

Bezbedno traženje biomarkera korišćenjem hibridne homomorfne enkripcione šeme

Seminarski rad u okviru kursa
Kriptografija
Matematički fakultet

Stanković Una
una_stankovic@yahoo.com

16. maj 2017.

Sažetak

Sa sve bržim razvojem tehnologija za sekvenciranje genoma, raste i potreba da se očuva bezbednost podataka, posebno jer se oni, sve češće, čuvaju u oblaku i odatle koriste za istraživanje. U radu će biti predstavljen protokol za rešavanje problema bezbednog uparivanja enkriptovanih podataka. Biće predložen efikasan način da se bezbedno pretraži pozicija koja odgovara podacima iz upita i izvuku neke informacije sa te pozicije. Nakon dekriptovanja, postoji samo mala količina poređenja sa informacijama iz upita koja se vrši nad običnim tekstom. Ovaj metod će biti primenjen da se nađe skup biomarkera u enkriptovanim genomima.

Sadržaj

1 Uvod	2
2 Postavka problema	2
3 Zaključak	2
Literatura	2

1 Uvod

Brz razvoj tehnologije sekvenciranja genoma dozvoljava nam da pristupimo velikim skupovima genoma i što bi moglo da nam omogući bitne pomake u medicinskim istraživanjima. Informacije koje dobijamo iz skupova genoma se najčešće koriste u medicini, biomedicinskim istraživanjima, kao i uslugama koje se direktno pružaju korisnicima, zbog čega je veoma važno da se ovim podacima rukuje sa oprezom i da se oni obezbede od neovlašćenog pristupa i korišćenja.

Predloženo je da se privatnost podataka očuva korišćenjem homomorfne enkripcije (engl. Homomorphic Encryption - HE), koja omogućava da izračunavanja budu prenesena u obliku šifrovanog teksta. Homomorfna enkripcija je vid enkripcije koji nam dozvoljava da izvodimo operacije nad enkriptovanim tekstom i dobijemo enkriptovani rezultat, koji kada dekriptujemo, je isti kao rezultat koji bismo dobili da smo operaciju primenili nad neenkriptovanim tekstom.

Jasuda i drugi autori (engl. Yasuda et al.) su dali praktično rešenje za pronalazak lokacije uzorka u tekstu izračunavanjem više Hamingovih rastojanja nad enkriptovanim podacima. Loter i drugi (engl. Lauter et al.) su dali rešenje kako da se bezbedno izvrše osnovni genomske algoritmi korišćeni u mnogim sprovedenim studijama.

Homomorfna enkripcija može biti primenjena za očuvanje privatnosti sekvence koju analiziramo, ali se pokazuje kao veoma nepraktična za analizu informacija kod kompletnog ljudskog genoma. Glavna ideja, koja će biti korišćena za bezbedno pretraživanje skupa biomarkera korišćenjem Ring-GSW homomorfne enkripcione šeme, je da se enkodira baza genoma kao jedan element polinomijalnog prstena. Operacija pretrage u bazi se radi vršenjem jednog množenja sa genomom upita. Potom vršimo proceduru izvlačenja kako bismo dobili DNK sekvencu i nakon dekripcije je poredimo sa DNK upita u običnom tekstu.

2 Postavka problema

Zadatak je da se na bezbedan način izračuna verovatnoća genetskih bolesti kroz uparivanje skupa biomarkera sa enkriptovanim genomima koji se čuvaju u javnom "oblaku" (engl. cloud). Zahtev je da ceo proces uparivanja bude izvršen korišćenjem homomorfne enkripcije tako da se nikakve informacije o bazi i upitu ne otkriju serveru prilikom izračunavanja. Pretpostavimo da klijent ima VCF fajl (engl. Variation Call Format), koji sadrži informacije o genotipu, kao što su broj hromozoma i pozicija u genomu. Klijent enkriptuje informacije koristeći homomorfnu enkripciju i server računa tačno poklapanje nad enkriptovanim podacima. Ishod je prisustvo/odsustvo specifičnih biomarkera. Na kraju, klijent dekriptuje rezultat uz pomoć tajnog ključa homomorfne enkripcije.

**3 Pretraga i izvlačenje podataka iz baze
uz očuvanje privatnosti**

4 Zaključak