Presented at the 17th International Workshop on Security Protocols, Cambridge, UK, April 1-3, 2009 (to appear in LNCS, Springer Verlag)

Brief Encounters with a Random Key Graph

Virgil D. Gligor Adrian Perrig Jun Zhao

ECE Department and CyLab
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213
{gligor,perrig,junzhao}@cmu.edu

Abstract. Random key graphs, also called uniform random intersection graphs, have been used for modeling a variety of different applications since their introduction in wireless sensor networks. In this paper, we review some of their recent applications and suggest several new ones, under the full visibility assumption; i.e., under the assumption that all nodes are within communication range of each other. We also suggest further research in determining the connectivity properties of random key graphs when limited visibility is more realistic; e.g., graph nodes can communicate only with a subset of other nodes due to range restrictions or link failures.

The notion of the random key graph (a.k.a. "uniform random intersection graph" [1]) was introduced for probabilistic key pre-distribution in wireless sensor networks [2]. Recently, these graphs have been used for a variety of different applications, such as clustering analysis [3], recommender systems using collaborative filtering [4], and cryptanalysis of hash functions [5]. While random key graphs are not Erdös-Renyi random graphs (e.g., the probabilities that graph edges exist are not necessarily independent), they have similar connectivity properties under the assumption of "full visibility"; i.e., they obey a similar "zero-one" law as Erdös-Renyi graphs for some scaling of their parameters, whenever each node is within communication range of other nodes [6, 7].

Recommender Systems Using Collaborative Filtering. Marbach [4] illustrates an application of random-key-graph connectivity for the modeling of recommender systems with collaborative filtering. In his model, there are N users and a pool of P_N objects. Each of the N users picks a set of K_N distinct objects uniformly and randomly from the pool, and ranks each object, where $K_N < P_N$. A recommender class C is defined as a set of users that gives the same ranking to any particular object that is ranked by at least one user of the set. This implies that (1) every subset of users in class C that rank a common object assigns the same ranking to that object, and (2) not every user in class C ranks all objects ranked by others in C; i.e., there exist users in class C that have not ranked objects ranked by others in class C.

A recommender system with collaborative filtering predicts the ranking a user would give an object that the user has not ranked, but that has been ranked by some other users in class C. However, to enable such predictions, the membership

of n users, $n \leq N$, in class C has to be identified. Hence, a question of interest is this: given P_N objects in the pool, how many objects need to be picked and ranked by each of the N users to identify the n members in recommender class C? That is, what is the lower bound of K_N?

To answer the above question, Marbach uses a graph, $G(K_N, P_N, N)$, and models recommender systems as follows. Let a user correspond to a node in graph $G(K_N, P_N, N)$. For any two users (nodes) of the graph, there is an edge between those users if and only if the following conditions hold: (1) there exists at least one object that is ranked by both users; (2) *any* object that is ranked by both users is given the same rank by those users. We note that both P_N and K_N are functions on n, and hence we denote them by P_n and K_n, respectively. Now let graph $G(K_n, P_n, n)$ be the sub-graph of $G(K_N, P_N, N)$, whose vertex set is class C and whose edge set is a subset of the $G(K_N, P_N, N)$ edges connecting vertices of class C. Marbach demonstrates that a necessary condition of identifying the membership of class C is that graph $G(K_n, P_n, n)$ be connected, and that $G(K_n, P_n, n)$ is almost surely connected if $\frac{K_n^2}{P_n} = \omega\big(\frac{\log n}{n}\big)$. Note that it is easy to show the probability that there exists an edge between two nodes of $G(K_n, P_n, n)$ is asymptotically equal to $\frac{K_n^2}{P_n}$ for certain values of n, K_n and P_n. In the context of secure wireless networks, Yağan and Makowski [6, 7] show that when $P_n > 3K_n$ and $\frac{K_n^2}{P_n} \sim \frac{c \log n}{n}$, where $c > 0$ is a constant, the probability that $G(K_n, P_n, n)$ is connected goes to 0 as $n \to \infty$, for $c < 1$, and goes to 1 as $n \to \infty$, for $c > 1$. In other words, for certain values of n, K_n and P_n, the connectivity of graph $G(K_n, P_n, n)$ follows a zero-one law.

