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 2011

Titre du sujet :	Monitoring the quality and evolution of the Internet access
Mention de thèse :	Informatique
HDR Directeur de thèse inscrit à l'ED STIC :	Barakat Chadi

Co-encadrant de thèse éventuel :

Nom :	
Prénom :	
Email :	
Téléphone :	

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

Description du sujet :

The Internet usage has shifted from communications between machines into content lookup and retrieval. Nowadays, Internet users spend most of their time looking for content (photos, clips, blogs, etc) or communicating with friends without caring about where they are, the most important is the content of the communication itself. We also see a large use of the Internet for multimedia applications whether streaming or real time. This shift has two main consequences. First, it changes the nature of Internet traffic from the traditional traffic we knew which was mainly composed of WEB traffic. On one side, the traffic is increasing and on the other side it is becoming more symmetric. Furthermore, more and more applications are used with new behavior as encrypting the communications or using non standard port numbers. The second consequence of this change is the appearance of big content players as google and the social networks that propose to users plenty of facilities to manage and share their content. These major content players make lot of profit from the deployed infrastructure and attract to themselves alone most of the

activities over the Internet. As a consequence, the ISPs start looking for solutions to understand the users' behaviors and the types of applications they use. They also start looking for solutions that allow them to play a role in the content distribution either by hosting content servers and making them close to clients, by proposing content themselves (the triple play formula including TV, Phone and Internet) or simply by discriminating some user applications to protect their network and to privilege some communications over others. In view of these changes, it becomes very important to understand their nature and to evaluate their impact on the three major players, the user, the content provider and the ISP itself.

The candidate will first go through the literature to understand the state of art in the field. The second phase will be dedicated to propose solutions based on traffic measurements (probing or passive measurements) that allow to infer the behavior of each player. The third phase, which can/should be carried out in parallel with the second phase is to deploy these tools, to model them and to test them by both simulations and experimentations. For instance, one might need tools to infer the applications used by users even in case of encryption and usage of non standard port numbers. One will also need solutions to detect if the ISP is practicing any discrimination, either positive or negative, on part of the traffic or on some specific applications (e.g. BitTorrent, VoIP, streaming). There is equally urgent need for end-user solutions to detect any deviation in the quality of the Internet access and to try to infer the root cause of this deviation (e.g. local problem, in the DSLAM or the router of the ISP, wireless, cable, elsewhere). These solutions should be accompanied by proofs that make the other parties agree on them. The final objective of this research will be to propose tools that make the game more transparent and that allow each player to be aware of the other players' behavior, especially those which are not documented in the signed contract. We will not take position towards this or that behavior, we will only propose tools that infer the behavior of each player and make it available to the others. Our solutions should be designed to "run in practice" and should be accompanied by an evaluation by modeling and trace-driven emulation/simulation. A deployment in the wild is also envisaged.

In summary, this thesis should lead to a better understanding of the quality of the Internet access and of the behavior of end users and operators. It should also provide solutions to detect any problem at the access, to understand its root cause, and to alert the concerned players to take the required actions.

The candidate should have a solid background in Internet protocols and basic programming languages C/C++. It is also required to have a solid background in statistics and network modeling.

References:

1- MLAB: Measurement Lab. URL: http://www.measurementlab.net/
2- Partha Kanuparthy, Constantine Dovrolis, DiffProbe: Detecting ISP Service Discrimination, in

the Proceedings of the IEEE INFOCOM Conference, March 2010.
3- Mohamad Jaber and Chadi Barakat, ''Enhancing Application Identification By Means Of Sequential Testing'', in proceedings of IFIP/TC6 Networking Conference, Aechan,
Germany, May 2009.
4- Jacobson, V., Smetters, D. K., Thornton, J. D., Plass, M. F., Briggs, N., Braynard, R. "Networking named content". Proceedings of the 5th ACM International Conference on
Emerging Networking Experiments and Technologies (CoNEXT 2009); 2009 December 1-4; Rome, Italy.
5- Christian Kreibich, Nicholas Weaver, Boris Nechaev, and Vern Paxson. "Netalyzr: illuminating the edge network". In Proceedings of the 10th annual conference on Internet
measurement (IMC '10). ACM, New York.
6- Jon Crowcroft. "Net neutrality: the technical side of the debate: a white paper". SIGCOMM Comput. Commun. Rev. 37, 1 (January 2007).

English version: