Trustworthy Social Informatics, A Relationship-Centric Networking Paradigm

S. Felix Wu
Computer Science Department
University of California, Davis
Davis, CA, USA
sfwu@ucdavis.edu

I. TRUSTWORTHY SOCIAL COMPUTING

A. A Social-Centric View of Cyber Security

While the trustworthy computing community has made significant progress toward a fundamentally secure system, for many critical cyber security issues, we are still lacking provable and yet practical/usable solutions to deal with them. For instance, we still don't have a good solution to automatically identify zero-day, previous unknown virus or to efficiently detect cyber vandalisms within the context of blogs, Wikipedia, or online social networks. Another example is the taint-checking approach, which treats information content directly from the network as un-trusted for certain limited contexts. However, many real world applications today require a more flexible/powerful trust model. On April 21st of this year, McAfee released a virus definition file disabling millions of hosts for an extended period of time. Via this unintentional fault, it was clearly demonstrated that a naïve trust model might lead to a very costly large-scale system/network failure for our society.

On the other hand, online social network/media services took off really well lately. For instance, Facebook allows human social relationships among half billions users (and still growing) being captured/stored digitally and, utilized within the context of many different online applications. This trend of online social networking has introduced new opportunities for social-centric computing paradigms and, at the same time, also raised concerns about new unknown vulnerabilities due to this new paradigm. For instance, personal information being leaked out of Facebook user profiles might be leveraged to make scams/phishing attacks much more believable to the victims. Worms such as Clickjacking/Likejacking or the earlier MySpace Samy worm have simply utilized social relationships to propagate themselves and potentially established another social Botnet infrastructure along the way such as the Koobface botnet. Finally, tools like "Snag-bar under Gamer Unite!" (for Farmville) could be used to spread malicious content to, potentially, millions of accounts in just a few seconds without the manual clicking. In a nutshell, online social networks like Facebook/Twitter seem to worsen our already complicated problem of cyber security and online privacy.

B. Social Informatics System

The term "***social informatics***" here refers to any digitized information related to the online social networks. Social informatics includes the social relationships and their related dynamics. For instance, social relationship might change, while some of them might be temporary, virtually private, or, mission-oriented. It also includes the communication activities or interactions over a graph of social relationships, and the policy, such as privacy, guarding those activities. In practice, the **availability** of online social informatics is constrained by the programming interface (open/propriertory). For instance, Facebook offers Graph/RESTful APIs for an application to access user profile, friendship, wall posts, message inboxes, live/news feeds, application membership, and others allowed via FQL (Facebook Query Language). On the other hand, under the DSL-FAITH API, in addition to the regular Facebook API's, each wall post or message can be associated with a potentially multiple-hop social path with trust scores between 0 and 1 [1]. In other words, the DSL API provides extra social informatics about the quality of a relationship chain being used to deliver a particular content to support the additional needs from certain social computing applications.

In this presentation, we will argue/discuss that why social informatics might play a crucial role for our future computing and network architecture. Our existing networking paradigm is mainly about processing information contents neglecting the relationship dimension. Given the availability of digitized social informatics, we can now possibly process at the same time BOTH the content AND the social relationship information related to either the content or the decision making process. The critical difference here is that now our information processing can be now socially aware and customized for the exactly same range of information content. However, *what types of and how much social informatics can be associated and utilized for information processing depend on the capability and functionalities of the social informatics system itself.* As an example, it is unclear about whether the social informatics provided by Facebook has been well architected to support our current and future social computing needs.

Under the context of social informatics system, we need to consider a number of fundamental security issues for the social computing paradigm. For instance, leveraging social informatics in computing immediately implies a potential violation of user privacy. And, different social informatics designs (e.g., Facebook versus Google+) raise different types of security threats. On the other hand, using Facebook again as an example, the privacy setting between a particular user and his direct friends might be already too loose in comparing to the application privacy setting for Farmville. Therefore, it is challenging to derive accurately a user's probable/specified intention of privacy settings and to determine the appropriate amount/types of social informatics that can be leveraged to support the applications on the behalf of the user. Thus privacy preserving methods within social norms is a difficult yet important research aim.