Trust Sub-Networks. Graph $G(K_n, P_n, n)$ can also be viewed as a trust sub-network for certificate evaluation in networks without a public-key infrastructure; e.g., ad-hoc networks [8]. In such networks, some certificates cannot be easily validated by clients (e.g., self-signed certificates), yet clients are expected to trust the validity of these certificates for use in security protocols. Let us assume that each of the n users of $G(K_n, P_n, n)$ is a certificate notary which serves a number of clients. These clients ask the notary to evaluate public-key or access control certificates that they receive from foreign sites and that they cannot validate. In a trust network, each notary evaluates K_n certificates it sees of the P_n being used among various sites of the Internet. If any of a notary u's clients asks for any of the K_n already evaluated certificates, u returns that certificate's evaluation (i.e., ranking). If u has not evaluated the certificate needed by one of its clients, u asks other notaries in trust network C for that evaluation. Notary u receives the certificate evaluation from any other notary with probability $\frac{K_n^2}{P_n}$, under the conditions established by the connectivity of graph $G(K_n, P_n, n)$. This evaluation is then sent by u to its client. Notary u returns an exception to its client if the certificate is not evaluated by any notary of the trust network C. We note that the robustness characterization of trust networks can benefit from the "redoubtable" and "unsplittable" properties of random key graphs [9].

An example of a practical system that resembles a trust sub-network is Perspectives [10]. The trusted notaries of the Perspective system can be viewed as the users of trust sub-network C. However, instead of returning a certificate eval-

uation, a trusted notary simply returns parameters of certificate use observed from its vantage point to its clients; e.g., use count observed.

Secure Connectivity to a Trusted Core of a Mobile Ad-hoc Network. Suppose that a subset of the nodes of a mobile ad-hoc network maintains full visibility; i.e., each node is within communication range of all other nodes [6, 7]. We call this subset a "trusted core" since the secure connectivity of its nodes can be assured by probabilistic key pre-distribution similar to that for sensor networks. Since, by definition, ad-hoc networks do not have an infrastructure for secure communication, the secure key connectivity of a trusted core can provide the keying infrastructure for other mobile nodes that connect with each other via the trusted core. If we assume that the trusted core comprises nodes of the user class C defined above and its key connectivity is provided by graph $G(K_n, P_n, n)$ under the choice of parameters determined by Yağan and Makowski [6, 7], then other mobile nodes can connect to each other either separately or via the trusted core. This is possible since the keying of the other mobile nodes can be viewed as the extension of the trusted core and can be performed in the same manner as the dynamic, ad-hoc, extension of a sensor network [2]. That is, each of the added mobile nodes can be pre-keyed with a set of K_n keys drawn from the same pool of size P_n of pre-distributed keys used for the trusted core. As long as the added mobile nodes remain in full visibility of the trusted core, their secure connectivity is assured with probability $\frac{K_{n'}^2}{P_{n'}}$, under the conditions established by the connectivity of the extended graph $G'(K_{n'}, P_{n'}, n')$, where $n' > n$ is the total number of nodes after the mobile nodes are added.

Analysis of Herding Attacks on Hash Functions[1]. Recently, Blackburn *et al.* [5] use the theory of uniform random intersection graphs (i.e., random key graphs) to analyze the complexity of the Kelsey and Kohno "herding attack" [11] on Damgård-Merkle hash-function constructions. In a herding attack, the adversary first commits to a hash value h and then is provided with a prefix p. The attacker needs to find a suffix s which "herds" to the previously committed hash value h. That is, given p, h, and a hash function $H()$, the adversary must find suffix s such that $H(p||s) = h$. To launch a herding attack, Kelsey and Kohno build a 2^k-diamond structure by repeatedly utilizing a collision-finding attack against the hash function. Blackburn *et al.* find a flaw in the original analysis [11] and point out that the message complexity of building the 2^k-diamond structure is k times that suggested by Kelsey and Kohno. In analyzing the complexity of the herding attack, Blackburn *et al.* first derive the threshold of perfect matching in a random key graph and then apply the obtained result to the diamond structure. A perfect matching of a random key graph is a set of pairwise non-adjacent edges where every vertex of the graph lies on one edge of the set.

