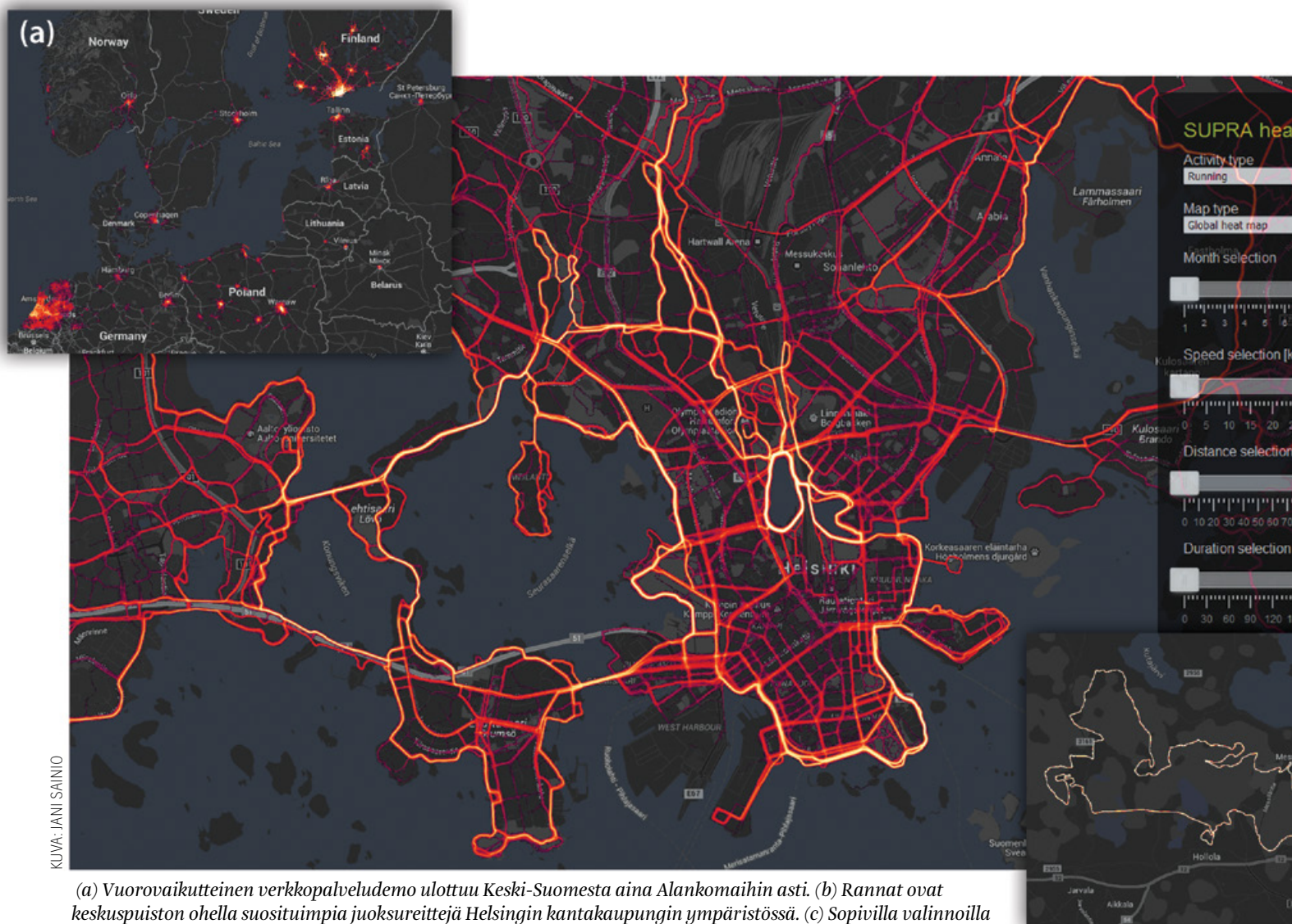


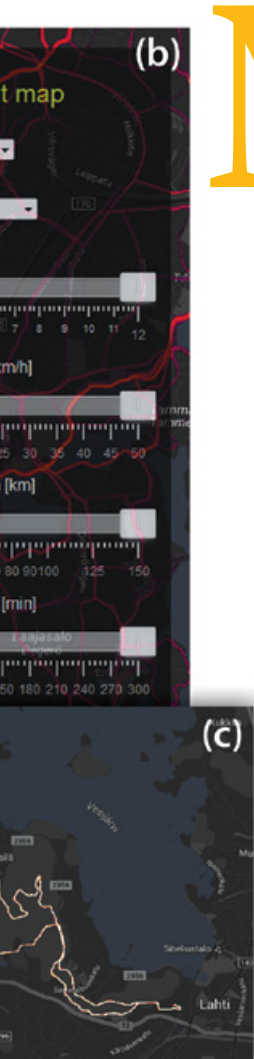
▶ JUHA OKSANEN, CECILIA BERGMAN, JANI SAINIO, JAN WESTERHOLM



KUVA: JANI SAINIO

Mistä löytyy paras pyöräilyreitti tai lenkkeilymaasto?

Supra toi uutta tietoa suosittujen reittien laskentaan.



Millaista lisäarvoa pystymme tuottamaan, jos kytkemme tehokkaan paikkatietoanalyysin suurelle yleisölle suunnattuun sijaintipohjaiseen palveluun? Kysymystä pohdittiin Paikkatietokeskuksen ja Åbo Akademin Supra-hankkeessa, jossa yhdistettiin hankkeen tutkimusryhmien aiemmat kokemukset tehokkaasti toteutetuista paikkatietoanalyseistä.

Hankkeessa tutkittiin mobiilin liikuntasovelluksen käyttäjäyhteisön keräämien kulkureittien hyödyntämismahdollisuuksia lisäarvopalvelujen pohjatietoina. Tutkimusideoita demonstroitiin vuorovaikutteisella web-sovelluksella, jolla voidaan tarkastella eri liikuntalajien harrastajien suosituimpia reittejä. Tietoja voidaan suodattaa liikuntasuorituksen pituuden, keskinopeuden, keston sekä vuodenajan suhteen. Tulokset myös osoittavat, miten joukkoistamalla kerättyä tietoa voidaan hyödyntää reitityspalveluiden rakentamisessa ja miten kaupunkisuunnitteluun on saatavissa kokonaan uudenlaista big data -analytiikkaan perustuvaa tietopohjaa ihmisten käyttäytymisestä kevyessä liikenteessä.

Yhtenä tutkimuskumppanina hankkeessa toimi Sports Tracking Technologies Oy (nyk. osa Amer Sports Suomi Oytä), joka tarjosi projektille tietokannastaan nimettömiksi muutettuja Sports Tracker-liikuntasovelluksen käyttäjien julkisiksi asettamia liikuntasuoritustietoja vuosilta 2010 – 2012 aina Alankomaista Keski-Suomeen asti.

