University POLITEHNICA of Bucharest

Faculty of Automatic Control and Computers

Department of Computer Science and Engineering

Rezumat

Doctoral Thesis

Diff-Anonym Algorithm to Provide Privacy for Patients' Location based Services

Based on Big Data in Cloud

Author: **Imad Al - Tameemi**

Scientific Supervisor: **Prof. dr. ing. Nicolae Ţăpuş**

Bucharest 2019

Rezumat

Utilizarea telefoanelor mobile, a computerelor, cloud computing etc., din ultimii ani, a generat încredere zero. Confidențialitatea informațiilor a fost ignorată în dezvoltarea rapidă a senzorilor, sistemelor, tehnicilor, puterii de procesare, capacităților de stocare a dispozitivelor și a rețelelor. Peste 7 miliarde de persoane utilizează în întreaga lume computere, telefoane mobile, servicii de cloud computing și alte tehnologii. Toate aceste tehnologii sunt conectate și sunt capabile să colecteze și să facă schimb de informații. O persoană are dreptul să selecteze informațiile personale ce pot fi divulgate. Informațiile despre locație sunt o parte importantă a informațiilor private referitoare la o persoană. Acest tip de informatii reprezinta obiectul principal al prezentului studiu, ce are în vedere cele mai importante preocupări legate de transmiterea datelor despre istoricul locațiilor către terți, în funcție de preferințele utilizatorilor. Utilizarea din ce în ce mai mare a rețelelor GSN sau a rețelelor geosociale ca Twitter, Instagram sau Facebook, a dus la afișarea publică a locației utilizatorului, prin serviciile de „check-in" ale rețelelor. Pe baza acestei informații despre locație, acțiunile utilizatorului pot fi refăcute și chiar poate fi anticipată localizarea persoanei în viitor. Astfel, riscul divulgării neautorizate este mare în legătură cu locații sensibile, cum ar fi spitale, adresa de domiciliu sau locația băncilor. Sunt posibile accesări neautorizate ale serverele de stocare a datelor, considerate servere lipsite de încredere, utilizate pentru colectarea informațiilor personale. Serverele de stocare prezintă, de asemenea, riscul serverelor cu un singur punct, caz în care un atac asupra unui singur dispozitiv poate compromite informațiile private stocate pe acest dispozitiv. În caz de dezvăluire a informațiilor private, adresa de domiciliu a unui utilizator ar putea fi descoperită, utilizatorul ar putea primi reclame nesolicitate sau chiar ar putea fi descoperite afecțiuni medicale ale persoanei. Drept urmare, aceste experiențe ar putea fi foarte neplăcute sau chiar ar putea genera complicații. Persoanele care au acces la informații despre locație aparțin unor grupuri sociale diferite, cum ar fi colegii, prietenii sau familia. O parte din interesul nostru este, de asemenea, tipul de informații distribuite către aceste grupuri. O persoană poate alege să partajeze anumite informații despre locație doar cu unele cercuri, de exemplu check-in-urile legate de viață de noapte nu ar trebui să fie vizibile pentru grupul familiei. Nivelurile de confidențialitate a datelor modifică modelele de afaceri, de la

asigurări sociale la comerț cu amănuntul și agricultură. Informațiile referitoare la toate activitățile sunt colectate la rate foarte mari, ceea ce oferă posibilități de ajustare a practicilor și procedurilor pentru a crește eficiența.

Într-un mediu comercial, se așteaptă vânzarea și distribuirea unui articol după dezvoltarea sa. Marii retaileri sunt capabili să stabilească nevoile de achiziție ale potențialilor clienți pe baza informațiilor despre clienți. Utilizarea de telefoane mobile, computere, servicii cloud computing etc., în ultimii ani, au generat încredere zero. A fost creat un concept nou, numit detectare participativă, în care indivizii sunt capabili să facă schimb prin tehnologiile pe care le folosesc, de informații sesizate în mediul înconjurător. Metodele și algoritmii de confidențialitate nu sunt suficiente pentru a proteja confidențialitatea individului. Există numeroase metode disponibile, cum ar fi tehnologii și sisteme care permit analiza constantă a comportamentului, acțiunilor și preferințelor, alături de tehnicile de extragere a datelor, analiză și catalogare. Agențiile guvernamentale, întreprinderile sau organizațiile care colectează și analizează date, cum ar fi întrebările de căutare, istoricul de apeluri telefonice sau informațiile despre locațiile utilizatorilor, au fost afectate de atacuri de securitate. Utilizatorii au început să nu mai aibă încrede în aceste entități, ceea ce a dus la evitarea utilizării a numeroase aplicații sau la introducerea de date false. Utilizatorii se simt, în general, mai confortabili atunci când aplicațiile solicită mai puține informații sau sunt opționale. O altă posibilitate este de a permite utilizatorilor să obscureze informații, dacă sunt acceptate de sistem. Această disertație prezintă, de asemenea, cele mai importante elemente care ar trebui păstrate în ceea ce privește informațiile despre locație sau datele personale. Opinia generală a utilizatorilor este că, prin ascunderea informațiilor personale, pot fi evitate atacurile de confidențialitate. Sunt mult mai puțin preocupați de informațiile despre locație. Primul capitol al disertației prezintă o imagine de ansamblu a fundalului subiectelor tezei și principalele preocupări ale confidențialității clienților pentru serviciile bazate pe locație în date mari. Al doilea capitol prezintă analiza literaturii pe teme în general. În plus, explicăm rezultatele cercetării luate în considerare și examinăm cercetările care vizează soluții pentru dezvăluirea generală a datelor versus dezvăluirea informațiilor despre locație în ceea ce privește impactul asupra clienților din punctul de vedere al confidențialității. În al treilea capitol este considerat efectul asupra confidențialității clienților privind divulgare a informațiilor generale versus divulgarea informațiilor despre locație. Analiza este realizată în două etape – crearea și publicarea uui chestionar online și analizarea datele colectate folosind acest sondaj. Scopul analizei a fost

identificarea efectului dezvăluirii informațiilor despre locație în comparație cu efectul dezvăluirii informațiilor generale. Pe baza rezultatelor, am ajuns la concluzia că orice dezvăluire a datelor are efecte negative. Al patrulea capitol include o prezentare a metodei propuse care asigură protecția datelor personale pentru datele normale și mari.