

Predstavenie možností sledovania majiteľov mobilných zariadení pomocou siete mobilných vysielateľov, zhodnotenie presnosti meraní, návrh možného využitia a spísanie zraniteľností riešenia.

Mgr. Marek Zeman, PhD., CRISC

V súčasnosti je stále ohýbaná téma sledovania občanov pomocou zamerania telefónu. Cieľom takéhoto sledovania je kontrolovanie občanov, či dodržiavajú karanténu po príchode zo zahraničia alebo po preukázaní ochorenia na vírus COVID-19. Druhým pertraktovaným cieľom je sledovanie, s kým sa pacient stretol od nakazenia po zahájenie karantény. Tento článok sa snaží o vytvorenie uceleného pohľadu a analýzu, aké dáta je možné získať, akým spôsobom sa dajú použiť a predstavíme si aj možné potenciálne chyby vyhodnotenia a spracovania dát.

Štát má za cieľ zaistiť údaje na identifikáciu chorých a potenciálne nakazených *„Toto opatrenie robíme s cieľom zbierať údaje na to, aby sme dokázali identifikovať chorých a ochrániť potenciálne nakazených a touto úpravou zavádzame špecifikáciu, na aký účel ideme využívať lokalizačné dáta pri mimoriadnej situácii v zdravotníctve.“* povedala vicepremiérka Veronika Remišová [1].

Následne sa predchádzajúce vyjadrenie pretavilo do dôvodovej správy ktorá bola predložená pri predkladaní zákona nasledovne: *„zavedenie možnosti využívania tzv. lokalizačných a prevádzkových údajov vznikajúcich v rámci elektronickej komunikácie zo strany Úradu verejného zdravotníctva na účely ochrany života a zdravia obyvateľov“* [5].

Z predchádzajúcich riadkov sa vyprofiloval zákon č.62/2020 v nasledovnom znení:

Zákon č. 62/2020 Z. z.Čl. IV

Zákon č. 351/2011 Z. z. o elektronických komunikáciách v znení zákona č. 547/2011 Z. z., zákona č. 241/2012 Z. z., zákona č. 352/2013 Z. z., zákona č. 402/2013 Z. z, nálezu Ústavného súdu Slovenskej republiky č. 139/2015 Z. z., zákona č. 247/2015 Z. z., zákona č. 269/2015 Z. z., zákona č. 391/2015 Z. z., zákona č. 397/2015 Z. z., zákona č. 444/2015 Z. z., zákona č. 125/2016 Z. z., zákona č. 353/2016 Z. z., zákona č. 386/2016 Z. z., zákona č. 238/2017 Z. z., zákon č. 243/2017 Z. z., zákona č. 319/2017 Z. z., zákona č. 56/2018 Z. z., zákona č. 69/2018 Z. z., zákona č. 177/2018 Z. z., zákona č. 30/2019 Z. z., zákona č. 94/2019 Z. z. a zákona č. 211/2019 sa dopĺňa takto:

§ 63 sa dopĺňa odsekmi 18 až 20, ktoré znejú:

„(18) Údaje, ktoré sú predmetom telekomunikačného tajomstva podľa odseku 1 písm. b) a d) spolu s informáciou o čase vzniku lokalizačného údaju podnik v čase mimoriadnej situácie alebo núdzového stavu v zdravotníctve, a to v príčinnej súvislosti so vznikom pandémie alebo šírením nebezpečnej nákazlivej ľudskej choroby,

- a) spracúva v anonymizovanej podobe na štatistické účely potrebné na predchádzanie, prevenciu a modelovanie vývoja ohrozenia života a zdravia,

- b) spracúva na účel identifikácie príjemcov správ, ktorým je potrebné oznámiť osobitné opatrenia Úradu verejného zdravotníctva Slovenskej republiky v záujme ochrany života a zdravia,
- c) spracúva výlučne v rozsahu potrebnom na identifikáciu užívateľov v záujme ochrany života a zdravia.

(19) Údaje spracúvané podľa odseku 18 podnik poskytuje Úradu verejného zdravotníctva Slovenskej republiky na základe odôvodnenej písomnej žiadosti. Žiadosť podľa predchádzajúcej vety musí obsahovať identifikáciu účastníka alebo užívateľa, alebo spôsob určenia identifikácie účastníkov alebo užívateľov, na ktorých sa poskytnutie týchto údajov vzťahuje.

(20) Úrad verejného zdravotníctva Slovenskej republiky môže údaje spracúvané podľa odseku 18 zbierať, spracúvať a uchovávať počas trvania mimoriadnej situácie alebo núdzového stavu v zdravotníctve, najdlhšie do 31. decembra 2020.“.

V našom článku budeme vychádzať z predchádzajúcich riadkov, t.j. nie je našou úlohou analyzovať, prečo zákon vznikol a zákonnosť spracovania údajov, ale či je možné údaje od operátorov použiť na spracovávanie a akej kvality dáta a softvér, ktorý ich spracováva musia byť. Analýzu si rozdelíme na tri časti:

1. Presnosť dát, ktoré mobilní operátori môžu získať a spracovávať. Chyby merania a nepresnosti z pohľadu technických parametrov.
2. Technologické obmedzenia a popis anomálií.
3. Ladenie modelu na analýzu dát a vyhodnocovanie pohybu a stretávania s potenciálnymi klientmi.

Ako funguje mobilná sieť

Presnosť lokalizácie mobilného telefónu je závislá od spôsobu, akým sú dáta zbierané a vyhodnocované a od prostredia, v ktorom sa telefón nachádza. Mobilný telefón môže komunikovať niekoľkými spôsobmi, či už pomocou WiFi, bluetooth alebo najčastejšie pomocou SIM karty s komunikačnými jednotkami operátora: Base Transmitting Station (BTS). BTS sieť čo je sieť vysielačov, ktoré sú umiestnené na území, v mestskom prostredí hustejšie. BTS zariadenia je možné použiť nielen na vysielanie a prijímanie hovorov, ale aj na dodatočnú analýzu signálu alebo prijímanie a vysielanie pridruženého signálu v čase, keď je telefón zapnutý. Telefón pri zapnutí a pripojení do siete operátora vysiela kód IMSI a IMEI, takže zariadenie aj klienta je možné počas pripojenia identifikovať. Naša analýza sa bude zaoberať len telefónmi s funkčnou a zapojenou SIM kartou, ostatné pripojenie nie je možné identifikovať.

V zásade sa dá rozdeliť prístup merania na štyri časti.

Prvou časťou je meranie pomocou trigonometrie a trilaterácie v rámci signálu vysielaného stanicami BTS. Trigonometria má svoje nepresnosti, používa sa zameranie pomocou troch BTS antén, každá anténa zameriava plochu okolo seba, t.j. 360 stupňov, do vzdialenosti maximálne 30km. Použitím troch BTS antén je možné zúžiť priestor na určitú oblasť, ktorá je však veľmi veľká. Existujú spôsoby na spresniť zamerania oblasti, či to je smerovaním antén iba na určitú oblasť alebo použitím technológie Timing Advance, ktorá meria časť, kým sa vráti signál, využíva teda fyzikálne vlastnosti rýchlosti prenosu signálu. Timing Advance má chybu okolo ½ km. Druhou možnosťou je využiť ďalšie spôsoby zameriavania s vyššou presnosťou, ako je

napríklad Round Trip time. Táto metóda zvyšuje presnosť teoreticky na 5 metrov, ale hodnota je teoretická a extrémne obmedzená na ideálne prostredie, v reálnom svete dostávame presnosť rádovo v niekoľkých desiatkach metrov. Tieto metódy si vyžadujú vyšší výkon a teda aj investície na strane operátora a aj v jednotlivých BTS, zároveň v zastavaných oblastiach je znížená presnosť riešení vzhľadom na tienenie žiarenia vysielačov budovami. Všetky metódy sú závislé aj od rýchlosti odozvy mobilného zariadenia. Výsledkom je, že na presné zameranie sa používajú matematické modely a logické obštrukcie. Príkladom logickej obštrukcie môže byť: ak je nameraná rýchlosť presunu zariadenia rýchlosťou 50km/h a podarí sa objekt umiestniť na cestu, je predpoklad, že je v aute a bude pokračovať jazdou po ceste, preto nasledovné umiestnenie bude sledovať cestu a zjednoduší sa zameriavanie. Vyššia presnosť predchádzajúcich zariadení sa dá dosiahnuť inštaláciou špeciálneho softvéru pre podporu jednotlivých technológií operátorom. Faktom je, že presnosťou merania v zásade sa zaoberá jediná norma a to je *Wireless E911 Location Accuracy Requirements*, ktorú vydala Federálna komisia pre komunikáciu 03. apríla 2015 [4]. Norma vyžaduje presnosť 50m pri 67% spoľahlivosti. Norma zároveň pamätá na zlepšovanie riešenia, preto sa počíta so zvýšením spoľahlivosti na 80% do piatich rokov od implementácie.

Predposlednou možnosťou je využitie GPS antény v zariadení, poprípade skombinovanie merania GPS antény s riešeniami z BTS siete. Poslednou možnosťou, ktorá je použiteľná v zastavaných oblastiach je využitie pripojenia do WiFi sietí a zamerania IP adresy. Tieto merania sú presnejšie, ale vyžadujú spustenie zo strany používateľa, používateľ by musel mať nainštalovanú aplikáciu pre mobilného operátora a aplikácia musí mať prístup ku týmto dátam.

Chyby merania a popis anomálií

Primárne sa zameriame na použitie zamerania pomocou BTS. Využitím GPS by sme získali veľkú presnosť, ale vlastník mobilného zariadenia musí súhlasiť s využitím a sprostredkovaním GPS dát pre operátora pomocou aplikácie. S pomocou GPS zariadenia v meste je presnosť v rádovo niekoľko metrov, často desiatok, podľa počtu viditeľných satelitov, ktoré sú zakryté budovami, pri pohybe objektu sú zakryté náhodne. V priestore, bez zakrývajúcich elementov je presnosť rádovo v metroch. Rovnakú presnosť vieme dosiahnuť zameraním WiFi vysielača alebo bluetooth vysielača.

Zameriavanie BTS má niekoľko nepresností.

- Zameranie je závislé od trigonometrie a trilaterácie, v mimomestských lokalitách je presnosť rádovo v stovkách metrov z dôvodu hustoty BTS. V mestských lokalitách je presnosť stále v desiatkach metrov, z dôvodu nehomogénneho signálu, odrazivosti a pravdepodobnosti výskytu rušivých prvkov/elementov.
- Často je zameraná skôr predpokladaná vzdialenosť ako smer, ak je zariadenie v dosahu jedného vysielača BTS:
- Presnosť ovplyvňujú prvky exteriéru ako sú veľké budovy, kopce. Pokiaľ nie je anténa BTS smerová, nie je možné identifikovať, či mobilné zariadenie zachádza za kopec, alebo vchádza do rokliny.
- Presnosť ovplyvňujú aj lokálne, časovo premenlivé, klimatické podmienky

- Každé nové meranie môže vyvolať zdanie pohybu.
- Meranie a presnosť BTS ovplyvňuje samotné vytáženie BTS. Pri veľkom vytážení sa presnosť merania znižuje v prospech využitia antény zákazníkmi na primárne funkcie.

Dopadom týchto nepresností je, že v prostredí hraníc okresov a štátov môže mobilné zariadenie prechádzať medzi jednotlivými krajinami. Napríklad v okolí rieky Moravy sa mobilné zariadenia prepája na BTS v Čechách alebo v Rakúsku. Na hrádzi povedľa Dunaja z Bratislavy do Čunova môže zariadenie prejsť tromi štátmi a v rámci Slovenska dvomi okresmi, pričom zariadenie neopustilo jeden štát a okres. Napr. pri výstupe na vrch Tríbeč zo strany okresu Zlatých Moraviec sa pred samotným vrcholom, mobilné zariadenie prepne na BTS anténu do okresu Topoľčany, hranica okresu prechádza samotným vrcholom. Pre nezávislého pozorovateľa sa toto prepnutie bude javiť ako prechod cez hranicu okresov.

Druhou pravidelnou chybou je, že zariadenie uložené v byte sa pri každom premeraní posunie, až sa dostane mimo bytu. Dokonca sa stáva, že zariadenie je zameriavané už len mimo bytu. V tom čase nie je možné presne zamerať zariadenie v budove.

Tretím dopadom je zvýšenie nepresnosti pri vytážení.

Štvrtým problémom je analýza priradenia mobilného zariadenia jednotlivým ľuďom. Pri priamom predaji je predpoklad, že operátor bude vedieť meno osoby, ktorej bolo mobilné zariadenie predané. Pri sprostredkovanom predaji, ako je odovzdanie mobilného telefónu firmou zamestnancovi, operátor nepozná mená jednotlivých klientov. Tieto osoby sa môžu pravidelne meniť. To isté platí ak sú telefóny centrálné nakupované jednou osobou pre domácnosť s cieľom maximálneho využitia zliav.

Celkovým dopadom je, že cieľ zákona spracúvať „*výlučne v rozsahu potrebnom na identifikáciu užívateľov v záujme ochrany života a zdravia.*“ je veľmi ťažko naplniteľné. Dôvodom je či už nepresnosť meraní, to znamená, že sa nedá s istotou identifikovať, či klient stretol chorého alebo nie. A zároveň je veľmi ťažké identifikovať, či klient dodržiava karanténu, klient na jednej strane nemusí byť známy mobilným operátorom a na druhej strane nemusí mobilné zariadenie so SIM kartou so sebou nosiť.

Ladenie modelu analýzy dát

V predchádzajúcej časti sme ukázali, že presné zameranie mobilu nie je možné. Naznačili sme, že je možné predikovať správanie držiteľa mobilného zariadenia na základe pridania a spracovania dodatočných informácií. Dáta od operátorov je nevyhnutné obohatiť o dodatočné informácie, ako je poloha budov, chodníkov, ciest, informácií o správaní držiteľov mobilov (ako sa správajú, pohybujú v bytoch, na prechádzke, v aute,...). Správanie v tomto prípade chápeme v hraniciach obvyklého správania anonymných osôb v priestore. Ak sú tieto údaje spracované pomocou následného programového vybavenia, ktoré zoberie do úvahy priestorové závislosti a správanie, môže byť pohyb klienta zásadne spresnený. To neznamena, že výsledok je bez štatistickej chyby, chyba bude však menšia. Dáta získané od mobilných operátorov patria do skupiny Big data a treba k nim pristupovať programovým vybavením a matematickými a štatistickými modelmi vhodnými pre túto skupinu dát. Modelov na spracovanie veľkých dát a vyselektovanie toho

správneho je práca pre matematika a odborníka, ktorý pozná obsah údajov, na niekoľko mesiacov. V tomto prípade nie je ľahké rozlíšiť koreláciu dát od príčinnej súvislosti. Výhodou je, že v rámci Slovenska existuje niekoľko firiem, ktoré sa zameriavajú na analýzu mobilných dát a majú k dispozícii dostatočné vedomosti aj skúsenosti.

Nie je možné bez využitia Big data prístupu spracovávať a dobre vyhodnotiť dáta z mobilných zariadení. Riešenia typu Big data pracujú s chybou matematického modelu, ktorá je závislá od výberu modelu, jeho presnosti a aj od nasadenia. Hlavným ohrozením v tomto prípade je obava rýchleho nasadenia zle vyladeného modelu, ktorý môže šikanovať pacientov a potenciálnych ľudí ohrozených vírom. Model spracovania dát so štatistickou chybou je použiteľný pri riadení firmy, pretože firma vie prijať a spracovať z chyby modelu vyplývajúce riziko, ťažšie sa bude presadzovať pri určovaní identifikácie neštandardného správania chorých pacientov.

O stave, keď matematický model ohrozuje chudobných, posilňuje rasizmus, či znásobuje nerovnosť píše matematicka *Cathy O'Neil* v knihe *Weapons of math destruction*. Tvrdí, že modely sú nejasné, to znamená, že nevieme, z čoho modely vychádzajú a akým spôsobom dáta spracovávajú, lebo je to predmetom obchodného tajomstva. Modely sú neregulované, to znamená, že môžu získavať a spracovávať viac dát ako je pôvodným účelom modelu a keďže neexistuje spoločný auditovací model, nie je možné bez hĺbkového auditu identifikovať dodržanie rozsahu a účelu spracovania. Výsledky modelov nie je možné potvrdiť a tým pádom predkladajú vlastné predsudky, pričom ovplyvňujú veľkú skupinu obyvateľstva. [3].

Zhrnutie

Ukázali sme, že dáta sú vhodné pri zaregistrovaní prichádzajúceho človeka zo vzdialenej cudziny na Slovensko. Sú vhodné aj na zaregistrovanie zásadného prechodu správnym celkom, či už ide o prechod medzi viacerými štátmi alebo medzi okresmi v rámci štátu. V našej analýze sme poukázali na nepresnosti pri presnom zameriavaní mobilných zariadení. Tieto nepresnosti idú až do úrovne, že objekt sa môže javiť v inom správnom územnom celku ak je na hranici alebo sa môže javiť vzdialený od domova, pričom domov neopustil. Výsledkom týchto informácií je, že nie je možné presné zameranie a jasné definovanie správania jednotlivých zariadení. To znamená, že nie je možné identifikovať, či sa majiteľ nepohyboval a bol celý deň na jednom, definovanom mieste a ak sa pohyboval, nevieme zdefinovať, po akej trajektórii to bolo. Takže nie je možné určiť ani s kým sa stretol a na akú vzdialenosť.

Predstavili sme, že jediná možnosť, ako spracovávať a vyhodnocovať dáta od mobilných operátorov je pomocou Big Data modelov. Modely typu big dáta sú matematickej podstaty a pracujú so štatistickou chybou. Veľkou neznámou je v súčasnosti pre verejnosť model, ktorým sa budú vyhodnocovať dáta od mobilných operátorov. Optimálne riešenie je model, jeho metódy, prístupy a nastavenia a uloženie a zabezpečenie dát verejne predstaviť, aby boli známe a spoločnosťou kontrolovateľné. Takýmto prístupom sa snažíme predísť nedorozumeniam a zvyšovaní nedôvery spoločnosti voči výsledku spracovaných dát ako aj spracovaniu dát. Nový prístup sa očakáva aj od spoločnosti, pretože musí akceptovať vety typu: „*S pravdepodobnosťou 78% ste porušili zákonom stanovenú karanténu, za ktorú je väzenie alebo pokuta.*“ To znamená,

že je potrebné vybudovať nový typ zmýšľania v spoločnosti a zároveň aj nové ochranné prvky jedincov. Príkladom takýchto prvkov sú špeciálne poistenia alebo vytvorenie prostredia pre jedincov, pre zabezpečenie a sústavné zbieranie dôkazového materiálu, pre prípad pochybenia modelu. Je možné využiť a štátom definovať niektorú zo spoločností so zameraním na spracovanie Big data modelov od operátorov a takejto spoločnosti prideliť úlohu spracovania, čím sa skráti ladenie modelu. Otázkou ostáva, do akej miery je súkromná spoločnosť schopná prezradiť parametre svojho modelu spracovania a do akej miery vie takáto spoločnosť bezpečne uložiť tento typ dokumentov v chránenom úložisku. Úvahy o spracovaní a uložení dát sú nad rámec dokumentu.

V súčasnosti sa najjednoduchším riešením ako obísť zameranie mobilného telefónu stále ukazuje ponechanie mobilného zariadenia doma, alebo použiť cudzí telefón so SIM alebo využiť iba voľné WiFi pripojenie a štandardné programy na chat a nemať v telefóne SIM. Použitie telekomunikačných dát podľa zákona je nejasné, pričom presnosť dát a ich spracovanie je veľmi volatilné. Samotná spoločnosť a jej politické špičky neprispievajú ku dôvere v spracovanie a následné použitie dát. Samotná myšlienka využitia dát zo zariadení, ktoré majú obyvatelia stále pri sebe pre prípad nákazy je vo svojej podstate dobrá. V rámci ladenia modelu budú nedostatky dát rýchlo objavené, hlavným problémom sa javí malá identifikovateľnosť klientov, pretože vysoké percento SIM kariet nie je zosobniteľné s konkrétnym človekom. Predpokladom je, že ďalším krokom bude rozšírenie zákona o nové zdroje dát na presnú identifikáciu klientov. Do úvahy pripadajú dáta zozbierané za iným účelom, štátom (napríklad pri kontrole prechodu medzi okresmi v rámci kontroly karantény) alebo dáta zozbierané a využívané súkromnými firmami (v rámci kontaktov pacientov v nemocniciach alebo zdroj dát môže byť zoznam klientov finančných inštitúcií alebo to môžu byť dáta klientov veľkých obchodných reťazcov).

Úlohou do budúcnosti bude podrobne sledovať tok dát, spracovanie dát, ich využitie správnym spôsobom v kontrolovanom prostredí a zmazanie dát po využití.

Literatúra

[1], <https://www.vicepremier.gov.sk/aktuality/podpredsednicka-vlady/lex-covid-data-od-operatorov-budu-sluzit-len-v-ramci-boja-proti-sireniu-pandemie/> zo dňa 12.04.2020

[2], <https://zive.aktuality.sk/clanok/74824/odborny-clanok-na-temu-lokalizacia/> zo dňa 12.04.2020

[3]. O'Neil, C (2016): Weapons of math destruction; Penguin books

[4] <https://www.federalregister.gov/documents/2015/03/04/2015-04424/wireless-e911-location-accuracy-requirements> zo dňa 14.04.2020

[5], <https://www.nrsr.sk/web/Dynamic/DocumentPreview.aspx?DocID=476589> zo dňa 14.04.2020