

▶ CLAUDIA BERGROTH, OLLE JÄRV, HENRIKKI TENKANEN, TUULI TOIVONEN

# Matkapuhelinaineisto paljastaa kaupungin dynamiikan

*Suomalaiset paikkatietoasiantuntijat ovat tottuneet käyttämään yhtä maailman tarkimmista väestörekisteriaineistoista. Yhä useammin tarvitaan kuitenkin tietoa myös ihmisten liikkumisesta. Matkapuhelinaineistot kertovat tästä paljon.*

**M**atkapuhelinyhtiöt, mobiilisovellukset ja sosiaalisen median alustat tallentavat ihmisten sijaintitietoja omiin tarkoituksiinsa yhä tarkemmin. Vaikka aineistot kerätään ensisijaisesti liiketoiminnan tueksi, on niillä mahdollista myös ymmärtää ja kehittää yhteiskunnan toimintoja.

Siinä missä sosiaalista mediaa tai urheilusovelluksia käyttää yleensä vain rajattu joukko ihmisiä, matkapuhelinten käyttö ulottuu lähes koko väestöön. Matkapuhelinyhtiöiden keräämä sijaintitieto kertoo parhaiten väestön tosiaikaisesta sijainnista millä ajanhetkellä tahansa. Tällaista dynaamista, ajassa muuttuvaa väestötietoa voitaisiin käyttää esimerkiksi liikenteen, terveystietojen ja kaupan toimintojen suunnitteluun tai pelastustoiminnan ohjaamiseen.

## Matkapuhelinaineistoa on monenlaista

**M**atkapuhelinyhtiöiden keräämää sijaintitietoa on karkeasti kahdenlaista. Verkkopohjainen aineisto kuvaa verkon antennien kuormitusta. Puhelintoimintojen tiedot puolestaan paljastavat yksittäisten SIM-korttien sijainnit.

Verkkokuormitusaineisto sisältää kaikkien puhelinten sijainnit huolimatta siitä, käytetäänkö puhelinta aktiivisesti vai ei. Valtavaa aineistoa ei yleensä tallenneta puhelimen tarkkuudella, vaan aineisto aggregoidaan antennitasolle, jolloin henkilöiden yksityisyydensuoja pysyy turvattuna. Aineiston avulla voidaan tunnistaa väestön jakauman tietyllä hetkellä ja selvittää väestön alueellista dynamiikkaa vaikkapa vuorokauden kuluessa.

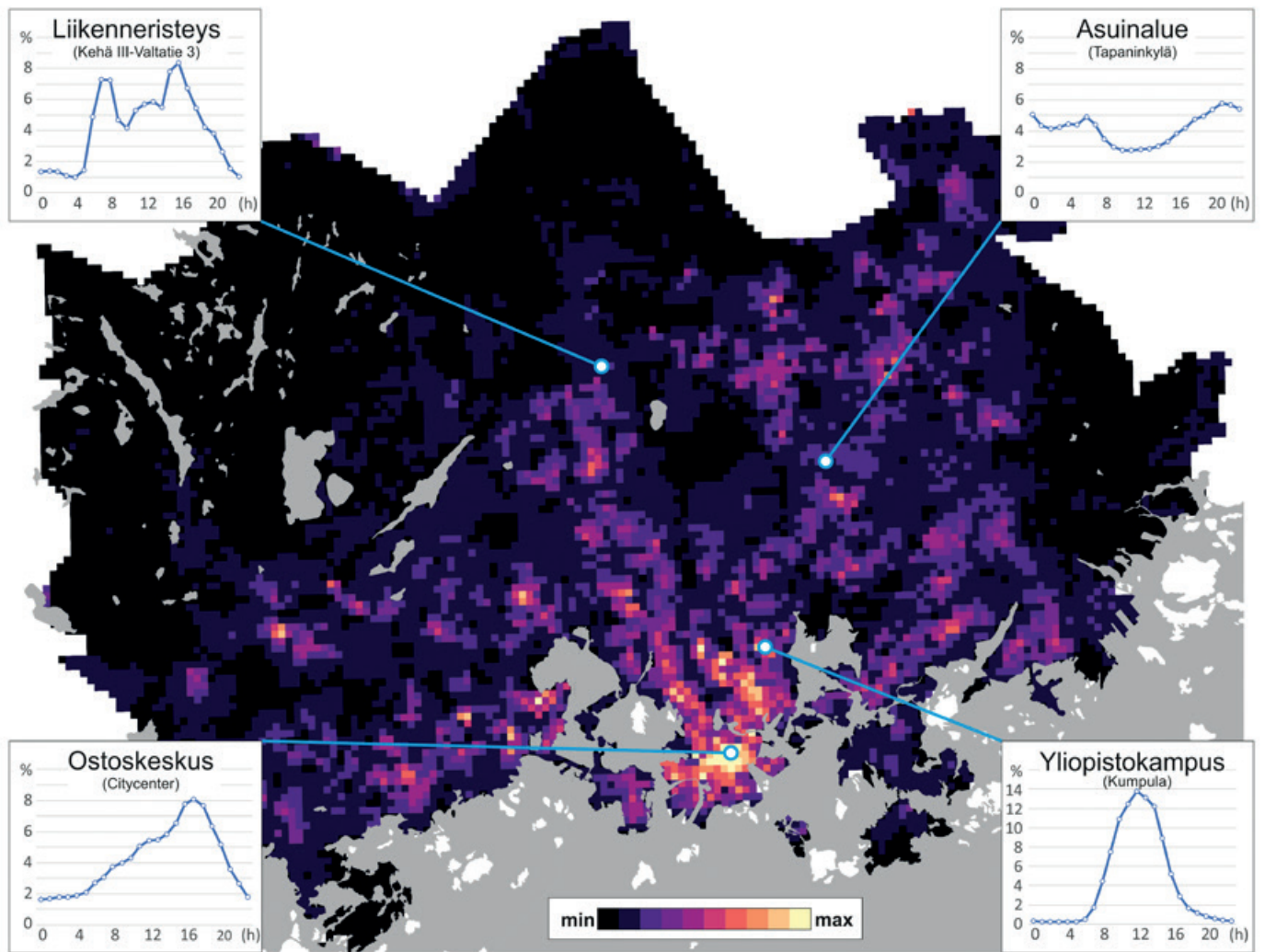
Puhelintoimintojen aineisto (call detail records, CDR) on käyttäjäkohtaista ja tallentuu lakisääteisesti laskutusta varten operaattorin tietokantaan. Näin käy aina, kun puhelimella soittaa, lähettää tekstiviestin tai käyttää datayhteyttä esimerkiksi sovelluksilla. Puhelintoimintojen sijainnit tallentuvat tukiasemien sijaintien tarkkuudella. Aineisto soveltuu hyvin ihmisten liikkumisrakenteiden, tärkeiden paikkojen (koti, työ) ja toiminnallisten alueiden tunnistukseen sekä sosiaalisten verkostojen tarkasteluun.

Vaikka operaattoreilla tätä tietoa on, yksilönsuojan kysymykset yleensä rajoittavat tiedon käyttöä esimerkiksi tutkimuksessa. Yksilötason tietoja voitaisiin kuitenkin yleistää ja saada tärkeää ymmärrystä väestön liikkeistä alueittain ja väestöryhmittäin.

## Tutkimus edistää matkapuhelinaineistojen käyttöä

**H**elsingin yliopiston Digital Geography Labissa (DGL) matkapuhelinaineistoa on käytetty tutkimuksessa jo jonkin aikaa, ensin Virossa ja nyt myös Suomessa. Alueellisen tutkimuksen kannalta matkapuhelinaineistojen haasteena on muuttaa pistemuotoinen tukiasematieto väestöpinnaksi. Yleisesti tässä käytetään teoreettisia kuuluvuusalueita, mutta haasteena on kuuluvuusalueiden datan yhdistäminen muihin aineistoihin, kuten tilastoruutuihin.

DGL:ssa kehitetty dasymetrinen interpolointimenetelmä jakaa kuuluvuusalueen aineiston esimerkiksi tilastoruutuihin hyödyntäen väestön ajankäyttötilastoja, maankäyttötietoa sekä rakennusten pinta-alaa ja käyttötarkoitusta. Väestön dynamiikkaa on analysoitu tilastoruututasolla ja sen vaikutusta on tutkittu esimerkiksi saavutettavuuslaskelmiin Tallinnassa ja Suomessa pääkaupunkiseudulla. Dynaamisessa saavutettavuusanalyysissä huomioidaan dynaamisen väestön rinnalla myös liikenneverkon ja palveluiden tarjonnan vuorokausirytmit. ▶



Kartta kuvaa pääkaupunkiseudun väestöjakaumaa arkipäivänä klo 12–13 esitettyinä 250 m x 250 m tilastoruuduilla. Diagrammit kuvaavat neljän yksittäisen ruudun väestömäärän vaihtelua vuorokauden aikana. Väestöjakauma perustuu verkkokuormitusaineistoon.

KUVA: CLAUDIA BERGROTH / DGL

## Aineistojen saatavuus vielä ongelma

DGL-työryhmässä tehty tutkimus ja erityisesti Claudia Bergrothin pro gradu -työn analytiikka paljastavat väestön vuorokauden aikaisen pulssin pääkaupunkiseudulla 250 m tilastoruutujen tarkkuudella tunneittain.

Aineiston tarkastelu osoittaa yksityiskohtaisesti, missä määrin väestöä siirtyy asuinalueilta vaikkapa työpaikkakeskittymiin tai puistoihin päivän aikana. Ruutukohtaiset pulssit näyttävät suunnataanko kotiin harrastusten tai kaupan kautta. Seudun yleinen väestöllinen pulssi näyttää, että Helsingin keskusta on tärkeä väestökeskittymä päiväsaikaan, mutta väkeä on paljon myös osakeskuksissa ja asuinalueilla. Myös työmatkaliikenteen alueet ovat merkittäviä väestökeskittymiä osan aikaa päivästä.

Matkapuhelinyhtiöiden keräämä tieto parantaisi yhteiskunnan ilmiöiden ja

prosessien ymmärtämistä, esimerkiksi ihmisten arjen alueet ja niiden mahdollinen eriytyminen, valtioiden välinen liikkuminen ja kesämökkeily tai eri ihmisryhmien sosiaalinen vuorovaikutus ja integraatio. Vaikka aineistoja on alettu kaupallistaa markkinoinnin, liikenteen ja matkailun sovelluksiin, on aineistojen saatavuus tutkimuskäyttöön edelleen ongelma. Yhteiskunnan hyvään ymmärtämiseen tarvitaan operaattorien ja tutkijoiden tiiviimpää yhteistyötä, vaikka nykyinen lainsäädäntö tutkimusyhteistyötä rajoittaakin. ◀

- ▶ Pääkaupunkiseudun dynaaminen väestö animoituina: <https://bit.ly/2DWdaKK>
- ▶ Tieteellinen artikkeli dynaamisesta saavutettavuusmallinnuksesta: <https://bit.ly/2RjAkNn>
- ▶ Tieteellinen artikkeli dasymetrisestä interpolointimenetelmästä: <https://bit.ly/2TrnEU>

CLAUDIA BERGROTH VIIMEIS-TELEE PRO GRADU -TYÖTÄÄN HELSINGIN YLIOPISTOLLA PÄÄKAUPUNKISEUDUN DYNAAMISEN VÄESTÖN ANALYSOINNISTA MATKAPUHELINAINESTOILLA. OLLE JÄRV ON VÄITELLYT MATKAPUHELINAINESTOJEN KÄYTTÖSTÄ VIROSSA. HÄN TOIMII POST-DOC-TUTKIJANA DGL:SSA. HENRIKKI TENKANEN ON VÄITELLYT DYNAAMISISTA SAAVUTETTAVUUSANALYYSI- TÄ DGL:SSA. HÄN ON POST-DOC-TUTKIJANA UNIVERSITY COLLEGE LONDONISSA. TUULI TOIVONEN ON GEOINFORMATIIKAN APULAISPROFESSORI JA VETÄÄ DGL-RYHMÄÄ HELSINGIN YLIOPISTOLLA. SÄHKÖPOSTI: ETUNIMI.SUKUNIMI@HELSINKI.FI