Yksityisyyden säilyttävistä suosiokartoista lisäarvoa

Kulkureittitietojen käsittely ja eri liikkumis-
muotojen suosion mittaaminen osoittautui
heti hankkeen alussa haastavaksi. Miten
voidaan taata, ettei yksittäistä sovelluksen käyttäjä voi tunnistaa suosiokartalta? Mikä itse asiassa on suosittu reitti tai tässä yhteydessä oikeammin paikka? Miten valtavan aineistomassan analysointi valjastetaan osaksi sijaintipohjaista palvelua ja mitä suosiokartat lopulta kertovat ihmispopulaation käyttäytymisestä?

Keskeistä suosiokarttojen laatimisessa oli huolehtia tiedon keränneiden liikuntasovelluksen käyttäjien yksityisyyden suojasta. Sovelsimme Supra-reititutkimuksessa yleisesti käytettyä k-anonymiteettiperiaatetta, jossa jokaisen käyttäjän tallentamat suoritukset sekoittuvat vähintään ennalta asetetun kokoisen käyttäjäjoukon suoritusten kanssa. Meidän suosiokartoissamme tämä tarkoittaa sitä, että kartalta on luettavissa informaatiota vasta silloin, kun vähintään viisi eri käyttäjää on kulkenut paikan läpi.

Intuitiivisesti ajatellen liikuntasuoritusten määrä

on paras suosion mittari. Todellisuudessa liikuntasuoritusaineistoa, niin kuin kaikkea muutakin joukkoistamalla kerättyä aineistoa, luonnehtii epä-tasainen osallistuminen (participation inequality), joka estää hätäisten johtopäätösten teon aineistosta.

