ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

Δρακωνάκης Κωνσταντίνος
Μεταπτυχιακός Φοιτητής

Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης
Επόπτης Μεταπτ. Εργασίας: Καθηγητής, Ε. Μαρκάτος

Παρασκευή, 5/10/2018, 18:00
Αίθουσα Τηλεδιάσκεψης Κ206,Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης

“ Διερεύνηση των Επιπτώσεων στην Ιδιωτικότητα από Δεδομένα Γεωγραφικής Τοποθεσίας σε Δημόσιες Ροές Δεδομένων ”

ΠΕΡΙΛΗΨΗ

Η δημοσιοποίηση δεδομένων που αφορούν την γεωγραφική τοποθεσία των χρηστών αποτελεί σημαντικό κίνδυνο κατά της ιδιωτικότητάς τους, καθώς αυτά τα δεδομένα μπορούν να οδηγήσουν στην κατάργηση της ανωνυμίας των χρηστών, στην αποκάλυψη ευαίσθητων προσωπικών πληροφοριών, και σε κάποιες περιπτώσεις, ακόμη και σε απειλές κατά της ζωής και της σωματικής τους ακεραιότητας. Στην παρούσα διπλωματική εργασία παρουσιάζουμε τον LPAuditor, ένα εργαλείο σχεδιασμένο ώστε να διεξάγει ενδελεχή αξιολόγηση των κινδύνων κατά της ιδιωτικότητας που προκαλούνται από την ευρεία διάθεση μεταδεδομένων προσδιορισμού γεωγραφικής τοποθεσίας. Αρχικά παρουσιάζουμε με ποιον τρόπο το σύστημά μας μπορεί να προσδιορίσει τις δύο πιο κύριες τοποθεσίες των χρηστών, συγκεκριμένα αυτές του χώρου κατοικίας και εργασίας τους, με μια άνευ προηγουμένου ακρίβεια, σε επίπεδο ταχυδρομικής διεύθυνσης. Με την χρήση δεδομένων από το Twitter δείχνουμε ότι οι τεχνικές μας ξεπερνούν σε αποτελεσματικότητα όλες τις προηγούμενες προσεγγίσεις κατά 18,9% - 91,6% για τον προσδιορισμό κατοικίας και κατά 8,7% -21,8% για προσδιορισμό του χώρου

εργασίας. Στη συνέχεια, παρουσιάζουμε μία πρωτότυπη προσέγγιση για τον αυτοματοποιημένο προσδιορισμό γεωγραφικών τοποθεσιών που έχει προηγουμένως επισκεφτεί ο χρήστης, οι οποίες αποκαλύπτουν περαιτέρω ευαίσθητες πληροφορίες που δεν γνωστοποίησε ο χρήστης εκ προθέσεως (συγκεκριμένα τοποθεσίες σχετικές με θρησκεία, ζητήματα υγείας και σεξουαλικού προσανατολισμού). Γενικά, δείχνουμε ότι τα μεταδεδομένα που αφορούν γεωγραφικές τοποθεσίες στο Twitter , μπορούν να προσφέρουν επιπλέον πληροφορία που σε συνδυασμό με το περιεχόμενο του tweet οδηγεί στην αποκάλυψη ευαίσθητων πληροφοριών που δεν αποσκοπούσε να αποκαλύψει ο χρήστης.

Ακόμη, εξετάζουμε την αναντιστοιχία που υπάρχει μεταξύ των ενεργειών των χρηστών και της έκθεσης πληροφοριών και διαπιστώνουμε ότι παλαιότερες εκδόσεις των επίσημων εφαρμογών του Twitter για κινητές συσκευές εφάρμοζαν μία πολιτική μη σχεδιασμένη για την προστασία της ιδιωτικότητας των χρηστών. Συγκεκριμένα, οι εφαρμογές αυτές εφάρμοζαν ακριβή μεταδεδομένα τοποθεσίας σε tweets που οι χρήστες είχαν χαρακτηρίσει με τοποθεσίες χαμηλότερης ακρίβειας (π.χ. πόλη). Τα αποτελέσματα της μελέτης μας δείχνουν ότι οι χρήστες είναι ιδιαίτερα προσεκτικοί σε ό,τι αφορά την ιδιωτικότητά τους και την αποκάλυψη δεδομένων ακριβής γεωγραφικής τοποθεσίας, το οποίο τονίζει ακόμη περισσότερο τις επιπλοκές της εν λόγω πολιτικής. Κατ᾽ ακρίβειαν, όταν οι χρήστες έχουν την δυνατότητα να επιλέξουν συγκεκριμένα το επίπεδο ακρίβειας των δεδομένων γεωγραφικής τοποθεσίας που διαμοιράζονται, παρατηρούμε μια μείωση της τάξης του 94.6% στα tweets που περιέχουν συντεταγμένες GPS. Συμβαδίζοντας με τρέχουσες προσπάθειες ώστε να δοθεί περισσότερος έλεγχος στους χρήστες σε ό,τι αφορά τα προσωπικά τους δεδομένα, ο LPAuditor μπορεί να αξιοποιηθεί από μεγάλες υπηρεσίες και να προσφερθεί σαν ένα εργαλείο που θα ενημερώνει τους χρήστες για ευαίσθητες πληροφορίες που έμμεσα αποκαλύπτονται από μεταδεδομένα γεωγραφικής τοποθεσίας τους.

Drakonakis Konstantinos
M.Sc. Thesis

Computer Science Department
University of Crete
Master's Thesis Supervisor: Professor, E. Markatos

Friday, 5/10/2018, 18:00
Room K206, Computer Science Dept., University of Crete

“Exploring the Privacy Implications of Location (Meta)Data in Public Data Streams ”

ABSTRACT

The exposure of location data constitutes a significant privacy risk to users as it can lead to de-anonymization, the inference of sensitive information, and even physical threats. In this work we present LPAuditor, a tool that conducts a comprehensive evaluation of the privacy loss caused by publicly available location metadata. First, we demonstrate how our system can pinpoint users’ key locations at an unprecedented granularity by identifying their actual postal addresses. Our experimental evaluation on Twitter data highlights the effectiveness of our techniques which outperform prior approaches by 18.9%-91.6% for homes and 8.7%-21.8% for workplaces. Next we present a novel exploration of automated private information inference that uncovers “sensitive” locations that users have visited (pertaining to health, religion, and sex/nightlife). We find that location metadata can provide additional context to tweets and thus lead to the exposure of private information that might not match the users’ intentions. We further explore the mismatch between user actions and information exposure and find that older versions of the official Twitter apps follow a privacy-invasive policy of including precise GPS coordinates in the metadata of tweets that users have geotagged at a coarse-grained level (e.g., city). The implications of this exposure are further exacerbated by our finding that users are considerably privacy cautious in regards to exposing precise location data. When users can explicitly select what location data is published, there is a 94.6% reduction in tweets with GPS coordinates. As part of current efforts to give users more control over their data, LPAuditor can be adopted by major services and offered as an auditing tool that informs users about sensitive information they (indirectly) expose through location metadata.