Hidden Is Safe? Location Protection against Machine-Learning Prediction Attacks in Social Networks

Xiao Han, Leye Wang, and Weiguo Fan

Abstract

User privacy protection is a vital issue in online social networks (OSNs). Even though users often intentionally hide their private information in OSNs, adversaries may conduct *prediction attacks* to predict the hidden information with advanced machine learning techniques. The private information that users intend to hide may still be at risk of being exposed. By taking *current city* on Facebook as a case, we propose a solution to estimating and managing the exposure risk of users' hidden information. First, we simulate an aggressive prediction attack by proposing a new current city prediction framework that integrates location indications from users' various published information including demographic attributes, behaviors, and relationships with advanced state-of-the-art machine learning algorithms. Second, we study the prediction attack results to model the pattern of prediction correctness (as correct predictions lead to information exposures), and construct an *exposure risk estimator*. The proposed exposure risk estimator can not only notify users of the exposure risks of their hidden current city, but also help them mitigate the exposure risk by overhauling and selecting countermeasures. Moreover, it can improve OSNs' privacy management by facilitating OSNs' empirical studies on the exposure risks of collective users. While taking the current city as a case, this work sheds light on the protection of other hidden information against machine-learning prediction attacks, and reveals several insightful implications for both practice management and future research.

Keywords: Private information protection, personal exposure risk, machine-learning, location prediction attack