Understanding Social Networks Properties for Trustworthy Computing

Abedelaziz Mohaisen, Huy Tran, Nicholas Hopper, Yongdae Kim
Department of Computer Science and Engineering
University of Minnesota, Minneapolis, MN 55455, USA

***Abstract*—The ever-increasing popularity of social networks opens new directions for leveraging social networks to build primitives for security and communication, in many contexts. Such primitives utilize the trust in these social networks to ensure collaboration and algorithmic properties exhibited in such networks to argue for the effectiveness of such primitives. Despite the importance of such properties and their quality to the operation of these primitives, less effort is made to measure these properties and understand the relationship among them and to other characteristics of social networks.**

We extend our earlier results measuring the mixing time, to investigate a new property used for building Sybil defenses, namely the expansion of social graphs. We measure the expansion of social graphs, and show quantitatively that, with a few exceptions, it is sufficient to support Sybil defense mechanisms based on expansion. We relate the mixing time of social graphs to graph degeneracy, which captures cohesiveness of the graph. We experimentally show that fast-mixing graphs tend to have a larger single core whereas slow mixing graphs tend to have smaller multiple cores. While this study provides quantitive evidence relating the mixing time to coreness of the graph, it also agrees with our previous observations about the tight-knit community structure in slow mixing social graphs.

***Index Terms*—Social networks, mixing time, expanders, measurements, Sybil defenses.**

I. Introduction

Leveraging social ties for building trustworthy computing services is becoming quite popular, and promises many primitives and applications for communication and security. Such applications and primitives benefit from both the trust exhibited in the underlying social networks—which rationalizes collaboration among nodes in the services built on top of social networks—and other algorithmic properties, which support the argument for the effectiveness of applications built on top of the social network. While most applications and primitives built on top of social networks share a commonality of purpose in using trust in the underlying graph (e.g., ensuring collaboration, rationalizing assumptions about the the nature of the underlying social graph and attackers' capabilities, among many others), the algorithmic properties used in them differ greatly [16]. In the following, we elaborate on some of these properties and the rationale of using them in building such applications and primitives.

One example of such applications is social network based Sybil defenses, which is a well-established vein of research, where several designs have been proposed in the literature [4], [10], [19], [22], [26], [27]. In these designs, trust is used to rationalize the difficulty of penetrating the social graph and establishing arbitrarily many links that thwart the utility of the defense mechanism, whereas the fast-mixing property is used to argue for the effectiveness of the detection of Sybil identities, and to yield a feasible solution. The same algorithmic property, of fast mixing social graphs, is used in a closely related direction for demonstrating the utility of social graphs as good mixers for potential deployment of anonymous communication networks on top of them [18]. An even more strict property than fast mixing, namely "good" expansion, is used for building a Sybil defense mechanism in [23].

Other examples of applications and properties used in building them include (node) betweenness for Sybil defense [19], betweenness and similarity for improving routing in delay tolerant networks [2], and closeness for efficient content sharing and anonymity [6], among others.

Despite their importance to the usability of such applications and primitives, less effort is spent understanding the quality of such properties in real-world social graphs, and even less effort is made to understand such properties and to relate them to other topological characteristics of such graphs. Most recently, we considered two such properties, and examined their quality in real-world social graphs: the mixing time of social graphs [17] and quality (and distribution) of shortest-path betweenness [15]. Both studies were motivated by the use of these properties for Sybil defenses and the assumptions made on the value of these properties in that context.

In this work, we extend upon our previous results to better understand these properties in social graphs. We consider both the mixing time and expansion of social graphs. We measure the expansion, as required by Sybil defenses such as [23], and show quantitatively that most social graphs have good expansion. We also relate the mixing time of social graphs to graph degeneracy, which captures the cohesiveness of the graph. We experimentally show that fast mixing graphs tend to have larger cores (formally defined below) whereas slow mixing graphs tend to have smaller cores.

This paper is organized as follows. In section II, we review related work to motivate for our measurements. In section III we outline the preliminaries and theoretical tools required for understanding the measurements and results in the paper. In section IV we outline measurements on both expansion and mixing characteristics of social graphs, and relate the latter one to coreness of nodes in the social graph. In section V we outline discussions headlines. In section VI, we draw concluding remarks and outline open directions of research.

II. Related work

Using social networks to build primitives and services by exploiting social structures, social network properties, and trust in social graphs has been a growing area of research, where several directions are investigated to solve challenging problems using social networks. Here, we limit ourselves to the recent work on using Social networks for Sybil defenses, and follow-up work on examining assumptions in such defenses.

