Comparing Linkage Graph and Activity Graph of Online Social Networks

Yuan Yao[1,2], Jiufeng Zhou[3], Lixin Han[3], Feng Xu[1,2], and Jian Lü[1,2]

[1] State Key Laboratory for Novel Software Technology, Nanjing, China
[2] Department of Computer Science and Technology, Nanjing University, China
YuanY@ics.nju.edu.cn, {xf,lj}@nju.edu.cn
[3] Department of Computer Science and Technology, HoHai University, China
{jfzhou,lhan}@hhu.edu.cn

Abstract. In the context of online social networks, the *linkage graph*—a graph composed of social links—has been studied for several years, while researchers have recently suggested studying the *activity graph* of real user interactions. Understanding these two types of graphs is important since different online applications might rely on different underlying structures. In this paper, we first analyze two specific online social networks, one of which stands for a linkage graph and the other for an activity graph. Based on our analysis, we find that the two networks exhibit several static and dynamic properties in common, but show significant difference in degree correlation. This property of degree correlation is further confirmed as a key distinction between these two types of graphs. To further understand this difference, we propose a network generator which could as well capture the other examined properties. Finally, we provide some potential implications of our findings and generator.

Keywords: Linkage Graph, Activity Graph, Online Social Networks, Degree Correlation, Network Generator.

1 Introduction

Researchers have made remarkable achievements in analyzing structural properties of the *linkage graph*, i.e., a graph composed of social links [10,15,1]. Several applications have used these properties, for example, to protect against Sybils [25] or to prevent unwanted communication [16,21]. Recently, researchers have suggested studying the *activity graph* of real user interactions instead, in order to enhance social network based applications [24,22]. We define linkage graph as a graph where nodes stand for the people in the social network and edges are their friendship links. Activity graph is correspondingly defined as a graph where nodes are the still the people but edges stand for their interactions.

Understanding the linkage graph and the activity graph, as well as their similarities and differences, is important for developing future online applications. Wilson et al. [24] have shown that, if operating on the activity graph instead of the linkage graph, the RE application [7] actually performs better while the

A. Datta et al. (Eds.): SocInfo 2011, LNCS 6984, pp. 84–97, 2011.

SybilGuard system [25] behaves less effectively. In addition, different online applications might rely on different underlying structures. We identify two categories of applications: *linkage-based applications* and *activity-based applications*. Linkage-based applications need available links, for example, to disseminate information, and therefore should be built on the linkage graph. By comparison, activity-based applications, such as trust inference, are based on a history of interactions, and therefore should be built on the activity graph.

Existing analysis on online social networks tends to experiment on bidirectional or undirected graphs, or map directed graphs to undirected graphs by removing the unidirectional edges (e.g. [10]). However, simply removing the unidirectional edges may result in a loss of information. Additionally, some applications rely on unidirectional edges, such as edges representing trust which is asymmetric in nature [8]. In view of this, we primarily focus on the analysis of directed networks.

In this paper, we mainly study the two graphs mapped from online social networks: the Flickr network which consists of social links, and the Epinions network of user interactions. These two graphs are both directed, and they both have timestamps on every edge for us to study dynamic properties. Dynamic properties are important as many online social networks keep evolving over time. To this end, in addition to analyzing several common static properties, we also explore some dynamic properties including densification, clustering, and diameter over time. Our results show that the two networks follow some common static properties, i.e., they both exhibit power-law degree distribution, high clustering coefficient, and small diameter. As to dynamic properties, we find that both networks follow densification law and relatively stable clustering coefficient over time. However, we do not observe diameter shrinking in Epinions' activity graph, while this shrinking exists in Flickr.

One of our major findings is the difference in degree correlation, also known as degree mixing [17], of the two graph types. Traditional social networks have essentially positive degree correlation, indicating that gregarious people tend to make friends with each other. This property can also differentiate social networks from other networks such as technological and biological networks [17]. However, this positive degree correlation does not always hold in online social networks. Based on additional experiments, we find that linkage graphs still have positive degree correlation whereas activity graphs show neutral degree correlation. We then confirm degree correlation as a key distinction between activity graphs and linkage graphs.

To further understand and capture the difference of degree correlation, we propose a network generator which could capture other static properties and dynamic properties as well. Our generator has only two parameters and follows the intuition that online linkage graphs have a high reciprocity relative to activity graphs. In addition to understanding the underlying factors of the emerged properties, the generator can also be used to evaluate algorithms on networks [8], and to generate realistic graphs when real data is unavailable or useless.

