# Štátna aplikácia proti koronavírusu je katastrofa, odosiela kontakty všetkých, sú v nej veľké chyby

- CZ24 News | 28. dubna 2020

SLOVENSKO: Neakceptovateľný je samotný principiálny spôsob fungovania aplikácie, píše technologický server DSL.

Štátom ponúkaná, oficiálne prevádzkovaná a propagovaná aplikácia pre boj s koronavírusom Covid19 ZostanZdravy, ktorej hlavným účelom je zatiaľ pomôcť vyhľadať predchádzajúce kontakty detekovaných infikovaných, má podľa analýzy servera DSL.sk závažné nedostatky. A to v oblasti ochrany súkromia aj v technickom prevedení.

Aplikácia bola vyvinutá v rámci dobrovoľnej aktivity, štát ju ale prevzal, oficiálne prevádzkuje a poskytuje občanom okrem iného aj propagovaním na štátnom webe korona.gov.sk. Oficiálnym uvádzaným prevádzkovateľom je Úrad verejného zdravotníctva, pričom v aplikácii sa spomína aj NCZI, Národné centrum zdravotníckych informácií, a cez jeho účet je aj aplikácií sprístupnená v obchodoch mobilných operátorov.

**Ochrana súkromia**

Neakceptovateľný je samotný principiálny spôsob fungovania aplikácie. Podobne ako iné iniciatívy je jej cieľom identifikovať cez Bluetooth iných užívateľov aplikácie, keď aplikácia do okolia cez Bluetooth Low Energy neustále vysiela identifikátor profilu užívateľa.

Pri bežnej činnosti neustále prijímaním takýchto správ od iných telefónov identifikuje kontakty s inými užívateľmi dlhšie ako nastavený limit, v súčasnosti sa používa 5 minút.

Následne keď bude niektorý užívateľ aplikácie testovaný ako pozitívny na nový koronavírus, aplikácia môže identifikovať kontakty infikovaného v predchádzajúcich dňoch vystavené riziku nákazy a tieto môžu byť upozornené napríklad notifikáciou.

Slovenská aplikácia ale tento cieľ dosahuje neakceptovateľným spôsobom, keď automaticky odosiela štátu všetky identifikované kontakty všetkých užívateľov aplikácie dlhšie ako aktuálne 5 minút. Odosielané je bežne ID profilu samotného užívateľa, ID profilu detekovaného kontaktu, presný čas začiatku kontaktu a presná dĺžka kontaktu. ID profilu je jedinečné číslo prideľované štátom pri registrácii aplikácie na serveroch, po jej nainštalovaní na zariadení užívateľa.

Hoci samotné priamo odosielané jednotlivé informácie sú anonymné, to neznamená že nemôžu byť deanonymizované. Naopak je zjavné, že tieto informácie pomocou ďalších informácií umožňujú identifikovať minimálne časť jednotlivých užívateľov a následne ich kontakty.

Hlavne ale samozrejme je takéto fungovanie neprimerané vzhľadom na daný účel, ktorý sa dá dosiahnuť bez takéhoto potenciálneho zásahu do súkromia. Na odosielanie všetkých kontaktov všetkých užívateľov nie je dôvod a prirodzeným riešením samozrejme je odoslanie kontaktov len infikovaných osôb po tom, ako boli detekované ako pozitívne.

Tak fungujú podľa dostupných informácií viaceré fungujúce a pripravované riešenia v civilizovaných krajinách západnej demokracie, okrem iného česká eRouška, návrhy iniciatív DP-3T, PEPP-PT,

pripravované riešenia Google a Apple. Tieto riešenia navyše tam nekončia a riešia aj ďalšie dôležité otázky ochrany súkromia eliminujúce možné spôsoby deanonymizovania, napríklad nevysielajú vždy ten istý identifikátor.

Už kvôli tomuto hlavnému princípu fungovania je súčasná podoba aplikácie Covid19 ZostanZdravy neakceptovateľná.

Navyše v informáciách v aplikácii a mobilných obchodoch sa zrejme explicitne neupozorňuje, že aplikácia odosiela všetky kontakty na server. Štát o tom nepochybne ale vie, keď to aj keď nie celkom jasne uvádza na webe korona.gov.sk formuláciou “anonymne zapíše obidve zariadenia na server”.

Zbierať všetky kontakty ľudí s ich dĺžkou a časom môže byť samozrejme potenciálne zaujímavé na nejaký vedecký výskum typu kontaktov v populácii použiteľný aj pri výskume šírenia nákazy, to sa na Slovensku evidentne nedeje a hlavne takýmto prípadným výskumom nie je potrebné narúšať súkromie všetkých užívateľov a realizovať sa má iba na vzorke populácie.

**Deanonymizovanie**

Analyzovať všetky možnosti deanonymizovania dát odosielaných aplikáciou štátu je mimo rozsahu tohto článku, ich neprimeranosť je zjavná. Uvádzame len niekoľko príkladov.