While the idea of using social networks to improve security in distributed systems has been investigated earlier in [12] and [3], the first formal attempt to use social networks for defending attacks is seen in SybilGuard [27], which pioneered the use of fast-mixing graphs for Sybil detection. The same work has been extended and improved in SybilLimit [26]. Both SybilGuard and SybilLimit did not measure the mixing time, which is necessary for their operation. However, SybilLimit showed end-to-end results demonstrating that real-world social graphs mix well enough to support the requirements of SybilLimit. Also using fast-mixing graphs, Danezis and Mittal introduced SybilInfer, an inference mechanism for Sybil nodes detection [4]. Tran et al. used the fast-mixing social graphs, and a ticket distribution mechanism that benefits from the mixing characteristics of social graphs, for building a Sybil resistant voting system called SumUp [22]. Lesniewski-Laas and Kaashoek used the same property to build a Sybil-proof DHT system. Though the authors of most of these designs performed experiments to show that their designs operate on real-world social graphs, none of these papers measured the mixing time directly from social graphs.

Using a slightly different assumption than the mixing time, namely the graph expansion assumption, Tran et al. introduced GateKeeper [23] that improves the guarantees of SybilLimit by reducing the number of Sybil identities introduced per attack edge to $O(1)$. While they provided experiments to demonstrate the operation of their design on real-world social graphs, the authors of GateKeepers did not attempt to measure the expansion of the graphs they used in their experiments, though that is doable as shown in earlier.

The use of assumptions in building Sybil defenses on top of social networks—like the "fast-mixing"—without verification motivated for investigating whether such systems work on various social graphs or not. In [24], Viswanath et al. conducted an experimental analysis of Sybil defenses based on social networks by comparing different defenses (SybilGuard [27], SybilLimit [26], SybilInfer [4], and SumUp [22]). They show that the different Sybil defenses work by ranking different nodes based on how well-connected are these nodes to a trusted node. Also, they demonstrate that the different Sybil defenses are sensitive to community structure in social networks whereas community detection algorithms can be used to replace the random walk based Sybil defenses.

In [17], Mohaisen et al. measured the mixing time in several social graphs (mostly in Table I) and demonstrated that social graphs are slower mixing than believed in literature. Furthermore, they noticed that mixing patterns in social graphs are associated with the underlying social model, where social networks with confined social models, and strict trust properties, are slow mixing whereas social networks with less strict trust properties are fast mixing. This observation is used in [16] to account for trust in social network-based Sybil defenses using modulated random walks.

Finally, Dell'amico et al. [5] measured the mixing time in four datasets (Epinion, OpenPGP, DBLP, and Advogato) using the sampling method and the model in (2). Their work is motivated by evaluating the effectiveness of wide range of work based on both the mixing time and transitive trust, for which the authors *claim* that real-world social networks meet theoretical assumptions of these defenses without showing that on any particular Sybil defense. The main conclusion made in [5] is that the mixing time is not associated with any of the known characteristics of the social graphs.

Our work is motivated by the latter part of the related work. In this paper we are interested in understanding the assumptions used for building systems on top of social networks. For that, we investigate two directions. The first direction is to understand what makes a given graph fast mixing, while another graph is not. We relate this to the core structure of the social graph. The second direction is measurement of the expansion of social graphs, and relating it to the mixing time.

III. Preliminaries and Theoretical Background

A. Graph (network) model

Let $G = (V, E)$ be a simple undirected and unweighted graph, where V is the set of vertices of G (correspond to social actors in the social graph) such that $|V| = n$ and E is the set of edges in G (correspond to the interrelationships between the social actors in the social graph) such that $|E| = m$. For every node $v_i \in V$, let the number of nodes in V adjacent to v_i be $\deg(v_i)$. For such G, we define the stochastic transition probability matrix $P = [p_{ij}]$ of size $n \times n$ where the $(i,j)^{\text{th}}$ entry in P is the probability of transitioning from node v_i to node v_j in one step, which is defined as follows

$$p_{ij} = \begin{cases} \frac{1}{\deg(v_i)} & \text{if } v_i \text{ is adjacent to } v_j, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

