followers, number of important neighbors, number of tweets that are red by many people	Graph based, tree based	3	Location model, role model	Context, content	Centrality analysis, information cascade modeling	Twitter	Lee et al. (2010)

Table 8 continued

Index	Influence definition(s)	Influence model(s)	Number of measurements	Measurement Model(s)	Measurement type(s)	Measurement algorithms(s)	Dataset	References
7	Number of important neighbors that perform interactions	Graph based	1	Location model	Context	Centrality analysis	Delicious	Lü et al. (2011)
8	Number of important followers based on interactions	Graph based	1	Location model	Hybrid	Centrality analysis	Weibo	Liao et al. (2013)
9	Number of actions from important neighbors	Graph based	1	Location model	Hybrid	Centrality analysis	Weibo	Zhang et al. (2012)
10	Number of important followers, number of retweets	Graph based, tree based	2	Location model	Hybrid	Centrality analysis, social network properties	Twitter	Kwak et al. (2010)
11	Number of people that are engaged in conversation	Graph based	1	Location model	Hybrid	Centrality analysis	Weibo	Li et al. (2013b)
12	Number of important followers based on interactions, size of tweets propagation	Linear based	2	Attribute model, diffusion model	Hybrid	Social network properties	Twitter	Leavitt et al. (2009)
13	Number of tweets that make users popular or make them active, diffusion size of tweets	Linear based	2	Attribute model, diffusion model	Context, content	Social network properties	Twitter	Anger and Kittl (2011)
14	Certain personal attributes	Linear based	2	Attribute model	Hybrid	Centrality analysis	Twitter	Yi et al. (2013)
15	Diffusion size of URL on Twitter	Tree based	2	Diffusion model	Content	Information cascade modeling	Twitter	Bakshy et al. (2011)
16	Diffusion size of content	Tree based	2	Diffusion model	Content	Information cascade modeling	Twitter	Reilly et al. (2014)
17	Number of interactions between users	Graph based	2	Location model	Hybrid	Centrality analysis	Twitter	Maharani et al. (2014)
18	Number of effective audience	Graph based, tree based	6	Role model	Context	Information cascade modeling	Sina	Sun et al. (2013)
19	Ability to have certain attributes and be connected to many users	Graph based	1	Location and attribute model	Hybrid	Centrality analysis	Flicker	Almgren and Lee (2016)

References

Aggarwal CC (2011) An introduction to social network data analytics. Springer, New York

Almgren K, Lee J (2015) Who influences whom: content-based approach for predicting influential users in social networks. In: International conference on advances in big data analytics, pp 89–99

Almgren K, Lee J (2016) Applying an influence measurement framework to large social network. J Netw Technol 7(1):7

Anger I, Kittl C (2011) Measuring influence on twitter. In: Proceedings of the 11th international conference on knowledge management and knowledge technologies. ACM, p 31

Bakshy E, Hofman JM, Mason WA, Watts DJ (2011) Everyone's an influencer: quantifying influence on twitter. In: Proceedings of the fourth ACM international conference on Web search and data mining. ACM, pp 65–74

Benesty J, Chen J, Huang Y, Cohen I (2009) Pearson correlation coefficient. In: Noise reduction in speech processing. Springer, pp 1–4

Black PE (2004) Sparse graph. In: Dictionary of algorithms and data structures. National Institute of Standards and Technology

Bonacich P, Lloyd P (2001) Eigenvector-like measures of centrality for asymmetric relations. Soc Netw 23(3):191–201

Brandes U (2001) A faster algorithm for betweenness centrality*. J Math Sociol 25(2):163–177

Cha M, Haddadi H, Benevenuto F, Gummadi PK (2010) Measuring user influence in twitter: the million follower fallacy. ICWSM 10(10–17):30

Ellison NB et al (2007) Social network sites: definition, history, and scholarship. J Comput Mediat Commun 13(1):210–230

Etherington D (2014) Flickr at 10: 1m photos shared per day, 170 % increase since making 1tb free. February 10:25–45

Facebook (2015) Stats. http://newsroom.fb.com/company-info

Freeman LC (1978) Centrality in social networks conceptual clarification. Soc Netw 1(3):215–239

Ghosh R, Lerman K (2010) Predicting influential users in online social networks. arXiv preprint arXiv:10054882

Granovetter MS (1973) The strength of weak ties. Am J Sociol 78:1360–1380

Hogg T, Lerman K (2012) Social dynamics of Digg. EPJ Data Sci 1(1):1–26

Katz E, Lazarsfeld PF (1955) Personal influence. The part played by people in the flow of mass communications. Transaction Publishers, New Brunswick

Kempe D, Kleinberg J, Tardos É (2003) Maximizing the spread of influence through a social network. In: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, pp 137–146

Kwak H, Lee C, Park H, Moon S (2010) What is twitter, a social network or a news media? In: Proceedings of the 19th international conference on world wide web. ACM, pp 591–600

