

# UNIVERSITE CLAUDE BERNARD – LYON I

## DIPLÔME NATIONAL DE DOCTORAT (Arrêté du 25 mai 2016)

Date de la soutenance : 14 décembre 2016

Nom de famille et prénom de l'auteur : **Zohra SAOUD**

Titre de la thèse : « Approche robuste pour l'évaluation des ressources sur le Web. »

### Résumé de la thèse

Cette thèse en Informatique s'inscrit dans le cadre de gestion de la confiance, et plus précisément des systèmes de recommandation. Ces systèmes sont généralement basés sur les retours d'expériences des utilisateurs (i.e., qualitatifs/quantitatifs) lors de l'utilisation des ressources sur le Web (ex. films, vidéos, et service Web). Les systèmes de recommandation doivent faire face à trois types d'incertitude liés aux évaluations des utilisateurs, à leur identité et à la variation des performances des ressources au fil du temps. Nous proposons une approche robuste pour évaluer la confiance en tenant compte de ces incertitudes.

Le premier type d'incertitude réfère aux évaluations. Cette incertitude provient de la vulnérabilité du système en présence d'utilisateurs malveillants fournissant des évaluations biaisées. Pour pallier cette incertitude, nous proposons un modèle flou de la crédibilité des évaluateurs. Ce modèle, basé sur la technique de clustering flou, permet de distinguer les utilisateurs malveillants des utilisateurs stricts habituellement exclus dans les approches existantes.

Le deuxième type d'incertitude réfère à l'identité de l'utilisateur. En effet, un utilisateur malveillant a la possibilité de créer des identités virtuelles pour fournir plusieurs fausses évaluations. Pour contrecarrer ce type d'attaque dit Sybil, nous proposons un modèle de filtrage des évaluations, basé sur la crédibilité des utilisateurs et le graphe de confiance auquel ils appartiennent. Nous proposons deux mécanismes, l'un pour distribuer des capacités aux utilisateurs et l'autre pour sélectionner les utilisateurs à retenir lors de l'évaluation de la confiance. Le premier mécanisme permet de réduire le risque de faire intervenir des utilisateurs multi-identités. Le second mécanisme choisit des chemins dans le graphe de confiance contenant des utilisateurs avec des capacités maximales. Ces deux mécanismes utilisent la crédibilité des utilisateurs comme heuristique.

Afin de lever l'incertitude sur l'aptitude d'une ressource à satisfaire les demandes des utilisateurs, nous proposons deux approches d'évaluation de la confiance d'une ressource sur le Web, une déterministe et une probabiliste. La première consolide les différentes évaluations collectées en prenant en compte la crédibilité des évaluateurs. La deuxième s'appuie sur la théorie des bases de données probabilistes et la sémantique des mondes possibles. Les bases de données probabilistes offrent alors une meilleure représentation de l'incertitude sous-jacente à la crédibilité des utilisateurs et permettent aussi à travers des requêtes un calcul incertain de la confiance d'une ressource.

Finalement, nous développons le système WRTrust (Web Resource Trust) implémentant notre approche d'évaluation de la confiance. Nous avons réalisé plusieurs expérimentations afin d'évaluer la performance et la robustesse de notre système. Les expérimentations ont montré une amélioration de la qualité de la confiance et de la robustesse du système aux attaques des utilisateurs malveillants.