Epätasaisella osallistumisella tarkoitetaan sosiaalisen median ja online-yhteisöjen ominaisuutta, jonka mukaan 90 prosenttia käyttäjistä on sisältöä tuottamattomia seuraajia, yhdeksän prosenttia tuottaa sisältöä ajoittain ja yksi prosentti tuottaa sisällöstä suurimman osan.

Jos suosiokartta tehtäisiin pelkäämään liikuntasuoritusten määrän perusteella, kertoisimme vain liikuntasovelluksen aktiivisimman käyttäjäjoukon käyttäytymisestä. Ongelmaa voitaisiin kiertää perustamalla suosiotieto käyttäjämääriin, mutta siinä tapauksessa jätämme liikuntasuoritusten määrän hyödyntämättä. Silloin esimerkiksi katuosuus, jossa 10 eri käyttäjää on tallentanut 10 suoritusta, tulisi yhtä suosituksi kuin katuosuus, jossa 10 eri käyttäjää on tallentanut 100 suoritusta. Ongelman ratkaisuksi keksittiin liikuntasuoritusten määrän korjaaminen käyttäjistä lasketulla diversiteetti-indeksillä. Indeksii kuvaa sitä, miten monta eri käyttäjää ja millaisella määräsuhteella paikalla on kulkenut. Näin suosituimmiksi tulevat paikat, joissa liikuntasuoritusten määrä on suuri ja joissa suoritustiedon ovat keränneet mahdollisimman monet eri henkilöt.

Konkreettisen lisähaasteen muodosti mobiilisovelluksella kerätyn reitittiedon laatu ja määrä

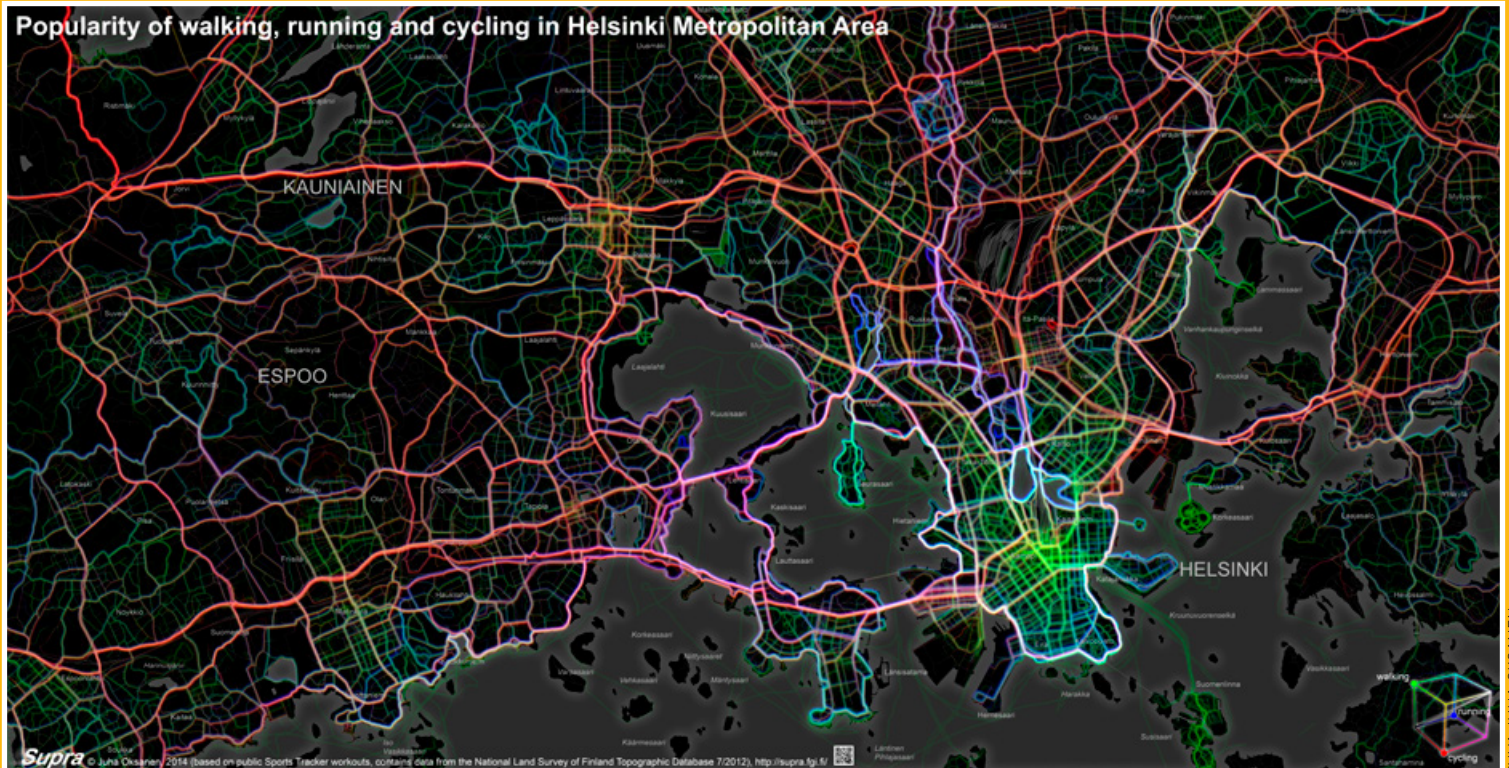
LISÄTIETOJA:

- Hankkeen verkkosivut <http://carde.fgi.fi/supra>
- Supra-demon verkkosivut <http://supra.abo.fi>

Supra-hanke oli Tekesin strateginen tutkimus-
avaus, joka toteutettiin vuosina 2013 – 2015.
Hankkeessa etsittiin ratkaisuja liikuntasuoritustie-
tojen tehokkaaseen analyysiin sekä lähtötietojen
epävarmuuden huomioiviin merivesitulvien
peittävyyskenaarioiden luontiin.

Hankkeen vastuullisina tutkijoina olivat Juha
Oksanen, Maanmittauslaitoksen Paikkatietokes-
kus (FGI) ja Jan Westerholm, Åbo Akademi (ÅA).
Tutkijoina hankkeessa toimivat Cecilia Bergman
(FGI), Jani Sainio (ÅA), Janne Kovanen (FGI),
Susanne Suvanto (FGI), David Eränen (ÅA) ja
Andreas Åkesson (ÅA).

Supra-hankkeen demopalvelu voitti Kansain-
välisen kartografisen seuran (ICA) karttanäyttelyn
sarjan Digitaaliset tuotteet Rio de Janeirossa
elokuussa 2015: <http://icaci.org/Map-of-the-Month-102015/>



Kartta esittää suosituimpia liikuntareittejä pääkaupunkiseudulla. Laskennassa on otettu huomioon sekä suoritusten määrä että kuinka monta henkilöä on paikan ohhi kulkenut.

sekä sen tehokas prosessointi. Tallennetut kulkureitit tekivät silloin tällöin selittämättömiä melkein yli valonnopeudella tehtyjä hyppäyksiä, ja eräskin pyöräilijä näytti menevän kulkureitin mukaan yli sadan kilometrin tuntivauhdilla Helsingistä Ouluun ja sieltä Kajaaniin rautatietä pitkin! Asettamalla sopivia rajoituksia saatiin esille yhteensä 2,8 miljardista GPS pisteestä koostunutta 800 000 kelvollista liikerataa.

Näihin tietoihin perustuen julkaistiin verkkopalvelu, jossa käyttäjä valitsee paikan ja zoom-tason lisäksi mistä liikuntamuodosta hän on kiinnostunut, mitkä kuukaudet otetaan mukaan ja lisärajoituksina vielä reitin pituus, kesto sekä keskinopeus.

Työmatkaliikkujat tallentavat reittejään aktiivisesti

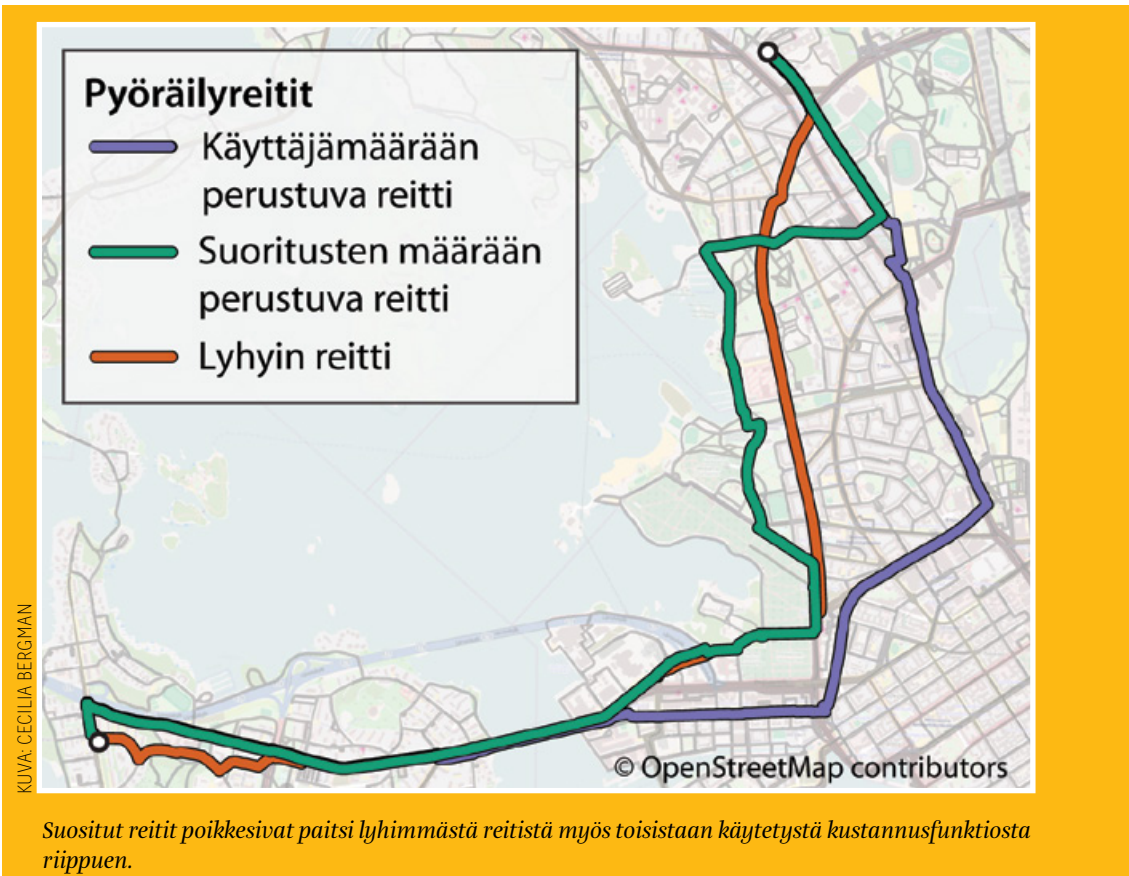
Koska käyttäjän valitsemien rajausten yhdistelmiä on miljoonia, ei verkkopalvelu voi perustua staattisiin valmiiksi laskettuihin suosiokarttoihin, vaan suosiokartta täytyy laskea erikseen jokaista käyttäjää varten. Annettujen rajoitusten puitteissa verkkopalvelin etsii mukaan tulevat reitit, laskee niiden käyttäjämäärät ja päättelee, näytetäänkö reittiosuutta ollenkaan vai liitetäänkö se suosiokarttaan ja millä painolla. Karttanäkymä on siten kovinkin erilainen riippuen asetetuista rajoituksista, ja käyttäjä voi tutkia hyvin tarkasti miten suo-

sittu jokin tietty katuosuus on erilaisin kriteerein. Kaikki tämä verkkopalvelimen suorittama laskenta ei kuitenkaan saa kestää liian kauan. Stressitestimme osoittivat, että palvelimen vasteaika tuhannen samanlaisen käyttäjän tapauksessa on kymmenissä millisekunneissa.

Kun tuloksia tarkastellaan mobiilin liikunta-sovelluksen kontekstin ulkopuolella ja ajatellaan suosiokarttojen käyttömahdollisuuksia laajemmin esimerkiksi osana kaupunkisuunnittelua, lisähaasteeksi muodostuu kysymys siitä, mitä joukkoa suosiokartta todellisuudessa edustaa.

Pyöräilyn tapauksessa havaitsimme, että Sports Tracker -sovelluksen käyttäjät eivät rajoitu tavoitteellisiin urheilun harrastajiin, vaan merkittävä osa reiteistä oli työmatka- tai muiden hyötyliikkujen tallentamia. Tästä kertoi myös suosiokarttojen korkea korrelaatio Helsingin kaupungin pyörälaskenta-aineiston kanssa, mikä mahdollistaa karttojen käytön kaupunkisuunnittelun pohjatietona. Kalibroimalla suosiokartta laskenta-aineistolla saavutetaan

Suosiokarttojen avulla voidaan tuottaa uutta tietoa esimerkiksi kaupunkisuunnitteluun.



Suosittu reitti poikkesivat paitsi lyhimmästä reitistä myös toisistaan käytetystä kustannusfunktioista riippuen.

sekä suunnittelun tarvitsema tieto absoluuttisista liikkujamääristä että big data -pohjaisen analyysin alueellinen kattavuus.

Reititys murroksessa

Tulevaisuuden reititys ei perustu pelkästään matkan pituuden tai matkaan kuluneen ajan optimointiin, vaan haluamme löytää tiedon siitä, mikä on paras reitti lähtöpisteen ja määränpään välillä. Suuret reittiaineistot mahdollistavat liikkujien paikallistuntemuksen hyödyntämisen reitityksessä ja antavat arvokasta tietoa verkon soveltuvuudesta erityyppisille liikkujille.

Lähestyimme parhaan reitin ongelmaa suosion näkökulmasta – vai tulisiko sanoa turvallisuuden, sillä onhan pyöräily Safety in numbers -hypoteesin mukaan sitä turvallisempaa, mitä enemmän väylällä on käyttäjiä. Yhdistimme pyöräilijöiden suoritukset liikenneverkkoon laskemalla todennäköisimmän reitin Markovin piilomalliin perustuvalla menetelmällä. Näin saatoimme hakea optimaalisen reittiehdotuksen ns. lyhimmän polun algoritmilla käyttäen kustannuksena normalisoitua suosiotietoa. Verkkona käytimme OpenStreetMapin (OSM,

openstreetmap.org) pyöräilijöiden käytettävissä olevia väyliä.

Tulosten mukaan käytetty kustannus eli impedanssi vaikuttaa olennaisesti reititykseen, ja myös tässä yksi vaihtoehto olisi hyödyntää diversiteettiindeksillä korjattua suoritusten määrää. Liikeratojen ja verkon yhdistämisessä haasteiksi osoittautuivat GPS-datan vaihtelevan tarkkuuden ohella verkolla itse suorituksen tallentamisen jälkeen tapahtuneet muutokset, OSM:n yksityiskohtaisuus sekä aukiot.

Kaikkien vapaasti muokattavissa oleva OSM osoittautui Helsingissä ainakin suosituimpien pyöräilyreittien osalta jo varsin kattavaksi aineistoksi. Joillakin

muilla alueilla verkon puutteet saattaisivat kuitenkin vaikuttaa merkittävästi palvelun laatuun ja osoittautua ongelmallisiksi loppukäyttäjän näkökulmasta. ◀

Käyttäjät voi tutkia reittien suosiota erilaisin kriteerein.

JUHA OKSANEN TOIMII PAIKKATIEDON VISUALISOINNIN JA ANALYYSIN TUTKIMUSRYHMÄN (GEOVA) JOHTAJANA GEOINFORMATIIKAN JA KARTOGRAFIAN OSASTOLLA MAANMITTAUSLAITOKSEN PAIKKATIEDOKUKKUNNASSA. JUHA.OKSANEN@MAANMITTAUSLAITOS.FI

CECILIA BERGMAN TOIMII TUKIJANA PAIKKATIEDON VISUALISOINNIN JA ANALYYSIN TUTKIMUSRYHMÄSSÄ GEOINFORMATIIKAN JA KARTOGRAFIAN OSASTOLLA MAANMITTAUSLAITOKSEN PAIKKATIEDOKUKKUNNASSA. CECILIA.BERGMAN@MAANMITTAUSLAITOS.FI

JANI SAINIO TOIMII TUKIJANA ÅBO AKADEMISSA. NYKYISIN HÄN ON ERIKOISTUTKIJANA TURUN KAUPPAKORKEAKOULUN CCR TUTKIMUSPALVELUISSA JA DATATIETEILIJÄNÄ LOKUS DIGITAL OY:SSÄ. JANI.SAINIO@UTU.FI

JAN WESTERHOLM TOIMII SUURTEHOLASKENNAN PROFESSORINA ÅBO AKADEMISSA. JAN.WESTERHOLM@ABO.FI