Leavitt A, Buchard E, Fischer D, Gilbert S (2009) The influentials: new approaches for analyzing influence on twitter. Web Ecol Project 4:1–18

Lee C, Kwak H, Park H, Moon S (2010) Finding influentials based on the temporal order of information adoption in twitter. In: Proceedings of the 19th international conference on World wide web. ACM, pp 1137–1138

Li H, Cui JT, Ma JF (2015) Social influence study in online networks: a three-level review. J Comput Sci Technol 30(1):184–199

Li J, Peng W, Li T, Sun T (2013a) Social network user influence dynamics prediction. In: Web technologies and applications. Springer, pp 310–322

Li X, Cheng S, Chen W, Jiang F (2013b) Novel user influence measurement based on user interaction in microblog. In: Proceedings of the 2013 IEEE/ACM international conference on advances in social networks analysis and mining. ACM, pp 615–619

Liao Q, Wang W, Han Y, Zhang Q (2013) Analyzing the influential people in sina weibo dataset. In: Global communications conference (GLOBECOM), 2013 IEEE. IEEE, pp 3066–3071

Lü L, Zhang YC, Yeung CH, Zhou T (2011) Leaders in social networks, the delicious case. PloS One 6(6):e21,202

Maharani W, Gozali AA, et al (2014) Degree centrality and eigenvector centrality in twitter. In: Telecommunication systems services and applications (TSSA), 2014 8th international conference on. IEEE, pp 1–5

McConnell RM, Spinrad JP (1994) Linear-time modular decomposition and efficient transitive orientation of comparability graphs. SODA 94:536–545

Merriam-Webster (2011) Influence. http://www.merriam-webster.com/dictionary/

Newman ME (2001) Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. Phys Rev E 64(1):016,132

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V et al (2011) Scikit-learn: machine learning in python. J Mach Learn Res 12:2825–2830

Pempek TA, Yermolayeva YA, Calvert SL (2009) College students' social networking experiences on facebook. J Appl Dev Psychol 30(3):227–238

Probst F, Grosswiele DKL, Pfleger DKR (2013) Who will lead and who will follow: identifying influential users in online social networks. Bus Inf Syst Eng 5(3):179–193

Rabade R, Mishra N, Sharma S (2014) Survey of influential user identification techniques in online social networks. In: Recent advances in intelligent informatics. Springer, pp 359–370

Reilly CF, Salinas D, De Leon D (2014) Ranking users based on influence in a directional social network. In: Computational science and computational intelligence (CSCI), 2014 international conference on, IEEE, vol 2, pp 237–240

Riquelme F (2015) Measuring user influence on twitter: a survey. arXiv preprint arXiv:150807951

Robertson S (2000) Evaluation in information retrieval. In: Lectures on information retrieval. Springer, pp 81–92

Sang J, Xu C (2013) Social influence analysis and application on multimedia sharing websites. ACM Trans Multimed Comput Commun Appl (TOMM) 9(1s):53

Scoot J (1992) Social network analysis theory and application. Sage, Newberry Park, CA

Singh S, Mishra N, Sharma S (2013) Survey of various techniques for determining influential users in social networks. In: Emerging trends in computing, communication and nanotechnology (ICE-CCN), 2013 international conference on, IEEE, pp 398–403

Stelzner MA (2011) Social media marketing industry report. Soc Media Exam 41:1–12

Sun B, Ng VT (2013) Identifying influential users by their postings in social networks. Springer, Berlin

Sun C, Zhang L, Li Q (2013) Who are influentials on microblogging services: evidence from social network analysis. In: PACIS, p 25

Tang L, Liu H (2010) Community detection and mining in social media. Synth Lect Data Min Knowl Discov 2(1):1–137

Twitter (2015) Twitter usage. https://about.twitter.com/company

Wang C, Chen W, Wang Y (2012) Scalable influence maximization for independent cascade model in large-scale social networks. Data Min Knowl Discov 25(3):545–576

Watts DJ, Dodds PS (2007) Influentials, networks, and public opinion formation. J Consum Res 34(4):441–458

Weng J, Lim EP, Jiang J, He Q (2010) Twitterrank: finding topic-sensitive influential twitterers. In: Proceedings of the third ACM international conference on Web search and data mining. ACM, pp 261–270

Yi X, Han Y, Wang X (2013) The evaluation of online social networks nodes influence based on users attribute and behavior. In: Frontiers in internet technologies. Springer, pp 9–20

Zafarani R, Abbasi MA, Liu H (2014) Social media mining: an introduction. Cambridge University Press, Cambridge

Zhang Y, Mo J, He T (2012) User influence analysis on micro blog. In: Cloud computing and intelligent systems (CCIS), 2012 IEEE 2nd international conference on, IEEE, vol 3, pp 1474–1478