Ecole Doctorale STIC

Sciences & Technologies de l'Information et de la Communication

ED STIC - Proposition de Sujets de Thèse

pour la campagne d'Allocation de thèses 2015

Axe Sophi@Stic :	SystèmesRéseaux|
Titre du sujet :	How social links are created on Twitter, and what is their privacy implication?
Mention de thèse :	Informatique
HDR Directeur de thèse inscrit à l'ED STIC :	Arnaud Legout

Co-encadrant de thèse éventuel :

Nom :	
Prénom :	
Email :	
Téléphone :	

Email de contact pour ce sujet :	arnaud.legout@inria.fr
Laboratoire d'accueil :	INRIA

Description du sujet :

Twitter is the most popular micro-blogging service in the world. It allows its users to exchange short messages (tweets) that are limited to 140 characters. It was created to enable people to find out what is currently happening with people and organizations they are interested in. The relation between users on Twitter is different from classical social networks like Facebook. Instead of bidirectional friendship link that are initiated by one user and accepted by another, Twitter uses the concept of following. Users can follow other users they are interested in, which means they subscribe for all the messages they sent. So, the links on Twitter are unidirectional, if someone follows you, you don't need to follow back. Twitter is a very interesting object of study because the unidirectional model of relationship is the closest to real-life communications, thus it has a huge societal impact.

The goal of this Ph.D. is to study how to influence users on Twitter by exploiting a transitive relationship called chain of trust. Typically, a chain of trust corresponds to follow relationships, but usually goes from a highly popular user to anonymous users. We propose to explore how social links are created and how this knowledge can be exploited by a malicious user to infringe privacy and influence information propagation. This work will involve a combination of measurement and modeling work. We will use statistics to model the reach of information introduced in the trust chain and we also propose to use game theory to model and understand why information is relayed by users. The combination of both will allow us to optimize defenses against trust chain manipulations. To explore the phenomenon in practice (and to calibrate the models), the student will have the opportunity to work on a unique datasets we collected in 2012 [1]. This dataset represents the entire Twitter social graph with more than 500 million accounts and 24 billion links.

[1] The Complete Picture of the Twitter Social Graph by Maksym Gabielkov and Arnaud Legout http://hal.inria.fr/hal-00752934

English version: