Tietoliikenne, tietoturva sekä datan keräämiseen ja keskittymiseen liittyvät isot kysymykset

Tietoliikenteen toimintaa oppii ymmärtämään tutustumalla palvelimien toimintaan, palvelinohjelmiin, verkkoselaimien toimintaan, Internetin fyysiseen infrastruktuuriin, protokolliin ja eri tapoihin, joilla tietoa voidaan siirtää laitteesta toiseen. Mitä tahansa elektronisella laitteella olevaa tietoa voidaan siirtää muun muassa äänellä, näkyvällä valolla, lämmöllä, muulla sähkömagneettisella säteilyllä, sähköjohtimilla tai mekaanisella värähtelyllä.

Internetin varassa toimii monenlaisia palveluita, kuten sosiaalisen median palveluita, pilvitallennuspalveluita, sähköpostipalveluita, pikaviestipalveluita ja videotoistopalveluita. Näitä palveluita käytetään verkkosivujen, mobiilisovellusten sekä muiden tietokoneohjelmien avulla. Toimiakseen ja säilyttääkseen käyttäjien dataa, nämä eri palvelut käyttävät useita erilaisia palvelimia. Palvelimet ovat verkkoon kytkettyjä tietokoneita, joissa on yksi tai useampi palvelinohjelma käynnissä.

Verkkopalveluiden, ohjelmien ja laitteiden käyttämiseen liittyy ainakin kuusi huomioitavaa seikkaa. Ensinnäkin, omaa nimeä, puhelinnumeroa tai osoitetta ei pidä tarpeettomasti kertoa verkossa kenellekään tai syöttää mihinkään palveluun.

Toiseksi, data on luonteeltaan sellaista, että se voi olla usein vaikeasti hävitettävissä. Vaikka joltakin laitteelta olisikin poistettu tiedostoja, niin on olemassa monia työkaluja, joilla poistettua dataa saa palautettua takaisin. Tästä syystä kannattaa esimerkiksi olla todella tarkkana, ennen kuin lähtee myymään vanhaa puhelintaan tai käytettyjä tallennusmedioitaan (kuten kiintolevyä). Käytetyistä puhelimista voi olla mahdollista palauttaa esimerkiksi edellisen omistajan puhelimessaan säilyttämiä kuvia, tallennettuja salasanoja, yhteystietoja ja tekstiviestikeskusteluja. Ellei ole vakuuttunut siitä, että on onnistunut tarkoitusta varten tehdyillä erikoistyökaluilla pyyhkimään nämä laitteet tyhjiksi, niin vanhat laitteet voi olla parasta jättää myymättä eteenpäin. Usein esimerkiksi verkkopalveluilla on useampia kopioita samasta

tiedosta hajautettuna eri laitteille, mikä edelleen vaikeuttaa tiedon hävittämistä. Eri palveluihin lähetettyyn dataan onkin turvallisinta suhtautua sillä tavalla, että kyseinen data ei tule koskaan katoamaan Internetistä. Johonkin palveluun lähetetty viesti tai kuva voi ilmaantua Internetin syövereistä vielä vuosikymmenien päästä.

Kolmanneksi, palveluntarjoajalla on lähtökohtaisesti tekniset edellytykset mielivaltaisesti lukea, muokata, poistaa ja kopioida mitä tahansa käyttäjän palveluun tallettamaa tai palvelun kautta lähettämää dataa. Tämä koskee esimerkiksi käyttäjän palveluun syöttämiä henkilötietoja, lataamia kuvia ja tiedostoja sekä palvelun kautta välitettyjä viestejä. Omien yksityisyyssuoja-asetuksien muuttamisella ei ole vaikutusta näihin teknisiin edellytyksiin. Lait tosin jossain määrin rajoittavat näiden tietojen käsittelyä palveluntarjoajien osalta. Tietomurron tai tietovuodon johdosta tämä data voi päätyä ei-toivotuille tahoille, tai siitä voi jopa tulla kaikille julkista tietoa.

Neljänneksi, useat eri palvelut, ohjelmat ja verkkosivut keräävät käyttäjistään dataa, kuten aktiivisuusdataa, sijaintidataa ja selaushistoriaa. Aktiivisuusdataa on esimerkiksi käyttäjän eri toimet palvelun tai ohjelman sisällä, kenelle käyttäjä on lähettänyt viestejä, kuinka useasti käyttäjä on lähettänyt viestejä hänen eri kontakteilleen ja mihin aikaan hän on niitä lähettänyt. Käyttäjien puolesta myös säilytetään ja välitetään runsaasti tietoa. Näistä esimerkkejä ovat pikaviestisovelluksilla lähetetyt viestit, sosiaalisen median palveluissa käydyt keskustelut, sähköpostiviestit ja pilvitallennuspalveluihin tallennetut tiedostot. Monesti välitettyjä ja vastaanotettuja viestejä myös säilytetään käyttäjien puolesta. Myös monet opetusteknologiatuotteet keräävät paljon tietoa.

Viidenneksi, verkkoselailu ei ole todellisuudessa kovinkaan yksityistä. Esimerkiksi Internet-yhteydentarjoajat saavat hyvinkin yksityiskohtaisia tietoja heidän asiakkaidensa verkkoselailusta: selatusta verkkosivusta riippuen, toisinaan palveluntarjoajien on mahdollista saada tietoa jopa asiakkaiden katsomista yksittäisistä kuvista tai artikkeleista, ja niihin liittyvistä katsomisajanhetkistä. Muut verkossa palveluita tarjoavat kaupalliset toimijat käyttävät erilaisia menetelmiä käyttäjien seuraamiseen verkossa. Myös voittoa tavoittelemattomat tahot voivat käyttää kolmannen osapuolen tarjoamia analytiikkapalveluita, joista esimerkkeinä Kela ja Terveyskirjasto, jotka molemmat käyttävät Google Analyticsiä verkkosivuillaan. Käyttämällä eri seurantamenetelmiä, on mahdollista hyvinkin tarkasti tallentaa tietoa yksittäisten käyttäjien verkkoselaamisesta. Seurantamenetelmien tukena ovat

muun muassa evästeet, seurantapikselit, käyttäjän IP-osoite, käyttäjän selailutottumukset ja selaimen/laitteen sormenjälki.

Kuudenneksi, laitteisiin asennetuista suljetun lähdekoodin ohjelmista ja käyttöjärjestelmistä on hankala tarkalleen tietää, mitä tietoa kyseiset ohjelmat ja käyttöjärjestelmät lähettävät. Joudumme tämän vuoksi usein luottamaan pelkästään ohjelmien tekijöiden sanaan siitä, mitä ohjelmat tarkalleen ottaen tekevät. Tämän lisäksi on tiedostettava, että käyttöjärjestelmien graafisten käyttöliittymien alla, on aina käynnissä monenlaisia eri prosesseja. Vaikka jokin tietoa lähettävä toiminto tai ohjelma näyttäisikin olevan kytkettynä pois graafisen käyttöliittymän perusteella, niin se ei välttämättä tarkoita sitä, etteikö kyseinen toiminto tai ohjelma olisi silti jollakin tavalla aktiivisena. Näistä syistä onkin aiheellista ottaa huomioon esimerkiksi tekijöiden toimintaan vaikuttavia taloudellisia kannustimia sekä tekijöihin kohdistuvia muita paineita, kun arvioimme eri ohjelmien luotettavuutta.

Esimerkkejä suljetun lähdekoodin käyttöjärjestelmistä ja ohjelmista ovat iOS, Windows, WhatsApp-sovellus, Instagram-sovellus, Snapchat-sovellus, Google Chrome -selain ja Google Play Store -sovellus. Android-käyttöjärjestelmän lähdekoodi on avoin, mutta käytännössä Android-puhelimet sisältävät useita suljetun lähdekoodin ohjelmia valmiiksi asennettuna.

Tällä hetkellä iso osa kaikesta kerätystä, välitetystä ja säilytettävästä datasta keskittyy muutamille suurille kaupallisille toimijoille. Kaupallisilla toimijoilla on useasti taloudellisia intressejä kerätä käyttäjistä mahdollisimman paljon dataa. Valtioilla voi puolestaan olla omia poliittisia intressejä päästä käsiksi kerättyyn, välitettyyn ja säilytettävään dataan. On myös muistettava se, että kaikissa valtioissa poliittinen ilmapiiri sekä paikallinen lainsäädäntö elävät jatkuvasti, joten vaikka valtiot eivät juuri nyt hyödyntäisivätkään dataa aktiivisesti, niin tilanne saattaa tulevaisuudessa muuttua. Datan keräämisen ja keskittymisen mahdollisia pitkän aikavälin yhteiskunnallisia vaikutuksia on hyvä pohtia.

Erilaisten dataa keräävien ohjelmien ja palvelujen käyttäminen ei ole pelkästään henkilökohtainen valinta, vaan omat valinnat vaikuttavat myös muiden ihmisten yksityisyyteen. Esimerkiksi toisesta ihmisestä kirjoittaminen jossakin palvelussa vaikuttaa tämän nimenomaisen henkilön yksityisyyteen, vaikkei hän itse käyttäisikään kyseistä palvelua. Samaan tapaan erilaisten tietojen ja tiedostojen säilyttäminen pilvitallennuspalveluissa, jotka sisältävät

jotakin tietoa muista ihmisistä, vaikuttavat kyseisten henkilöiden yksityisyyteen.

Käyttämällä ohjelmia ja laitteita jotka jatkuvasti nauhoittavat ympärillä kuuluvaa puhetta, myös muiden ihmisten puhe tai keskustelut voivat joutua nauhoitetuiksi, täysin ilman heidän tietämystään tai suostumustaan. Koska useimmiten sähköpostiviestit ovat sen palveluntarjoajan luettavissa, joka sähköpostipalvelua ylläpitää, niin omalla sähköpostipalveluntarjoajan valinnalla pakottaa myös muut ihmiset käsittelemään heidän omia henkilötietojaan kyseisen palveluntarjoajan kautta. Jokin dataa keräävä palvelu saattaa myös muuttua niin suosituksi, että palvelun käyttämättä jättäminen ei ole monelle realistinen vaihtoehto, ellei ole valmis osittain jättäytymään sosiaalisen kanssakäymisen ja muun yhteiskunnan ulkopuolelle.

Nykytilanteeseen uusina haasteina ovat myös tulossa autonomiset ajoneuvot, autonomiset dronet, lisätyn todellisuuden laitteet, valvontakameroiden jatkuva lisääntyminen ja nopeasti kehittyvät ihmisentunnistusalgoritmit. Edellä mainitut autonomiset sekä lisätyn todellisuuden laitteet tarvitsevat toimiakseen kameroita ja muita sensoreita. Kuka hallitsee näiden edellä listattujen laitteiden avulla kerättyä dataa ja kenellä on pääsy tähän dataan? Minkälaisia vaikutuksia sillä voi olla ihmisten yksityisyyteen? Toisaalta kannattaa kysyä, että minkälaisia asioita nämä autonomiset tai lisätyn todellisuuden laitteet voivat mahdollistaa?

Kurssilta Algoritmisen ajattelun kehittäminen

18.05.20