Our findings and generator, while preliminary, could provide potential implications for a variety of online applications, including viral marketing, social search, and trust and reputation management. In this paper, we concentrate on two specific applications from the linkage-based applications and the activity-based applications, respectively.

The rest of the paper is organized as follows. Section 2 covers related work on property analysis and network generator. Section 3 presents our results of networks in several aspects, including static and dynamic properties. Section 4 presents our network generator and discusses its simulation results. Section 5 discusses some implications of our findings and generator for different applications. Section 6 concludes the paper.

2 Related Work

There is a rich set of research on analyzing the structural properties of graphs mapped from online social networks. Kumar et al. have studied the datasets of Flickr and Myspace, and they found that the networks are composed of three groups: singletons, giant component, and isolated communities [10]. They also analyzed how nodes evolved among these three groups. For example, isolated communities might merge into the giant component. However, they mapped the networks to undirected graphs by leaving out the unidirectional edges. In contrast, all the measurements we choose to study are based on directed graphs.

Mislove et al. have taken a set of measurements on Flickr, LiveJournal, Orkut, and YouTube [15]. They found that power-law exists in both out-degree and in-degree distribution, and nodes with high out-degree tend to have high in-degree. They also found that high clustered nodes are usually of low degree, and the clusters connect to each other through a relatively small number of high-degree nodes. Different from their work, we put special emphasis on the measurements of degree correlation which is confirmed as a key indicator to distinguish linkage graphs and activity graphs.

Ahn et al. have analyzed the Cyworld network, and observed a multi-scaling behavior in degree distribution [1]. In addition, they compared the explicit linkage graph and the implicit activity graph constructed by messages exchanged on Cyworld's guestbook. They only focus on static properties, while we also concern dynamic properties.

Wilson et al. have studied the activity graph and linkage graph of Facebook [24]. Similar to Ahn et al. [1], they built the activity graph through real interactions between friends, and compared it to the linkage graph. However, findings based on this technique might be biased, because the extracted activity graph is actually a sub-graph of the linkage graph. Viswanath et al. also checked the activity graph of Facebook [22]. They found that although the individuals' behavior changed rapidly, the properties of the network as a whole remained relatively stable. Different from the preceding work, we compare the linkage graph and activity graph based on two distinct datasets to eliminate bias.

A parallel body of work focuses on network generators. We give a brief history of the network generators based on our examined properties, and detailed discussion can be found in [5].

The BA generator [2] effects on a large body of later generators. The idea of preferential attachment from the model is thought to be the cause of power-law degree distribution. New nodes connect to existing high-degree nodes with greater probability, and make their degree even higher, forming the heavy tail of the degree distribution. As to the clustering of networks, a well-known idea is perhaps derived from the edge copying generator [9]. Its basic idea is to copy the links of an existing node with some probability when a new node arrives. This idea is analogical to the process of making friends in human communities. The small world generator [23] is another well-known generator that can meet the clustering property of social networks. Based on a ring lattice and the rewiring process, networks generated by this generator also have low average distance between nodes. However, all the preceding generators do not exhibit the dynamic properties of networks.

Forest fire generator [11] could meet a set of dynamic and static properties of social networks. However, when trying to generate the Flickr network, the generated network does not have a positive degree correlation while other properties are met. Based on our experiments, this positive degree correlation is a key distinction between activity graphs and linkage graphs of online social networks. Our generator incorporates reciprocity and the fire burning idea from the forest fire generator. The results show that our generator can generate networks with positive degree correlation, and at the same time capture other examined properties. The two generators could complement with each other, as forest fire generator could meet several examined properties with neutral degree correlation.

3 Structural Property Analysis

In this section, we study several static and dynamic properties of the two graph types. We first describe our chosen datasets, and then present our results of these properties, including degree distribution, clustering coefficient, and diameter. After that, we give emphasis to the degree correlation property which is further confirmed as a key indicator to distinguish linkage graphs and activity graphs. Overall, we find that the two graphs are surprisingly similar to each other in many properties except degree correlation.

3.1 Datasets

The Flickr dataset [15,14] consists of a linkage graph with each edge representing a social friendship link. This data is continuous crawled for 104 days, form February 3rd, 2007, to May 18th, 2007. According to the timestamps, we cut the data into 27 snapshots over time. The first snapshot is the initial graph of February 3rd, 2007, and we will refer to this snapshot as G_{F0}. Each of the remaining snapshots consists of four more days' data.