B. Graph cores and node coreness

For the same graph model defined above, and for any k where $1 \leq k \leq k_{\max}$, let $G_k = (V_k, E_k)$ be a subset of G such that $|V_k| = n_k$ and $|E_k| = m_k$, and with the constraint that for all $v_i \in V$, $\min\{\deg(v_i)\} = k$; i.e., minimum degree of any node $v_j \in V_k = k$. G_k is said to be a $k-$core of G if, in addition to the above condition, it is a maximal and connected graph. If we relax the connectivity condition of the $k-$core, we get a set of cores (≥ 1) where each of such cores satisfy the degree condition. Such set of cores is represented as a (possibly disconnected) graph $G'_k = (V'_k, E'_k)$ where $|V'_k| = n'_k$ and $|E'_k| = m'_k$. An efficient algorithm for decomposing a simple graph to its $k-$cores by iteratively pruning nodes with degree less than k has the complexity of $O(m)$ where m is the number of edges in the graph [1]. We define the node-relative

size of G_k as $\nu_k = \frac{n_k}{n}$ and the edge-relative size of G_k as $\tau_k = \frac{m_k}{m}$. Similarly, for G'_k we define $\nu'_k = \frac{n'_k}{n}$ and $\tau'_k = \frac{m_k}{m}$. Finally, for every node $v_i \in V$, we define the coreness c as the largest c s.t. $v_i \in G_c$ where G_c is a $c-$core.

C. Mixing time

Moving from a node to another on G is captured by the Markov chain which represents a random walk over G. A random walk of length w over G is a sequence of vertices in G such that each node is selected at its predecessor in the random walk following the transition probability defined in (1). The Markov chain is said to be *ergodic* if it is irreducible and aperiodic, where in such case it has a unique stationary distribution π and the distribution after w steps converges to π. The stationary distribution of the Markov chain is a probability distribution that is invariant to the transition matrix P (i.e., $\pi P = \pi$). The mixing time, T, is defined as the minimal length of the random walk to reach the stationary distribution. More precisely, the mixing time of a Markov chain on G parameterized by a variation distance ϵ is defined as

$$T(\epsilon) = \max_i \min\{t : |\pi - \pi^{(i)}P^t|_1 < \epsilon\}, \tag{2}$$

where $\pi = [\deg(v_i)/2m]^{1\times n}$, for all $v_i \in V$, (for simple graph) is the stationary distribution, $\pi^{(i)}$ is the initial distribution concentrated at vertex v_i, P^t is the transition matrix after t steps, and $|\cdot|_1$ is the total variation distance, defined as $\frac{1}{2}|\sum_{j=1}^{n}|\pi_j^{(i)} - \pi_j|$. We say that a Markov chain is "fast mixing" for the graph [4], [10], [26], [27] when $\epsilon = \Theta(\frac{1}{n})$, and $T(\epsilon) = O(\log n)$. An interesting result on bounding the mixing time for simple graphs is provided in [21], and used in [17] for measuring the mixing time of several real world social graphs. In short, the method uses the second largest eigenvalue, μ, where the mixing time of the social graph is bounded by $\frac{\mu}{2(1-\mu)}\log(\frac{1}{2\epsilon}) \leq T(\epsilon) \leq \frac{\log(n)+\log(\frac{1}{\epsilon})}{1-\mu}$.

Measuring the mixing time can be done using both definitions [17]. However, since using the second largest eigenvalue modulus accounts only for the poorest mixing source in the social graph (correspond to the $\max_i$ in Eq. (2)), measuring the mixing time using the definition in Eq. (2) is highly desirable to observe the richer patterns of mixing presenting various sources in the social graph. In [17], we used this observation to variety of mixing patterns in the same social graph. By varying sources of walks, and sampling such sources from G, one can get a good estimate about the distribution of the mixing across different sources, and thus the entire graph.

D. Graph expansion

Let $S \subset V$, where $0 < |S| \leq 1/2|V|$, be any set of vertices in the graph. The (vertex) expansion factor α is defined as

$$\alpha = \min_{0<|S|\leq\frac{n}{2}} \frac{|N(S)|}{|S|} \tag{3}$$

Where the minimum is over all nonempty sets S of at most $n/2$ vertices—S need not be connected and thus the number of possible configurations of S is exponential in n—and $N(S)$ is the set of vertices not in S but each of which is connected by an edge to another node in S. In [23], where this property is used for building a Sybil defense, the definition of S is further restricted so as S is connected. Such restriction reduces the number configurations of S to a linear factor of n. Here, we elaborate on how to measure α in such setting.

We estimate the expansion factor of a graph by constructing an envelope Env_d (the same terminology used in [23]) formed by all nodes that are within a (shortest-path) distance i from a core node. The expansion Exp_i of the envelope consists of all of its immediate neighbors. We define the expansion factor as $\alpha_i = |Exp_i|/|Env_i|$. In our experiments, by letting each of the nodes in the graph to be the core, we calculate the expansion factor α_i, with $0 \leq i \leq d-1$, where d is the diameter of the graph, by building a tree rooted at the core that expands in the breadth-firth search manner. Let L_i be the number of nodes at level i in the tree, we have

$$\alpha_i = \frac{L_{i+1}}{\sum_{j=0}^{i} L_j} \tag{4}$$

To estimate the expansion characteristics of the social graphs, we run our experiment by letting each node in the graph to be the core and build a breadth-first search tree rooted at that core. We then count the number of nodes in each level of the tree to calculate the expansion factor as in Eq. (4).

IV. Results

A. Social graphs

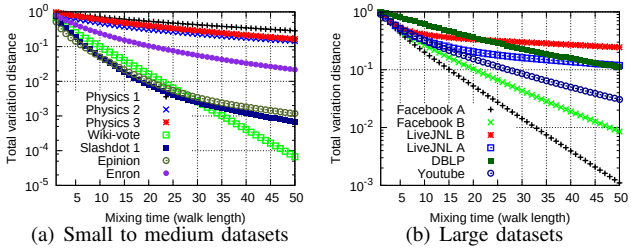
The social graphs which we use in this work are in Table I. These graphs are considered benchmarks and are used in many recent studies including [10], [16], [17], [23], [24], [26] for studying their mixing characteristics, or bringing insight on the usability of social network-based Sybil defenses. These graphs are shown to possess various mixing characteristic, and different underlying social models. First, these graphs represent variety of social network sizes, making them rich of social structures. Second, qualities of links in these graphs differ greatly and express different social structures which make them good choice to study the tradeoffs between the algorithmic properties and trust models used in social networks based Sybil defenses (see [16]). For more details on these graphs characteristics see our work in [17].

Measurements—To motivate for the first part of this work, and based on our prior work in [17], we use the sampling technique as explained in section III, to measure the mixing time of some of the social graphs in Table I. The results are shown in Figure 1.

In Figure 1(a), we observe that—despite their apparent differences in size—both of Wiki-vote and Enron have relatively similar mixing characteristics. On the other hand, while Wiki-vote and "Physics 2" have relatively similar size, they have entirely different mixing characters. Similar observations can be made on other social graphs in Figure 1, and by observing their size in Table I. In the following subsection, we try to understand the reason of this behavior, by relating the mixing characteristics of social graphs to its core structure.

TABLE I
DATASETS, THEIR PROPERTIES AND THEIR SECOND LARGEST EIGENVALUES OF THE TRANSITION MATRIX

Dataset	Nodes	Edges	μ	Dataset	Nodes	Edges	μ
Wiki-vote [7]	7,066	100,736	0.899418	Epinion [20]	75,879	508,837	0.998133
Slashdot 2 [9]	77,360	546,487	0.987531	DBLP [11]	614,981	1,155,148	0.997494
Slashdot 1 [9]	82,168	582,533	0.987531	Facebook A [25]	1,000,000	20,353,734	0.982477
Enron [8]	33,696	180,811	0.996473	Facebook B [25]	1,000,000	15,807,563	0.992020
Physics 1 [8]	4,158	13,428	0.998133	Livejournal A [13]	1,000,000	26,151,771	0.999387
Physics 2 [8]	11,204	117,649	0.998221	Youtube [13]	1,134,890	2,987,624	0.997972
Physics 3 [8]	8,638	24,827	0.996879	Rice-cs-grad [14]	501	3255	na

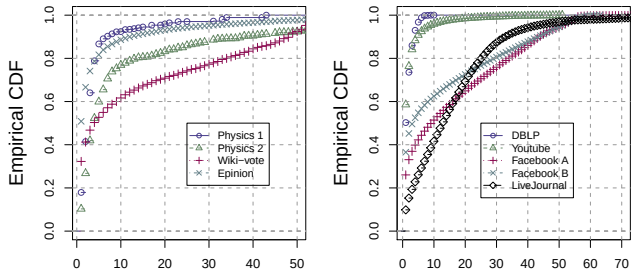


(a) Small to medium datasets (b) Large datasets

Fig. 1. Measurement of the mixing time of different social networks (shown in Table I) using the sampling method for 1000 random sources.

B. Measuring cores of social graphs

Recall the definition of the core (G'_k) in section III. We use the definition, and the algorithm provided in [1], to iteratively prune nodes with degree less than k to obtain the $k-$core of the graph, for all possible k. In particular, in this experiment we are interested in size of each core (i.e., the number of nodes with coreness k for every k, or ν'_k and ν_k). By performing the experiment on representative graphs from Table I, with various sizes, we plot the empirical CDF for the coreness number of each node in each of the graphs as shown in Figure 2.



(a) Small datasets (b) Large datasets

Fig. 2. Coreness distribution (ECDF) for the different social graphs in Table I.

Generally speaking, and despite some ambiguous cases (see explanation in section V), Figure 2 shows that—by comparing the distribution of the core number to that of the mixing characteristics in Figure 1—fast-mixing social graphs tend to have relatively larger portion of nodes with higher coreness (relative size). This is, the fast mixing graphs have larger core—or set of cores, when G'_k is used—whereas slower mixing graphs tend to have smaller cores. For example, by comparing the Wiki-vote dataset to that of the Enron dataset in 2(a) and 1(a), we find them consistent in both graphs (reasoning about the inconsistency of Enron is in section V). The results in 2(a) and 1(a) are also consistent, in that faster mixing graphs have larger core(s). In 2(a) and 1(a), notice that all graphs have relatively similar sizes.

TABLE II
NUMERICAL RESULTS OF OPERATING GATEKEEPER [23] ON TOP OF DIFFERENT SOCIAL GRAPHS WITH DIFFERENT CHARACTERISTICS. 10 ATTACKERS ARE SELECTED RANDOMLY AND 99 DISTRIBUTERS ARE SAMPLED IN EACH CASE (ATTACK EDGES ARE 131, 145, 277, AND 344, RESPECTIVELY). f IS A SECURITY PARAMETER, HONEST ACCEPTANCE PERCENT IS OF THE WHOLE GRAPH SIZE AND SYBIL IS PER ATTACH EDGE.

Dataset	Accept.	$f = 0.1$	$f = 0.3$	$f = 0.5$
Physics 1	Honest	89.90%	70.50%	54.40%
	Sybil	8.4	1.7	0.7
Facebook	Honest	98.40%	79.00%	51.30%
	Sybil	10.1	1.8	0.7
LiveJournal	Honest	97.00%	78.60%	53.20%
	Sybil	3.7	0.7	0.3
Slashdot	Honest	97.00%	81.10%	55.20%
	Sybil	3.1	0.8	0.4

C. Measuring the expansion of social graphs

To motivate for this measurement, consider the measurements in Table II. In this experiment, we run Gatekeeper [23] on four different datasets with different characteristics. Unsurprisingly, we observe that results of operating Gatekeeper on such graphs are quite anticipated given that that such graphs are experimented on other social network-based Sybil defenses (in [17] and [24]), despite that other defenses require social graphs to be fast-mixing whereas GateKeeper requires the graphs to be expanders with good expansion factor. Hence, we establish that the property in action is closely related to the mixing time.

To further understand the expansion characteristics of these social graphs, we recall the definition in section III. We recall that the expansion used in GateKeeper is a reduced version of the general expansion measurements, and requires only a linear number (in n) of expansion measurements. Particularly, each expansion measurement, following the model in (4), requires the construction of a breadth first search tree, which runs in $O(m)$. For all sources in the graph, where each is considered a source of expansion, the running time of a naive implementation of our algorithm takes $O(nm)$, which is manageable for small to medium sized social graphs.

We develop this algorithm to compute the expansion for all datasets in Table I. Each source node has $d-1$ measurements of sets of nodes and their neighbors, where d is the diameter of the graph. To visualize the tendency of the expansion as the set size increases, we aggregate the unique sizes of sets, and find the number of neighbors to each of them. For each set of neighbors to a unique size of nodes, we compute the maximum, minimum, and expected number of neighbors. Such statistics of measurements for a selected set of datasets are

