

Abstract

Nell'era corrente, grazie alla digitalizzazione dei processi aziendali e delle pubbliche amministrazioni, sono disponibili grandi collezioni di dati. Gli utenti sono diventati fornitori diretti dei dati, ad esempio quando pubblicano contenuti sui social network o utilizzano servizi sul web, i quali diventeranno proprietà delle aziende che gestiscono tali servizi. Per questo motivo, gli utenti devono essere consapevoli delle problematiche di privacy dei propri dati e le aziende devono garantire la protezione dei dati personali, in conformità con le nuove regolamentazioni. Tuttavia, è necessario non limitare il trattamento dei dati da parte di aziende/enti pubblici. Pertanto, è necessario sviluppare metodologie dedicate all'identificazione di possibili minacce alla privacy durante le attività di utilizzo dei servizi online e sviluppare strategie di privatizzazione che non degradino l'informazione intrinseca dei dati.

Questa tesi fornisce prove sperimentali inerenti problematiche riscontrabili quando gli utenti forniscono dati personali per accedere ai servizi fruibili online. Allo scopo di rendere gli utenti consapevoli riguardo la privacy dei dati, la tesi descrive nuove metodologie/strumenti per supportare le aziende nel trattamento dei dati personali. In particolare, le metodologie proposte sfruttano le correlazioni di dati, espresse in termini di dipendenze funzionali rilassate (RFD), per definire strategie di privatizzazione volte a salvaguardare la privacy dell'utente e per discriminare account malevoli nel dominio dei social network. Infine, sono stati progettati e implementati due strumenti automatici per offrire supporto, in termini di privatizzazione, quando gli utenti utilizzano servizi fruibili online.

Abstract

Nowadays, thanks to the digitalization of business processes and public administrations, many significant Big data collections are available. Users are direct suppliers of data when publishing contents on social networks. However, when using a service on the web, users must often provide their data, which will become property of the company running the service. To this end, users need to be aware of the privacy issues related to the management of their data, whereas companies need to ensure the protection of users' personal data, also according to new laws and regulations issued by governments. On the other hand, there exists the necessity not to limit the processing of data by companies and other public institutions. Thus, it is necessary to devise methods devoted to the identification of possible privacy threats during users' online activities, and to develop privatization strategies that possibly do not downgrade the significance of data.

This dissertation provides experimental evidence of several threats for users when providing their personal data for accessing online services, aiming to increase their awareness, and it describes new methodologies and tools to support companies when processing personal data of their users. In particular, the proposed methodologies exploit data correlations expressed in terms of relaxed functional dependencies (RFDs) to define privatization strategies, aiming to safeguard user's privacy, and to detect malicious accounts in social networks. Finally, two automatic tools have been designed and implemented to help users better understand privacy threats during their online activities.