Table 1. High level statistics of the two chosen online social networks as directed graphs

Graph	Flickr	Epinions
Initial nodes	1,834,425	93,598
Initial edges	24,828,241	588,616
Final nodes	2,570,535	131,828
Final edges	33,140,018	841,200
Time span	104 days	31 months
Snapshots	27	31
Node growth	40,13%	40.84%
Edge growth	33.48%	42.91%
Average node growth per snapshot	1.54%	1.36%
Average edge growth per snapshot	1.29%	1.43%

The Epinions dataset [13] consists of an activity graph as each edge stands for a user interaction. We use the data from January 12nd, 2001 to August 12nd, 2003. Similar to Flickr, we cut the Epinions data into 31 snapshots. The first snapshot is the initial graph of January 12nd, 2001, and we will refer to this snapshot as G_{E0}. Every additional month of the remaining data forms a snapshot of the network. High level statistics of the two graphs can be seen in Table 1. Although the datasets have different time span, their growth rates are similar to each other. For the sake of simplicity, we will only give the results of G_{F0} and G_{E0} for static analysis.

3.2 Static Properties

Power-law degree distribution [2] indicates that the count of nodes with degree k, versus the degree k, forms a line on a log-log scale. This skewed distribution is very common in social networks, and we observe this distribution in both graphs. The degree of Epinions is about one order of magnitude smaller than that of Flickr, but the power-law coefficient of the four distributions are quite close to each other. To verify this, we use the fitting method described by Clauset et al. [6] and apply it to the four distributions. We do not give the figures of these static properties due to space limit, but the statistics can be seen in Table 2.

Clustering coefficient is widely used as an indicator of community structure [5], or the transitivity of a graph [18]. We use the definition of clustering

Table 2. Statistics of static properties on G_{F0} and G_{E0}

Graph	In-degree coef.[1]	Out-degree coef.[1]	Clustering Coef.	Diameter
G_{F0}	$1.76, 0.0122$	$1.76, 0.0188$	0.2090	7.5173
G_{E0}	$1.78, 0.0287$	$1.75, 0.0146$	0.2357	5.8646

[1]Values in these columns are the power-law coefficient estimate and corresponding Kolmogorov-Smirnov goodness-of-fit, respectively.

coefficient by Watts and Strogatz [23]. The two networks both have a high clustering coefficient, and as Table 2 shows, the global clustering coefficient value is 0.2090 for G_{F0} and 0.2357 for G_{E0}. As to the local clustering coefficient (not shown here), nodes with lower degree tend to cluster more compactly at the edge of the graph, and nodes with higher degree stay in the middle of the graph to maintain high connectivity.

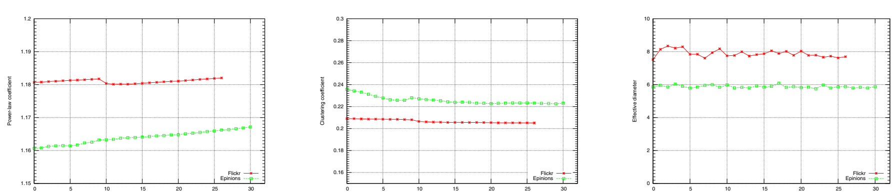
Diameter can be are calculated by many approaches, such as characteristic path length, average diameter, and effective diameter. The effective diameter [20] is defined as the minimum number of hops in which 90% of all nodes can reach each other, and we use the smoothed version described by Leskovec et al. [11]. To reduce the randomness impact of the effective diameter algorithm, every diameter value in this paper is an average of four calculations. Our results show that, G_{F0} has an effective diameter of 7.5173, and the value for G_{E0} is 5.8646.

Unlike the results by Wilson et al. [24], we find that the two kinds of graphs are quite similar in examined static properties. Although the diameter of Flickr is a little bigger, we find later in the dynamic analysis that the Flickr network is in a diameter shrinking stage while Epinions are much stable in diameter. In addition, Flickr and Epinions both are small-world networks, with high clustering coefficient and low diameter.