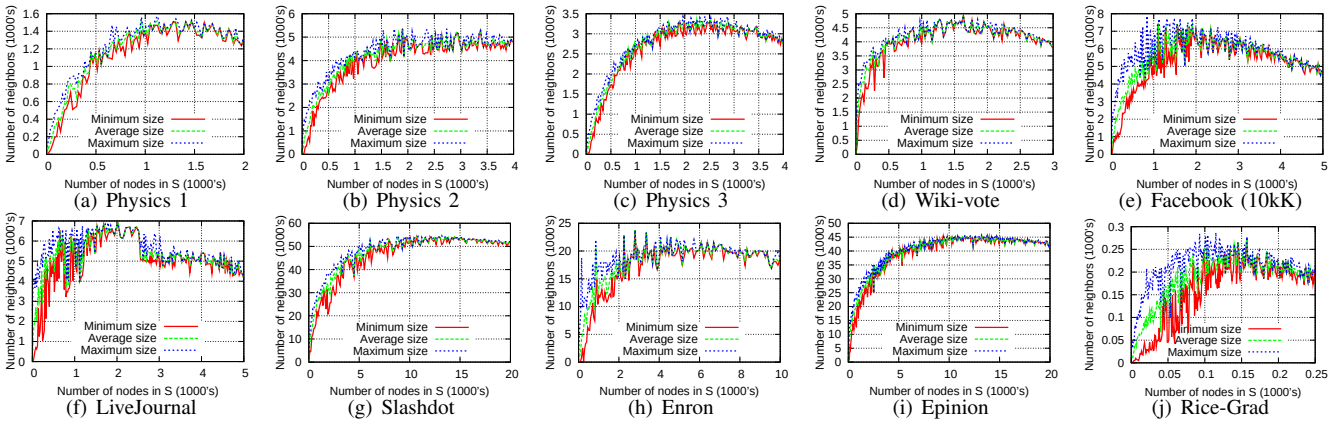


Fig. 3. A measured of the expansion of sets of nodes of different sizes beginning from every nodes in the given graphs as potential core for the expansion.

shown in Figure 3.

To further investigate the expansion of social graphs when compared to each other, we use the model in (4). For all sets of nodes with the same size, we compute the expected expansion as the average of all sizes of different neighbor sets. The result of the average expansion is shown is shown in Figure 4.

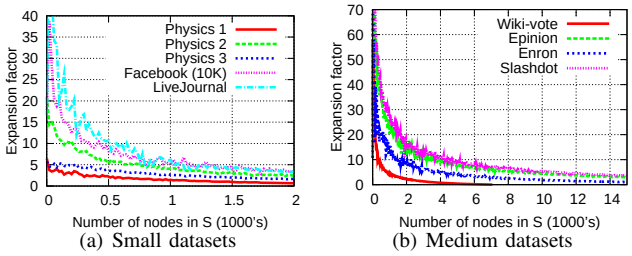


Fig. 4. Expected expansion for various set sizes of various social graphs.

V. Discussion

Our hypothesis in reasoning about the mixing time of social graphs is natural and stems out of the following observation. Given that the mixing characteristic of a particular node is not merely dependent upon the node itself and its position in the graph—but rather nodes in the social graph that connect such node to other nodes and thus make the overall graph fast-mixing, fast-mixing social graphs *should* have a dense and well-connected "core" of nodes. Such core is the backbone of flows—whether they are random or guided—between nodes. To this end, in section IV-B we investigate the existence or absence of such cores in social graphs, and try to relate findings to their mixing characteristics. We particularly found that fast mixing social graphs tend to have a large portion of nodes in their core; i.e., large ν'_k for a relatively large k.

We perform the same experiment of core computing, and use the notion of "node relative size" in section III. The results of the relative size of cores for different k values—which can also be derived from Figure 2—are shown in Figure 5 (5(a) through 5(e)). As we hypothesized, the relative size of the core is generally large for fast mixing graphs, except for a few cases (see below for explanation).

To test whether these are multiple cores or a single core with the given size, we compute the number of resulting cores from applying the algorithm for computing the cores and record the resulting number of cores at each time step (as k increases). The results are in Figure 5 (5(f) through 5(j)).

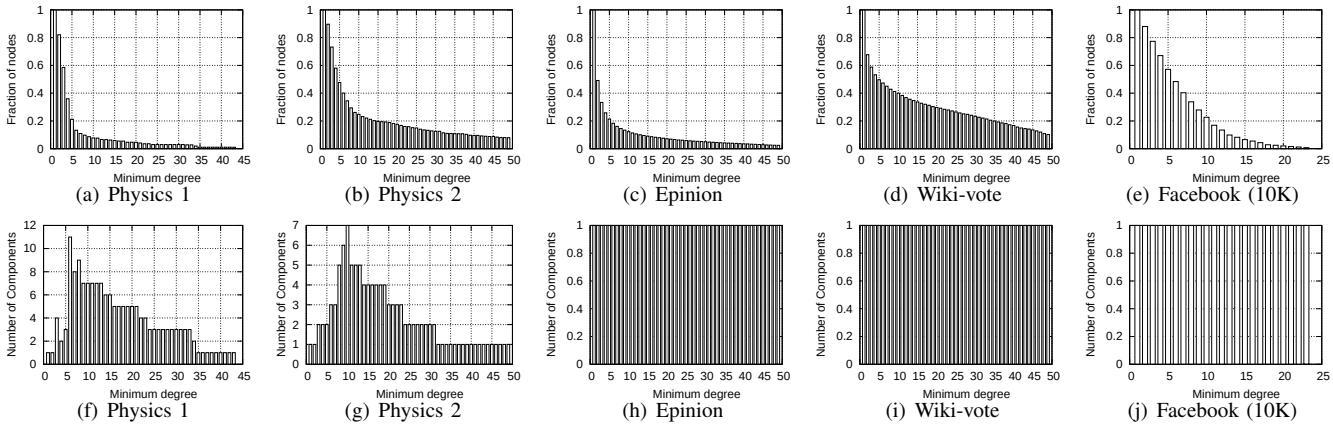
We notice that whereas slower mixing social graphs tend to have small ν'_k, and multiple cores as k grows, fast mixing social graphs have single core as it is the case of Epinion in Figure 5(h) and Wiki-vote in Figure 5(i) (or fewer cores, as it is the case of Physics 2 in Figure 5(g) versus Physics 1 in Figure 5(f)), and generally have larger ν_k (or ν'_k, respectively).

Even more interesting, we notice that this observation could be used to resolve the ambiguity of the core distribution in Figure 2 concerning Epinion and Physics 2. In Figure 2(a), it is shown that Physics 2 has a relatively larger core than that of Epinion, even though Epinion is faster mixing than Physics 2 (as shown in Figure 1). Figure 5 helps us understand the reason: while Physics core consists of three isolated components, Epinion has only single core with that size. Also, it is worth noting the difference in size (from $\sim 10,000$ in Physics 2 to $\sim 75,000$ in Epinion) which makes the raw size of the single core in Epinion even much larger than that of Physics 2, and establish a consistency in our observations concerning cores and the mixing time.

Finally, while the expansion factor used for building Gatekeeper is claimed to be different from that of the mixing time [23], and to be further strict, we notice that both are analogous to each other. In particular, we observe that the measurements in Figure 4 of the expansion factor can be interpreted as a scale of the measurements in Figure 1. Also, when operating on such networks, we observe similar tendency in the behavior between GateKeeper and other Sybil defenses (results not included). In total, we also establish that the expansion argument used in [23] is *generally* valid in the context of real-world social networks, despite that real-world social networks are not random graphs, as claimed in [23].

VI. Conclusion and Open Problems

In this work, we extend upon our previous results of measuring the mixing time, and consider a new property that is used for building Sybil defenses, namely the expansion of social graphs. For the expansion, we have shown that qualities of expansion in social graphs generality support the requirements and assumptions of such systems, as GateKeeper. We further



(a) Physics 1 (b) Physics 2 (c) Epinion (d) Wiki-vote (e) Facebook (10K)

(f) Physics 1 (g) Physics 2 (h) Epinion (i) Wiki-vote (j) Facebook (10K)

Fig. 5. A measurement of the relative size of cores, and their numbers, for different social networks.

establish that the whereas the expansion is used to advocate the operation of GateKeeper, the mixing characteristics of social graphs capture the assumption of its operation as well.

For the mixing time of the social graphs, we relate this property to graph degeneracy, which captures cohesiveness of the graph. We experimentally established that fast mixing graphs tend to have larger single core whereas slow mixing graphs tend to have smaller multiple cores.

While the main concern in this paper is limited to relating the mixing time to other characteristics of social graphs, and examining the validity of expansion assumptions in the context of social networks Sybil defenses, the same problems concerning the usability of social networks to build trustworthy computing applications are still wide open. In particular, it is open problem to understand the trust value of these social networks, and their potential. In particular, formal models of attackers supported by experimental and analytical evidence would be a fruitful direction worth investigation

Another open problem is to investigate the expansion and mixing characteristics of dynamic social graphs. Given that social graphs are dynamic, understanding the long-term impact of evolution, and how this impacts the underlying social structure, and properties used for building trustworthy applications, is an interesting direction worthy of investigation.

Acknowledgement—

We would like to thank Nguyen Tran for providing the code of Gatekeeper, and Alan Mislove and Ben Zhao for providing their social graphs used in this work. We would like also to thank Pan Hui for inviting us to submit this work to SIMPLEX. This research was supported by NSF grant CNS-0917154 and a generous research grant from KAIST.

REFERENCES

[1] V. Batagelj and M. Zaversnik, "An O (m) algorithm for cores decomposition of networks," *Arxiv preprint cs/0310049*, 2003.

[2] E. M. Daly and M. Haahr, "Social network analysis for routing in disconnected delay-tolerant manets," in *MobiHoc*. New York, NY, USA: ACM, 2007, pp. 32–40.

[3] G. Danezis, C. Lesniewski-Laas, M. Kaashoek, and R. Anderson, "Sybil-resistant DHT routing," *ESORICS*, pp. 305–318, 2005.

