Encrypted Databases:
New Volume Attacks against Range Querie

Aceasta lucrare se bazeaza pe exploatarea erorii umane de a porni de la premisa ca datele criptate sunt complet in siguranta. Scopul discutiei nu se axeaza pe algoritmi si metode de criptare, nici macar pe criptarea in sine, ci pe un aspect al datelor vesnic prezent. Este vorba de volumul acestora, sau mai bine spus, de relatia dintre volum si numar de inregistrari. In mod evident, criptarea datelor poate expanda sau reduce volumul de date, dar aceasta nu schimba un lucru legat de acesta: proportionalitatea directa dintre numarul de inregistrari si volumul de date.

Orice baza de date si aplicatie construita peste aceatsa pentru a facilita accesu la date, are ca scop simplificarea procesului de interogare. In urma unei interogari, un numar de inregistrari sunt obtinute din baza si transmise catre aplicatie. De cele mai multe ori, aplicatia si serverul vor rula pe masini (sau containare) diferite. In acest caz, comunicarea se va face printr-o interfata, deci date vor fi transmise printr-un mediu mai mult sau mai putin expus. In acest punct, prin monitorizarea traficului, se pot face estimari legate de volumul de date transmise, chiar daca gradul de precizie poate varia. Cu cate mai multe astfel de transmisii sunt interceptate, cu atata sansele sa se obtina informatii relevante creste.

Scopul nu este spargerea criptarii, nici obtinerea datelor efective, ci estimarea unor cantitati. Aceasta tehnica poate parea inofensiva, insa in domenii in care orice fel de scurgere reprezinta o problema critica, gradul de risc creste exponential. Spre exemplu, prin astfel de metode se pot determina lucruri precum numarul de bolnavi de o anumita varsta sau boala din baza de date a unui spital. Relevanta acestor date variaza puternic, insa numarul este mai important decat numele pentru companii de asigurari spre exemplu, unde nu este de interes intrebarea “cine?” ci “cati?”.

Acest model de atac poate fi dat peste cap de diferite tehnici, precum: expandarea artificiala a volumelor cu date aleatoare, inserarea de inregistrari false, constrangerea intervalelor de cautare sau modificarea acestora.

Metoda prezentata in acest articol tinteste sa depaseasca d.p.d.v. al performantei algoritmi prezentati in trecut in mediul stiintific in acelasi domeniu, dar si sa demonstreze cum se pot depasi metodele de protectie prezentate anterior.