3.3 Dynamic Properties

The dynamic properties we study include the densification law and diameter shrinking discovered by Leskovec et al. [11], and the clustering coefficient over time. *Densification law* indicates that the network is getting denser, and the densification follows a power-law pattern: $E(t) \sim N(t)^{\alpha}$, where $E(t)$ and $N(t)$ are the edge and node numbers at time t, respectively. We find both graphs exhibit densification over time. This means both graphs are in a stage when more and more edges are created [10]. Moreover, we find that the coefficient of the densification power-law increases a little over time (See Fig. 1(a)).

We also check the *clustering coefficient and effective diameter over time* of Flickr and Epinions, as shown in Fig. 1(b) and Fig. 1(c). We find that the



(a) The densification power-law coefficient over time

(b) The clustering coefficient over time

(c) The effective diameter over time

Fig. 1. Some dynamic properties on Flickr and Epinions with horizontal axis representing the snapshot number over time

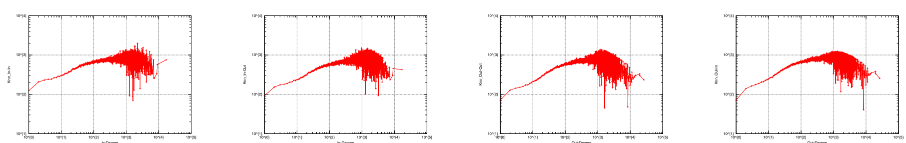
clustering coefficient is relatively stable with a slight decline over time in both graphs, and the diameter shrinking only appears in Flickr, while Epinions exhibit a stable effective diameter over time. These two results are quite consistent with the results by Viswanath et al. [22], who also find their activity graph strikingly stable in clustering coefficient and diameter. In addition, the shrinking phenomenon of diameter is also found in earlier work [11,10].

Analyzing dynamic properties can help to predict network growth, as well as assess the quality of graph generators. Overall, except for the slight difference in diameter over time, the two graphs are again similar to each other in examined dynamic properties.

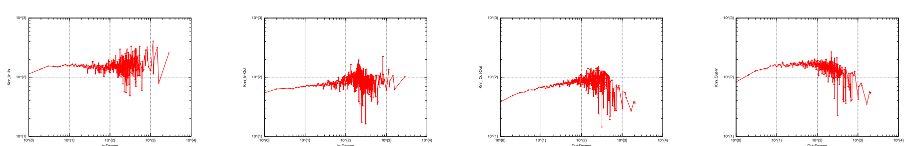
3.4 Degree Correlation

Degree correlation reflects how frequently nodes with similar degrees connect to each other. Degree correlation of a graph can be measured by k_{nn} distribution and corresponding assortativity coefficient r. k_{nn} of an undirected graph is a mapping between degree k and the average degree of all neighbors connected from nodes of that degree k. Assortativity coefficient r, ranging between -1 and 1, gives a quantitative measurement of degree correlation. For example, positive r value indicates a preference of high-degree nodes connecting to each other, and a random graph's r value should be 0 in theory.

We can define four kinds of k_{nn} and assortativity r on our directed graphs. As an example, k_{nn}^{out-in} can be defined as a mapping between an *out*-degree k (horizontal axis in Fig. 2(d)) and the average *in*-degree of all neighbors connected from nodes of that degree k (vertical axis in Fig. 2(d)). We can further calculate r values according to the formulae given by Newman [17]. As shown in Fig. 2(d), the k_{nn}^{out-in} distribution of Flickr is significantly upward and the corresponding r value is 0.2329 (shown in Table 3). This upwardness of k_{nn}, along with the significantly positive value r, indicates that Flickr nodes with high out-degree



(a) k_{nn}^{in-in} of G_{F0} (b) k_{nn}^{in-out} of G_{F0} (c) $k_{nn}^{out-out}$ of G_{F0} (d) k_{nn}^{out-in} of G_{F0}



(e) k_{nn}^{in-in} of G_{E0} (f) k_{nn}^{in-out} of G_{E0} (g) $k_{nn}^{out-out}$ of G_{E0} (h) k_{nn}^{out-in} of G_{E0}

Fig. 2. The k_{nn} distribution of G_{F0} and G_{E0}

Table 3. Degree correlation analysis of activity graphs and linkage graphs. All activity graphs studied here are based on interactions and linkage graphs based on friendship links.