Pri zvažovaní primeranosti nie je samozrejme potrebné uvažovať čo sa s dátami naozaj bude robiť ale čo by sa s nimi dalo robiť.

Viaceré osoby sú deanonymizované priamo informáciami, ktoré poskytujú štátu respektíve hygienikom zo samotného princípu fungovania aplikácie. Infikovaní majú hygienikom priamo poskytnuť svoj identifikátor v aplikácii, aby mohli byť upozornené kontakty minimálne notifikáciou. Ľudia v karanténe využívajúci túto aplikáciu v režime karantény sú zase dostatočne identifikovaní minimálne mobilným telefónnym číslom.

Ďalšie informácie, ktoré infikovaná osoba reálne poskytne hygienikom, už priamo umožňujú deanonymizovať ďalších ľudí. Pri štandardnom šetrení infikovaná osoba spomína na dlhšie stretnutia so známymi osobami, pričom takéto informácie sú typicky doplnené aj časom stretnutia. To je špecificky účelom napríklad pripravovaného riešenia, keď si bude užívateľ spomínať na kontakty podľa histórie polohy v čase od mobilných operátorov. Keďže v kontaktoch odoslaných aplikáciou je aj čas stretnutia, tieto informácie spolu identifikujú aj profily aké majú identifikované osoby v kontakte s infikovaným v aplikácii.

Podobne samotné informácie odosielané aplikáciou umožňujú zrejme často dobre deanonymizovať ďalšie osoby. Napríklad kontakty už identifikovanej osoby, ktoré sú pravidelne celú noc v blízkosti tohto kontaktu, sú zrejme osoby bývajúce s ňou alebo v jej blízkosti, atď.

Dáta sa dajú tiež deanonymizovať samozrejme aj skombinovaním s inými ďalšími dátami.

Následne tak tieto informácie poskytnuté štátu identifikujú aj prípadné stretnutia dvoch vyššími spôsobmi identifikovaných ľudí, z ktorých ani jeden nebol infikovaný a na identifikovanie ich stretnutia nie je žiadny dôvod.

**Dokonca aj odosielanie polohy**

Pri analýze sme sa zamerali na aplikáciu pre mobilnú platformu Android, ktorá je výrazne masovejšia ako iOS. Analyzovaná bola verzia 1.0.2, ktorá bola vydaná 17. apríla a doteraz je stále poslednou verziou.

Okrem neakceptovateľného princípu centrálneho zbierania všetkých kontaktov sa v aplikácii nachádza ďalšie problematické štandardné nastavenie.

Aplikácia využíva konfigurovanie cez službu Firebase Remote Config od Google a v aktuálnej konfigurácii vracanej zo služby sa poloha miesta kontaktu vo všeobecnosti neodosiela.

Podľa štandardných zabudovaných nastavení priamo v aplikácii pre prípad, že by sa nepodarilo stiahnuť konfiguráciu, by sa spolu s kontaktami odosielala ale dokonca aj poloha tohto stretnutia. Konkrétne s GPS súradnicami zaokrúhlenými na tri desatinné miesta, čo znamená priemerné zníženie presnosti v horizontálnom smere len o menej ako 40 metrov a vo vertikálnom o menej ako 60 metrov.

K použitiu štandardnej konfigurácie môže prísť, ak sa z nejakých dôvodov nepodarí stiahnuť konfiguráciu z Firebase Remote Config. Zrejme je to maximálne výnimočná situácia, reálne ale môže nastať. Takáto základná konfigurácia je tak nevhodná.

Vzniká tiež otázka, či takáto konfigurácia nebola v minulosti aj reálne nastavovaná z Firebase Remote Config a prípadne dokedy. Naznačujú to aj informácie na štátnom webe, ktoré hovoria "aplikácia zaznamená približnú GPS polohu" kontaktu, čo sa pri súčasnom nastavení už nedeje.

Tvorcovia pre DSL.sk potvrdili, že sa tak reálne na začiatku po sprístupnení aplikácie najskôr dialo a aj vo vracanej konfigurácii z Firebase Remote Config bola hodnota znamenajúca odosielanie aj polohy všetkých kontaktov so súradnicami zaokrúhlenými na tri desatinné miesta. To bolo podľa nich pôvodne zamýšlané na "dezinfekciu miesta". Takéto nastavenie malo byť v platnosti podľa nich ale iba približne 8 hodín, malo sa dotknúť len pár užívateľov a všetky pozície boli zo serverov okamžite odstránené.

Ďalej poloha ľudí v karanténe využívajúcich aplikáciu v tomto režime sa podľa nastavení priebežne neodosiela štátu, podľa zdrojových kódov sa zrejme ale odosiela presná poloha všetkých ich kontaktov s inými užívateľmi. Tento scenár sme ale netestovali. O spôsobe monitorovania polohy ľudí v karanténe sa dá diskutovať, odosielanie polohy kontaktov s inými osobami ale sleduje aj polohu týchto iných osôb pri týchto kontaktoch. Pritom sa nemuseli nijako previniť a osobu v karanténe ani nemusia poznať, len sa nachádza v ich blízkosti.

Tvorcovia fakticky pre DSL.sk potvrdili, že poloha sa v tomto prípade naozaj odosiela a aktuálne bola do Google Play poslaná aktualizácia, ktorá tak už robiť nebude.

**Vážna chyba**

Navyše v aplikácii pre Android je vážna programátorská chyba, ktorá znehodnocuje zbierané dáta.

Aplikácia cez Bluetooth LE vysiela číslo profilu užívateľa ako štyri bajty vo vysielanej správe. Momentálne sa najvyššie číslo profilu pohybuje nad 25 tisíc. Aplikácia pre Android ho spätne z týchto bajtov počíta chybným spôsobom, pre ktorý pre približne polovicu čísiel profilov vypočíta zlé číslo profilu, navyše pri tom stratí informáciu a z vypočítaného čísla sa skutočné číslo profilu kontaktu nedá priamo vypočítať.

Konkrétne používa výpočet content[16] << 24 | content[17] << 16 | content[18] << 8 | content[19], pričom content je pole bajtov správy. V Jave je ale byte číslo so znamienkom a pri bitovom or sa najskôr zrejme operandy skonvertujú na int. V prípade, že posledný bajt čísla má hodnotu od 128 do 255 a v Jave je teda záporný, jednotky vo všetkých bitoch vyšších bajtov príslušnej int hodnoty spôsobia že celý výsledok bude mať vo vyšších bajtoch samé jednotky a výsledkom bude teda iba tento posledný bajt a to ako záporné číslo.

Napríklad pre číslo profilu kontaktu 5000 v skutočnosti Android aplikácia vypočíta z prijatej Bluetooth LE správy jeho hodnotu ako -120, keďže posledný bajt 5000 je 136 znamenajúci pri reprezentácii bajtu so znamienkom -120.

Približne polovica kontaktov registrovaných Android aplikáciou tak má zlé číslo identifikovaného kontaktu, prišlo k strate vyššieho bajtu a z uložených a následne odoslaných informácií sa nedá priamo vypočítať správna hodnota. Ak vzájomný kontakt zachytil a odoslal aj druhý telefón, čo sa nemusí diať vždy, v takýchto prípadoch sa budú dať informácie o identifikovaných kontaktoch často ale nie vždy s určitosťou jednoznačne opraviť.

Aplikácia odosiela informácie o kontaktoch na servery raz za hodinu, pričom posiela naraz informácie o všetkých ešte neodoslaných kontaktoch. Takéto odosielanie ale skončí v súčasnosti neúspechom, ak obsahuje informácie o aspoň jednom kontakte kvôli tejto chybe so záporným číslom profilu. Server vracia v takom prípade chybu 500, aplikácia považuje odosielanie za neúspešné a všetky kontakty sa pokúsi odoslať neskôr znovu. Chybné dáta sa podľa tvorcov napriek chybovému kódu 500 reálne na server dostávali.

Tvorcovia aplikácie chybu identifikovali skôr ako sme im chybu a jej závažnosť v nedeľu nahlásili, nie je ale jasné kedy ju identifikovali a najmä či doteraz plne chápali jej vážnosť. Chybu v zdrojových kódoch aj na GitHube opravili ešte 21. apríla.

Chybu na GitHube ale označujú len ako “potenciálne pretečenie”, opravená aplikácia nebola vydaná ani v čase kontroly v nedeľu 26. apríla, stále boli prideľované aj čísla profilov spôsobujúce problém v doterajších verziách Android aplikácie a server aj v nedeľu vracal chybu 500 pre odosielané kontakty so zápornými číslami profilov.

Okamžite ako bola chyba zanalyzovaná bola podľa tvorcov opravená a poslaná do Google Play, kde čaká na schválenie. Kedy sa tak ale stalo odmietli uviesť.

**Stanovisko tvorcov**

Napriek použitiu štátom je aplikácia stále a nepochopiteľne tvorcami naďalej vyvíjaná na dobrovoľnej báze a bezplatne. Tvorcovia sú ale radi, že môžu pomôcť štátu v tejto dobe, a vyzývajú aj ďalších vývojárov, aby pomohli zlepšiť tento projekt.

Podľa tvorcov v čase vzniku prvej verzie aplikácie kládli dôraz na ochranu obyvateľstva, teraz plánujú zlepšiť ochranu súkromia. Zvažujú zrejme prejsť na vyššie spomínaný spôsob, kedy budú so štátom zdieľané kontakty užívateľa len v prípade ak bude užívateľ infikovaný. Kedy by mohla byť k dispozícii nie je jasné.

Verzia s opravenou chybou pre Android má byť sprístupnená, keď Google Play sprístupní už poslanú aktualizáciu. Či sa opravia kontakty pokazené chybou v Android aplikácii neuviedli.