



## Rapport de thèse

Ashwin Rao

### *Improving Transparency and End-User Control in Mobile Networks*

La thèse de Ashwin Rao traite principalement du contrôle des informations personnelles dans les applications mobiles. Il s'agit d'un sujet d'actualité qui mérite une forte attention et qui revêt une importante dimension pratique. La thèse est composée de cinq chapitres qui présentent trois contributions successives. La première est centrée sur la réalisation de la plate-forme « Meddle » qui vise à redonner transparence et contrôle aux utilisateurs de mobiles. La seconde est constituée de l'utilisation de Meddle afin de recenser les fuites de données personnelles sur un jeu d'applications populaires. La dernière exploite Meddle afin de caractériser le trafic YouTube et les stratégies de streaming associées.

L'avènement des mobiles et de leurs nombreuses applications a fait émerger un énorme marché dont les acteurs - concepteurs des systèmes d'exploitation et des applications, opérateurs des "stores", FAI - convoitent les bénéfices en monnayant les données des utilisateurs. De plus, l'extraction de ces informations de nature personnelle consomment des ressources tant au niveau de l'énergie des terminaux que de la bande passante.

La contribution principale de la thèse est d'avoir proposé une plate-forme « Meddle » qui offre la possibilité de mesurer et de contrôler le trafic internet mobile de façon à accroître la transparence pour l'utilisateur. Le principe adopté est de rediriger l'ensemble du trafic mobile vers un équipement intermédiaire qui est inséré sur le chemin entre le terminal et le serveur. La solution est fondée sur la redirection du trafic mobile sur un VPN utilisé de manière native. Cette approche présente l'avantage d'être indépendante du système d'exploitation et de l'opérateur. Néanmoins, cette solution permet également d'intercepter le trafic utilisateur qui aurait été codé au préalable par une application. Meddle est donc naturellement utilisée pour identifier les fuites d'informations personnelles dans les applications mobiles majeures. Une seconde contribution concerne l'analyse du trafic YouTube. Plusieurs observations intéressantes ont pu être réalisées. Elles montrent l'impact de celles-ci sur les performances du système dans son ensemble.

Le chapitre 1 présente les bases nécessaires pour la compréhension du contexte et des contributions de ce travail, de même que les références essentielles à l'état de l'art. Il s'agit de présenter les difficultés et les principes fondateurs des solutions adoptées.

Le chapitre 2 présente la solution retenue pour redonner transparence et contrôle aux utilisateurs de mobiles. Le principe de redirection utilisé est clairement exposé de même que les principales difficultés techniques qu'il a fallu traiter. Cette solution semble efficace afin d'intercepter et d'analyser le trafic produit sans intervention de l'utilisateur ni de modification de son terminal (à part la configuration du tunnel). Les limitations éventuelles de la solution sont clairement identifiées et discutées (par exemple le fait de ne pas supporter IPv6 ou le blocage éventuel des VPNs par certains ISPs). Un intérêt particulier a également été consacré à une évaluation de la charge induite par la solution « Meddle » sur des paramètres importants tels que le délai ou la consommation d'énergie. Enfin, les aspects juridiques sont essentiels et sont brièvement discutés. Le fait que cette solution soit « centralisée » peut éventuellement être dissuasif ?



Le chapitre 3 apporte une description intéressante du diagnostic d'applications. Plusieurs jeux de données complémentaires sont utilisés. L'effort se concentre sur la classification du trafic ainsi que sur les fuites de données privées observées sur ces applications. L'analyse est conduite en utilisant une analyse de traces «enregistrées (« offline »). Tcpdump est utilisé pour sauvegarder l'ensemble du trafic intercepté par Meddle. La classification de trafic est principalement réalisée en s'appuyant sur les entêtes HTTP et SSL afin d'identifier les applications, les OSs et les services. Des techniques originales sont développées afin d'améliorer l'efficacité de la correspondance permettant la classification des applications. Enfin, l'étude se poursuit par un recensement et un filtrage des fuites d'informations personnelles utilisées par les principales applications mobiles.

Le chapitre 4 poursuit l'étude exploitant Meddle en vue de mener une analyse approfondie de la caractérisation du trafic YouTube. Il est constaté que le trafic dépend fortement des navigateurs utilisés, des technologies Flash/HTML5, et que les dernières versions ont un comportement très différent des solutions passées. De même, le trafic généré varie suivant que l'application communique en WiFi ou en utilisant le réseau cellulaire. L'étude est conduite aussi bien dans le cas de PCs que de mobiles (iOS et Android).

Le chapitre 5 conclut la thèse, résume les contributions et ouvre de nouvelles perspectives de recherche.

En résumé, l'ensemble des travaux démontre une parfaite maîtrise du sujet ainsi que la capacité de conduire un travail finalisé avec une composante développement significative. Les objectifs sont annoncés et permettent une construction efficace de l'exposé des travaux. La démarche est claire et progressive. Le sujet traité est pertinent. Il est de nature expérimentale et permet d'apprendre des comportements importants de systèmes de communications présents très largement utilisés. Les résultats ont un intérêt pratique immédiat. Par ailleurs, la thèse est bien rédigée. L'ensemble fournit une contribution solide qui a été bien publiée.

Pour toutes ces raisons, j'émet un avis très favorable à l'attribution du grade de docteur en Informatique de l'Université de Nice Sophia-Antipolis à Ashmin Rao.

Paris, 28 Novembre 2013

Serge FDIDA  
Professeur, Université Pierre et Marie Curie,  
Sorbonne Universités