Graph Type	Network	Size	Symmetry	r^{in-in}	r^{in-out}	$r^{out-out}$	r^{out-in}
Activity graph	Epinions	93,598	30.5%	0.0135	0.0698	-0.0164	-0.0556
	Advogato[1]	7,421	39.8%	-0.0250	0.0003	-0.0008	-0.0683
	Wiki-Vote[2]	7,115	14.4%	0.0051	0.0071	-0.0189	-0.0832
	Wike-Talk[2]	2,394,385	5.6%	-0.0566	-0.0482	0.0231	-0.0809
Linkage graph	Flickr	1,834,425	62.3%	0.3383	0.2614	0.1830	0.2329
	Facebook [22]	46,952	62.9%	0.1830	0.2131	0.2719	0.2435
	LiveJournal [15]	5,204,176	73.5%	0.1759	0.3633	0.3763	0.1796

[1] Advogato dataset, available at http://www.trustlet.org/datasets/advogato/
[2] Snap datasets, available at http://snap.stanford.edu/data/

have a strong tendency to connect to nodes with high in-degree. The other three kinds of k_{nn} and assortativity r can be defined and calculated accordingly, as shown in Fig. 2.

Compared to the upward tendencies of the four k_{nn} distributions of Flickr (Fig. 2(a)-2(d)), Epinions have relatively flat k_{nn} distributions (Fig. 2(e)-2(h)). This difference is again reflected by the assortativity coefficients. While Flickr has significantly positive assortativity coefficients, the r values of Epinions are much neutral. This neutrality indicates that the edges of Epinions are relatively random, recalling that r values of random graphs should be 0 in theory.

Existing work already shows that social networks tend to have positive r values while the r values of other networks tend to be negative [17]. However, this rule does not hold in activity graphs of online social networks. To further study the degree correlation of the two graph types, we analyze additional datasets as shown in Table 3. We observe from the table that all activity graphs we studied tend to have neutral r values around 0, while linkage graphs have significantly positive r values. One possible reason for the neutrality of activity graphs could be the relative randomness of interactions, while linkage graphs have strong reciprocity. The reciprocity of the datasets are also shown in the forth column of Table 3, and we will discuss this further in the next section.

4 Network Generator

In this section, we present our generator for generating online social networks. The generator captures the local behavior that forms positive degree correlation, and focuses especially on generating linkage graphs. Our goal in developing this generator is to understand the local behavior of the global degree correlation property, and to generate a network that exhibits all examined properties in the previous section.

4.1 Generator Description

The fire burning idea of forest fire generator aims at meeting a set of properties, including power-law degree distribution, high clustering, small diameter, densification, and shrinking diameter [11]. However, the generator does not consider degree correlation. By exploring the parameter space of forest fire generator, we find that the generated networks can exhibit only neutral assortativity coefficients when other examined properties are met. In contrast, networks by our generator could achieve significant positive assortativity coefficients of degree correlation, and exhibit other properties at the same time.

Forest fire generator was proposed based on the intuition of how authors find references in citation networks. In this generator, new nodes can only connect to old nodes because authors cannot refer to unpublished papers. This is not the case of online social networks, because these networks allow old nodes to connect to new ones. Actually, the linkage graphs are quite reciprocated with high link symmetry [10,15]. In view of this, we incorporate the idea of *symmetry* into our generator, while retaining the fire burning idea in order to obtain the other properties.

Algorithm 1. AddNode(v, P_b, P_s)