[4] G. Danezis and P. Mittal, "SybilInfer: Detecting sybil nodes using social networks," in *Network & Distributed System Security Conference*, 2009.

[5] M. Dellamico, , and Y. Roudier, "A measurement of mixing time in social networks," in *Int'l Workshop on Sec. & Trust Management*, 2009.

[6] T. Isdal, M. Piatek, A. Krishnamurthy, and T. Anderson, "Privacy-preserving P2P data sharing with OneSwarm," in *SIGCOMM*. ACM, 2010, pp. 111–122.

[7] J. Leskovec, D. P. Huttenlocher, and J. M. Kleinberg, "Predicting positive and negative links in online social networks," in *WWW*. ACM, 2010, pp. 641–650.

[8] J. Leskovec, J. Kleinberg, and C. Faloutsos, "Graphs over time: densification laws, shrinking diameters and possible explanations," in *KDD*. New York, NY, USA: ACM, 2005, pp. 177–187.

[9] J. Leskovec, K. J. Lang, A. Dasgupta, and M. W. Mahoney, "Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters," *CoRR*, vol. abs/0810.1355, 2008.

[10] C. Lesniewski-Lass and M. F. Kaashoek, "Whānau: A sybil-proof distributed hash table," in *NSDI*, 2010, pp. 3–17.

[11] M. Ley, "The DBLP computer science bibliography: Evolution, research issues, perspectives," in *SPIR*. Springer, 2009, pp. 481–486.

[12] S. Marti, P. Ganesan, and H. Garcia-Molina, "DHT routing using social links," *Peer-to-Peer Systems*, pp. 100–111, 2005.

[13] A. Mislove, M. Marcon, P. K. Gummadi, P. Druschel, and B. Bhattacharjee, "Measurement and analysis of online social networks," in *Internet Measurement Conference*, 2007, pp. 29–42.

[14] A. Mislove, B. Viswanath, K. P. Gummadi, and P. Druschel, "You are who you know: Inferring user profiles in Online Social Networks," in *WSDM*, New York, NY, February 2010.

[15] A. Mohaisen, T. AbuHmed, H. Kang, Y. Kim, and D. Nyang, "Mistaking friends for foes: An analysis of a social network-based Sybil defense in mobile networks," in *Proceeding of the 5th ICUIMC*. ACM, 2011.

[16] A. Mohaisen, N. Hopper, and Y. Kim, "Keep your friends close: Incorporating trust into social network-based sybil defenses," in *INFOCOM'11*. Piscataway, NJ, USA: IEEE Press, 2011, pp. 336–340.

[17] A. Mohaisen, A. Yun, and Y. Kim, "Measuring the mixing time of social graphs," in *IMC*. ACM, 2010, pp. 383–389.

[18] S. Nagaraja, "Anonymity in the wild: Mixes on unstructured networks," in *Privacy Enhancing Technologies*, ser. LNCS, 2007, pp. 254–271.

[19] D. Quercia and S. Hailes, "Sybil attacks against mobile users: friends and foes to the rescue," in *INFOCOM*. IEEE, 2010, pp. 336–340.

[20] M. Richardson, R. Agrawal, and P. Domingos, "Trust management for the semantic web," in *ISWC*, ser. LNCS. Springer, 2003, pp. 351–368.

[21] A. Sinclair, "Improved bounds for mixing rates of mc and multicommodity flow," *Comb., Prob. & Comp.*, vol. 1, pp. 351–370, 1992.

[22] N. Tran, B. Min, J. Li, and L. Subramanian, "Sybil-resilient online content voting," in *USENIX NSDI*, 2009.

[23] N. Tran, J. Li, L. Subramanian, and S. S. Chow, "Optimal sybil-resilient node admission control," in *INFOCOM*, Shanghai, P.R. China, 4 2011.

[24] B. Viswanath, A. Post, K. P. Gummadi, and A. Mislove, "An analysis of social network-based sybil defenses," in *SIGCOMM*, 2010.

[25] C. Wilson, B. Boe, A. Sala, K. P. Puttaswamy, and B. Y. Zhao, "User interactions in social networks and their implications," in *EuroSys*. New York, NY, USA: ACM, 2009, pp. 205–218.

[26] H. Yu, P. B. Gibbons, M. Kaminsky, and F. Xiao, "SybilLimit: A near-optimal social network defense against sybil attacks," in *IEEE Symposium on Security and Privacy*, 2008, pp. 3–17.

[27] H. Yu, M. Kaminsky, P. B. Gibbons, and A. Flaxman, "SybilGuard: defending against sybil attacks via social networks," in *SIGCOMM*, 2006, pp. 267–278.