Further Research. The full visibility assumption, although interesting from a theoretical point of view, is often impractical in mobile ad-hoc (and sensor) networks where communication range is constrained and sometimes unreliable.

[1] This example replaces the analysis of a key-collision attack presented at the SPW 2009 workshop.

A set of interesting research questions arises when one considers the connectivity properties of random key graphs under the "limited visibility" constraint, already encountered in sensor networks [2]. Some robustness properties of random key graphs have already been considered; e.g., redoubtable and unsplittable graphs [9]. However, robustness properties of random-key-graph connectivity under the limited visibility assumption have not been considered to date. Of similar interest are questions of connectivity in random key graphs when some of the links of individual nodes fail.

Note: Since the presentation of this paper in April 2009, additional interesting properties of random key graphs have been investigated; e.g., k-connectivity [12], diameter size [13]. Random key graphs have also been shown to be good candidates for "small world" models [14].

Acknowledgement

This research was supported in part by CyLab at Carnegie Mellon under grants DAAD19-02-1-0389 and MURI W 911 NF 0710287 from the Army Research Office. The views and conclusions contained here are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either express or implied, of ARO, CMU, or the U.S. Government or any of its agencies.

References

1. Blackburn, S., Gerke, S.: Connectivity of the Uniform Random Intersection Graph. Discrete Mathematics, vol. 309, no. 16, pp. 5130-5140 (2009)
2. Eschenauer, L., Gligor, V.D.: A Key-Management Scheme for Distributed Sensor Networks. In: 9th ACM Conference on Computer and Communications Security, Alexandria, VA (2002)
3. Goehardt, E., Jaworski, J., Rybarczyk, K.: Random Intersection Graphs and Classification. In: Studies in Classification, Data Analysis, and Knowledge Organization, 33, H.J. Lens and R. Decker (eds.) Springer Verlag, pp. 67-74, Berlin (2007)
4. Marbach, P.: A Lower Bound on the Number of Rankings Required in Recommender Systems Using Collaborative Filtering. In: IEEE Conference on Information Sciences and Systems, pp 292-297, Princeton University, NJ (2008)
5. Blackburn, S.R., Stinson, D.R., Upadhyay, J.: On the Complexity of the Herding Attack and Some Related Attacks on Hash Functions. `http://eprint.iacr.org/` Report 2010/030 (2010)
6. Yağan, O., Makowski, A.: On the Random Graph Induced by a Random Key Predistribution Scheme under Full Visibility. In: IEEE International Symposium on Information Theory, Toronto, ON (2008)
7. Yağan, O., Makowski, A.: Zero-One Laws for Connectivity in Random Key Graphs. Technical Report, Institute for Systems Research, University of Maryland, February (2009)
8. Eschenauer, L., Gligor, V.D., Baras, J.: On Trust Establishment in Mobile Ad-Hoc Networks. LNCS, vol. 2845/2004, pp. 47-66 (2004)

9. Di Pietro, R., Mancini, L.V., Mei, A., Panconesi, A., Radhakrishnan, J.: Redoubtable Sensor Networks. ACM Transactions on Information and System Security (TISSEC), vol. 11, no. 3, March (2008)
10. Wendlandt, D., Andersen, D., Perrig, A.: Perspectives: Improving SSH-style Host Authentication with Multi-Path Probing. In: USENIX Annual Technical Conference, June (2008)
11. Kelsey, J., Kohno, T.: Herding Hash Functions and the Nostradamus Attack. LNCS, vol. 4004, pp. 183-200 (2006)
12. Rybarczyk K.: Sharp Threshold Functions for Random Intersection Graphs via a Coupling Method. The Electronic Journal of Combinatorics, 18(1), pp. 36-47 (2011)
13. Rybarczyk K.: Diameter, Connectivity, and Phase Transition of the Uniform Random Intersection Graph. Submitted to Discrete Mathematics, July (2009)
14. Yağan, O., Makowski, A.: Random Key Graphs? - Can They Be Small Worlds? In: International Conference on Netowrks & Communications, pp. 313-318 (2009)