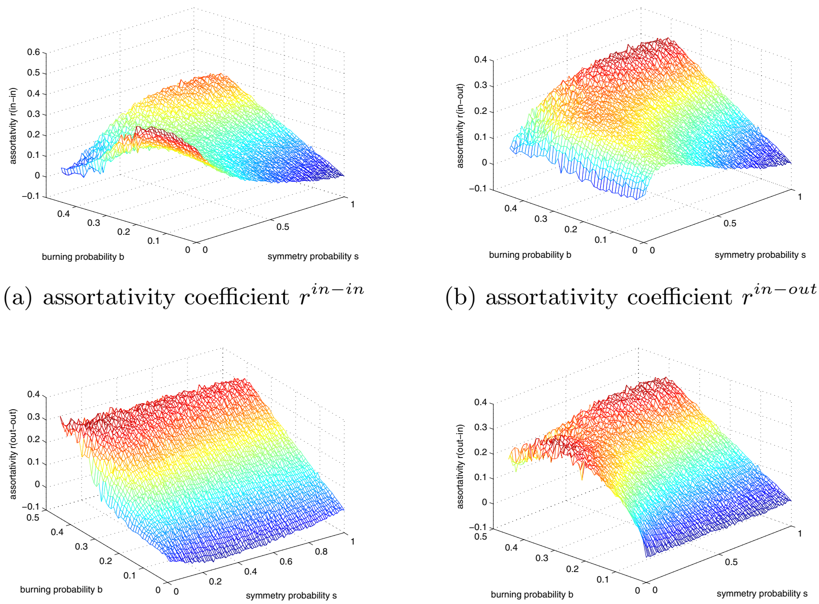
Upon a new node v arrives;
v chooses an ambassador node w **uniformly** at random;
Set $S := \emptyset$;
Queue Q.add(w);
while Q is not empty **do**
    generate a random number x, that is **geometrically** distributed with means $1/(1-P_b)$;
    node $c := Q$.head();
    Q.remove(c); S.add(c);
    Set $T :=$ the outlink ends of c that are not in S;
    for $i := 1$ to min(x, T.size()) **do**
        node $t :=$ a randomly sampled node in T without replacement;
        Q.add(t);
        S.add(t);
    end for
end while
for node u in S **do**
    v forms a link to u;
    u forms a link to v with probability P_s;
end for

Our generator has two parameters, the *burning probability* P_b which is in charge of the burning process, and the *symmetry probability* P_s which indicates backward linking from old nodes to new ones. Consider the arrival of a new node, v, it follows the process as shown in Algorithm 1.

The P_b controls a BFS-based forward burning process, as users in these kinds of networks can check the out-linked friends of their out-linked friends. The fire burns increasingly fiercely with P_b approaching 1. Meanwhile, the P_s adds fuel to the fire as it brings more links.

4.2 Experiments

We partially explore the parameter space of our generator, in order to understand the degree correlation property which is measured by assortativity coefficients. Our exploration covers burning probability P_b from 0 to 0.45 and symmetry probability P_s from 0 to 1, with step size 0.01 for both. One problem about generating online social networks is the large size of the datasets. For simplicity, we fix the number of nodes to 90,000 in our experiments.



(a) assortativity coefficient r^{in-in}

(b) assortativity coefficient r^{in-out}

(c) assortativity coefficient $r^{out-out}$

(d) assortativity coefficient r^{out-in}

Fig. 3. The assortativity coefficients over the partial parameter space of our generator

Fig. 3 shows our results of four assortativity coefficients over the parameter space of our generator. The vertical axis of every subfigure represents the assortativity coefficient, and the two horizontal axes represent the burning probability and the symmetry probability, respectively.

We observe that our generator can generate significantly positive assortativity coefficients in general. This is probably because P_s gives chances for big nodes to connect back to big nodes, while the links in forest fire generator are created much randomly. In addition, with the increase of P_s and P_b, the graphs generated by our generator show upward trend of r values. This is because symmetry has an increasing impact when more links are burned.

As discussed above, degree correlation is one of the major differences between linkage graphs and activity graphs, and the fraction of symmetric links of linkage

Table 4. Examined properties of our generator with symmetry probability $P_s = 0.7$ and burning probability $P_b = 0.472$

nodes	r^{in-in}	$r^{out-out}$	Power-law coef.	Diameter	Diameter over time
90,000	0.2231	0.2031	2.30/2.33	8.1225	shrinking
edges	r^{in-out}	r^{out-in}	Densification coef.	CC	CC over time
691,296	0.2135	0.2129	1.18	0.45	stable

graphs is relatively high. Based on our exploration of the parameter space, we also find that adding symmetry to network generator can produce significantly positive assortativity coefficients. Symmetry is a reflection of reciprocity in linkage graphs, and therefore, we argue that reciprocity is a key factor that could lead to positive degree correlation and positive assortativity coefficients.

As an example, we try to generate a graph that can meet all the observed properties of Flickr. When the number of nodes is fixed at 90,000, the edge number of Flickr is around 666,000 which can be estimated by the tendency of the densification power-law coefficient. With symmetry probability $P_s = 0.7$ and burning probability $P_b = 0.472$, the results are shown in Table 4. The generated graph has significantly positive r values which are similar to Flickr. Moreover, the graph follows static properties of Flickr including power-law degree distribution, small diameter, and high clustering coefficient, as well as dynamic properties of shrinking diameter, stable clustering coefficient, and densification.