II. FAITH (FACEBOOK APPLICATIONS: IDENTIFITATION, TRANSFORMATION, & HYPERVISOR)

During my presentation, I will discuss and demonstrate how to architect a trustworthy social informatics system to possibly support the social computing paradigm. We would like to determine ***the appropriate boundary of social informatics*** that should be supported by a trustworthy social informatics system. In order to derive the best possible answer to this research challenge, we have been developing experimentally a set of social computing applications and learn from this experience to advance our knowledge under this direction.. For instance, we will show how to leverage social informatics to determine the trust toward a piece of information (e.g., software program or information content) from a user, a social community, or a software agent. This academic study will not only extend the notion of social trust but also offer potentially alternate solutions to important trustworthy computing problems such as unknown virus, spam/scam, or DDoS. We will also discuss the possibilities of future Internet architecture and operating system kernel design based on the social computing paradigm.

To realistically support the trustworthy social informatics, we have recently developed the FAITH system [3], like a social kernel, to monitor and control the social informatics and activities between social informatics service providers, such as Facebook and Google+ and their associated applications. Our goal is to detect and control from the user's perspective whether certain Facebook applications request an anomalous amount of social informatics for a particular group of users.

Applications function as social informatics consumers, which leverage social information to provide valuable online social interactions among users. In contrary, Facebook functions as a social informatics provider, which offers FAITH its social graph and informatics. FAITH functions differently depending on different points of view. To Facebook, FAITH is nothing but an ordinary application fetching social informatics. To applications, FAITH supplies the transformed social informatics upon request. From the users' perspective, FAITH is a multi-functional application-level proxy. It transforms and logs the social informatics upon users' requests to manage social information more securely and transparently. FAITH allows users to specify rules which transform their social informatics used by OSN kernel such that the social router utilizes the transformed social graph instead of the original Facebook social graph. During the process of producing content, applications may send multiple requests to FAITH to access the social informatics of Facebook or the functionalities of web services. In the case of requesting social informatics, FAITH sends requests to Facebook and then logs and passes back the transformed informatics to applications. In the case of utilizing the functionality of the web services, FAITH sends other requests to services, and also logs and passes back the results to applications. Currently, FAITH has 15 applications of various kinds with ~100 active users and it's running on the GENI (Global Environment for Network Innovations) testbed.

III. SOCIAL-CENTRIC FUTURE INTERNET ARCHITECTURE

We believe that the Internet should not merely communicate bits and bytes syntactically. The Internet is also about communicating ***relationships*** based on social informatics among users and a graph of related content [2]. While packets are being forwarded at the network layer, relationship status is being updated and, furthermore, leveraged to conduct future communication activities. Using DDoS as an example, from the traditional network-layer perspective, some resources (such as bandwidth of a critical link) have been over-consumed syntactically. From the relationship perspective, our relationship status has been anomalously updated and consumed in a very short period of time. A critical difference here, though, is that, at the core of the network, we need to continuously track the attack signatures, if those do exist, while, at the relationship layer, users can more precisely inform the core regarding the expected and desired relationship updates. As human relationships are relatively more stable than the dynamics of bit/byte patterns, the resolution of DDoS might be much easier in the relationship domain than in the packet domain. Furthermore, if the attacker likes to mimic "natural human relationship" at the relationship layer (to evade from being detected), then he is forced to blast much more traffic within a smaller number of relationship paths. By trying to evade from social profiling, the attacker will probably be forced to reveal his attack signatures at the network layer. Thus, the network layer should make routing decisions also based on relationship information, instead of only relationship-insensitive identifiers such as destination IP addresses or URLs.

[1] M. Spear, X. Lu, N. Matloff, and S. F. Wu, KARMANET: LEVERAGING TRUSTED SOCIAL PATHS TO CREATE JUDICIOUS FORWARDERS, In Proceedings of the First International Conference on Future Information Networks (IFCIN '09), pp. 218-223, Beijing, China, 2009.

[2] Networking: Four ways to reinvent the Internet, by Katharine Gammon, Nature 463, 602–604 (03 February 2010) | doi:10.1038/463602a

[3] Design and Implementation of FAITH, An Experimental System to Intercept and Manipulate Online Social Informatics, R. Lee, R. Nia, S. Ye, J. Hsu, K. Levitt, J. Rowe and S. F. Wu, ASONAM'2011.