5 Discussion

In this section, we discuss some potential implications based on our findings and experiments. Specially, we concentrate on two applications: the linkage-based application of *information dissemination*, and the activity-based application of *trust inference*.

5.1 Information Dissemination

Online social networks have become a popular way to disseminate information. This kind of applications should be built on the linkage graph, as linkage graph is a common mechanism for information dissemination in the content sharing networks, such as Flickr and Youtube. In addition, Cha et al. have found that a large portion of information spread through social links [3,4], making the underlying structure worth a thorough study. They observe that information is limited to the vicinity of the source, although the small world property implies that information could spread widely through the network. We give a possible explanation based on the observation of the positive degree correlation: high-degree nodes tend to connect to each other, and thus their impact is limited within the high-connected core of the network. They also find that information spreads relatively widely when the source node is of high degree. Consequently, in order to widely spread the information, the information sources should include some high-degree nodes in the core of the graph, as well as some nodes at the edge of the graph.

5.2 Trust Inference

In the context of online social networks, we may have to interact with unknown participants. It is necessary to evaluate the trustworthiness of these unknown participants in the open environment, and one particular approach among others is to infer trust through existing trust relationships.

In order to assist trust inference, researchers have proposed to study the structural properties of the underlying social networks [19,8,12]. We believe that it is suitable to infer trust through activity graph of real user interactions, rather than through the linkage graph of social links in the online environment. First, edges in a linkage graph may be the result of reciprocity, and these edges cannot indicate trust. Second, we need explicit trust ratings along the way to the unknown participant to carry out the computation of trust inference [8]. Activity graphs can mitigate this problem, because we can obtain trust ratings from feedback of every interaction [12].

Golbeck and Hendler have mentioned that graph generator is necessary for evaluating algorithms about trust inference on networks [8]. However, they conduct their experiments on networks generated by the small world generator, and this generator captures only the clustering and diameter properties of social networks. To make the results more convincing, we need generators that could generate more realistic graphs. Our generator can capture several dynamic properties of social networks, while retaining conciseness with only two parameters.

6 Conclusion

In this paper, we have studied several structural properties on two direct graphs mapped from online social networks. We recognize the two graphs as linkage graph and activity graph, respectively. Our results show that the two graphs are very similar to each other in several common static and dynamic properties, but quite different in degree correlation. We analyze several additional datasets and confirm that degree correlation is a key indicator of the two graph types. To further understand this property, we propose our network generator and find that reciprocity is a key factor for this difference.

Future developers should consider and take advantage of the structural properties of the corresponding underlying network, and develop their applications accordingly. Moreover, our findings and generator together could be used to detect anomalies, predict network evolving behavior, and generate realistic graphs of online social networks.

Acknowledgments. The authors would like thank Alan Mislove for sharing their Flickr data, and Paolo Massa for sharing their Epinions data. This work is supported by the National Natural Science Foundation of China (No. 60736015, 60721002, 61073030), the National 973 Program of China (2009CB320702), the National 863 Program of China (2009AA01Z117), and the ”Climbing Program” of Jiangsu Province, China (BK2008017).

References

1. Ahn, Y.Y., Han, S., Kwak, H., Moon, S., Jeong, H.: Analysis of topological characteristics of huge online social networking services. In: Proceedings of the 16th International Conference on World Wide Web, WWW 2007, pp. 835–844. ACM, New York (2007)
2. Barabási, A., Albert, R.: Emergence of scaling in random networks. Science 286(5439), 509–512 (1999)
3. Cha, M., Mislove, A., Adams, B., Gummadi, K.P.: Characterizing social cascades in flickr. In: Proceedings of the First Workshop on Online Social Networks, WOSP 2008, pp. 13–18. ACM, New York (2008)
4. Cha, M., Mislove, A., Gummadi, K.P.: A measurement-driven analysis of information propagation in the flickr social network. In: Proceedings of the 18th International Conference on World Wide Web, WWW 2009, pp. 721–730. ACM, New York (2009)
5. Chakrabarti, D., Faloutsos, C.: Graph mining: Laws, generators, and algorithms. ACM Comput. Surv. 38 (2006)
6. Clauset, A., Shalizi, C.R., Newman, M.E.J.: Power-law distributions in empirical data. SIAM Review 51, 661–703 (2009)
7. Garriss, S., Kaminsky, M., Freedman, M.J., Karp, B., Mazières, D., Yu, H.: RE: reliable email. In: Proceedings of the 3rd Conference on Networked Systems Design & Implementation, NSDI 2006, vol. 3, pp. 297–310. USENIX Association, Berkeley (2006)
8. Golbeck, J., Hendler, J.: Inferring binary trust relationships in Web-based social networks. ACM Transaction on Internet Technology 6, 497–529 (2006)
9. Kleinberg, J.M., Kumar, R., Raghavan, P., Rajagopalan, S., Tomkins, A.S.: The web as a graph: Measurements, models, and methods. In: Asano, T., Imai, H., Lee, D.T., Nakano, S.-i., Tokuyama, T. (eds.) COCOON 1999. LNCS, vol. 1627, pp. 1–17. Springer, Heidelberg (1999)
10. Kumar, R., Novak, J., Tomkins, A.: Structure and evolution of online social networks. In: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2006, pp. 611–617. ACM, New York (2006)
11. Leskovec, J., Kleinberg, J., Faloutsos, C.: Graphs over time: densification laws, shrinking diameters and possible explanations. In: Proceedings of the 11th ACM SIGKDD International Conference on Knowledge Discovery in Data Mining, KDD 2005, pp. 177–187. ACM, New York (2005)
12. Liu, G., Wang, Y., Orgun, M.A.: Optimal social trust path selection in complex social networks. In: Proceedings of the 24th AAAI Conference on Artificial Intelligence, AAAI 2010, pp. 1391–1398 (2010)
13. Massa, P., Avesani, P.: Trust-aware collaborative filtering for recommender systems. In: Chung, S. (ed.) OTM 2004. LNCS, vol. 3290, pp. 492–508. Springer, Heidelberg (2004)
14. Mislove, A., Koppula, H.S., Gummadi, K.P., Druschel, P., Bhattacharjee, B.: Growth of the Flickr social network. In: Proceedings of the First Workshop on Online Social Networks, WOSP 2008, pp. 25–30. ACM, New York (2008)

15. Mislove, A., Marcon, M., Gummadi, K.P., Druschel, P., Bhattacharjee, B.: Measurement and analysis of online social networks. In: Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement, IMC 2007, pp. 29–42. ACM, New York (2007)
16. Mislove, A., Post, A., Druschel, P., Gummadi, K.P.: Ostra: leveraging trust to thwart unwanted communication. In: Proceedings of the 5th USENIX Symposium on Networked Systems Design and Implementation, NSDI 2008, pp. 15–30. USENIX Association, Berkeley (2008)
17. Newman, M.: Mixing patterns in networks. Physical Review E 67(2), 026126 (2003)
18. Newman, M.: The structure and function of complex networks. SIAM Review 45, 167–256 (2003)
19. Pujol, J.M., Sangüesa, R., Delgado, J.: Extracting reputation in multi agent systems by means of social network topology. In: Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS 2002, pp. 467–474. ACM, New York (2002)
20. Tauro, S., Palmer, C., Siganos, G., Faloutsos, M.: A simple conceptual model for the Internet topology. In: Global Telecommunications Conference, GLOBECOM 2001, vol. 3, pp. 1667–1671. IEEE, Los Alamitos (2001)
21. Tran, T., Rowe, J., Wu, S.F.: Social email: A framework and application for more socially-aware communications. In: Bolc, L., Makowski, M., Wierzbicki, A. (eds.) SocInfo 2010. LNCS, vol. 6430, pp. 203–215. Springer, Heidelberg (2010)
22. Viswanath, B., Mislove, A., Cha, M., Gummadi, K.P.: On the evolution of user interaction in Facebook. In: Proceedings of the 2nd ACM Workshop on Online Social Networks, WOSN 2009, pp. 37–42. ACM, New York (2009)
23. Watts, D., Strogatz, S.: Collective dynamics of 'small-world' networks. Nature 393(6684), 440–442 (1998)
24. Wilson, C., Boe, B., Sala, A., Puttaswamy, K.P., Zhao, B.Y.: User interactions in social networks and their implications. In: Proceedings of the 4th ACM European Conference on Computer Systems, EuroSys 2009, pp. 205–218. ACM, New York (2009)
25. Yu, H., Kaminsky, M., Gibbons, P.B., Flaxman, A.: SybilGuard: defending against Sybil attacks via social networks. In: Proceedings of the 2006 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, SIGCOMM 2006, pp. 267–278. ACM, New York (2006)