

MS-E2177 Operaatiotutkimuksen projektityöseminaari

# Junien peruuntumistodennäköisyyksien hyödyntäminen veturinkuljettajien työvuoroluetteloiden suunnittelussa

Projektisuunnitelma

21.2.2017

Tapio Hautamäki, 345312 (projektipäällikkö)

Juhani Mutikainen, 218245

Johanna Piipponen, 292180

## 1 Tausta

Veturinkuljettajat VR:n junaliikennöintidivisioonassa ajavat sekä kaukoliikenteen henkilöjunia että tavarajunia. Yhteen työvuoroon voi kuulua sekä henkilö- että tavarajunia, ja yksi työvuoroluettelo on kolmen viikon pituinen. Työvuoroluettelot annetaan kuljettajille etukäteen ja luetteloon on hankala tehdä muutoksia myöhemmin, vaikka muutoksille on usein tarvetta. Henkilöjunia perutaan harvoin, mutta tavarajunien kannalta tilanne on toinen: tavarajunia voidaan perua työvuoroluetteloiden laatimisen jälkeen tai niitä voi tulla lisää.

Kaikilla veturinkuljettajilla on täyspäiväiset työsopimukset, joten työvoiman tarjonta ei joustu tarjolla olevien työtuntien mukaan. Osa veturinkuljettajista on kuitenkin "heittomiehiä", joille työvuoroluetteloita ei tarvitse julkaista viikkoja etukäteen vaan vasta edellisenä iltana. Nämä heittomiehet hoitavat ne ylimääräiset työvuorot, joille ei ole kuljettajia, ja heitä on enintään 17 % kuljettajista. Työvuoro jää heittomiehille ajettavaksi, kun vuoron oikea kuljettaja on yllättäen poissa, kun tulee tarve lisäjunien ajamiselle, tai kun vuoroa ei ole suunniteltu työvuoroluetteloon. Kaksi ensimmäistä tapausta tapahtuvat aina satunnaisesti, mutta päätös siitä, mitkä työvuorot kiinnitetään työvuoroluetteluihin ja mitkä jätetään heittoon, on suunnittelijoiden harkittavissa.

Työvuorojen sijoittaminen joko työvuoroluetteluihin tai heittoon vaikuttaa siihen, kuinka monta tuntia veturinkuljettajien työajasta käytetään hyödyksi. Veturinkuljettajien työaika kolmen viikon jaksolla on enintään 114 h 45 min. Tällä hetkellä työvuorojen sisältämien junien peruuntumisriskiä ei analysoida järjestelmällisesti. Syntyy tilanteita, joissa vuoron kuljettajien työtunnit jäävät täyttymättä junien peruuntumisten vuoksi, heittomiesten joustavuutta ei hyödynnetä riittävästi, ja lisäjunia joudutaan ajamaan pakon edessä ylitöinä. Tavoite on, että kiinteää vuoroa ajavat kuljettajat saavat maksimityötuntinsa täyteen eli junia ei peruuntuisi vuorosta. Toisaalta heittomiehet ajaisivat ne työvuorot, jotka ovat jääneet yli sekä lisäjunat, joista ei ollut tietoa työvuoroja laadittaessa.

## 2 Tavoitteet

Tämän projektin tavoitteena on tuottaa työvuorosuunnittelijoille päätöksentekotukea siitä, millaiset työvuorot tulisi sijoittaa työvuoroluetteluihin ja mitkä jätetään heittoon niiden sisältämien peruuntumisherkkien junien vuoksi. Asiakkaalle tulisi tuottaa analyyttinen työkalu, joka jaottelee työvuorot kolmeen kategoriaan: mitkä vuorot tulisi sijoittaa työvuoroluetteluihin, mitkä vuorot tulisi ehkä sijoittaa luetteluihin ja mitä vuoroja ei ainakaan sijoiteta luetteluihin. Työkalulla pyritään tuottamaan helppokäyttöistä tietoa erityisesti keskimmäisen kategorian vuoroista päätöksentekoa avustamaan.

Tavoitteena on analysoida, millaisia vaikutuksia yksittäisen junan peruuntumisella on työvuoron ajamiseen ja kannattaisiko vuoron niiden perusteella olla työvuoroluettelossa vai heitossa. Junan peruuntumistodennäköisyyttä estimoidaan kulkuhistoriadataan perusteella, jota on saatavilla vuoden 2014 kesästä lähtien. Peruuntumistodennäköisyyden mallintamista varten testataan niin aikasarjamalleja kuin klassifiointimalleja.

### 3 Tehtävät

Tiimin ja asiakkaan välisten keskustelujen perusteella työkalun kehittäminen jakautuu kolmeen päätehtävään. Ensimmäiseksi selvitetään, miten junien peruuntuminen vaikuttaa työvuoroon. Työvuoron rakenteen eli sen sisältämien junatyyppien, asematyyppien, taukojen jne. perusteella arvioidaan junien peruuntumisen vakavuutta. Toiseksi tiimin tulee tutustua junien kulkuhistoriatietoihin, jotta junien peruuntumistodennäköisyydet saadaan selville. Kolmanneksi tiimin tulee selvittää, miten nämä tiedot saadaan työvuorosuunnittelijoille käyttökelpoiseen muotoon.

Projekti voidaan jakaa yhteentoista tehtävään, joilla vastataan sekä asiakkaan että kurssin vaatimuksiin.

1. **Projektin suunnittelu asiakkaan kanssa:** Projektin aihe on esitelty lyhyesti alustustilaisuudessa, mutta tarkempi määrittely tapahtuu tiimin ja asiakkaan välisissä tapaamisissa. Tapaamisissa selvennetään epäselviä asioita ja syvennetään ymmärrystä. Tapaamisista laaditaan tiimin käyttöön kokousmuistiot.
2. **Dataan tutustuminen:** Projektia varten käytettävissä on aito työvuoroaineisto VR:ltä ja junien kulkutiedot Liikenneviraston avoimesta tietokannasta. Työn onnistuminen nojaa näiden aineistojen ymmärtämiseen. Työvuoroaineistoa käsitellään tapaamisissa, jotta tiimi voi hyödyntää junaliikennöintidivisioonan kokemusta.
3. **Projektisuunnitelma:** Projektisuunnitelman kirjoittaminen on osa kurssisuoritusta. Sen palautuspäivämäärä on 24. helmikuuta. Projektisuunnitelma kirjoitetaan asiakkaan kanssa tehdyn suunnittelun pohjalta ja siihen pyydetään kommentit asiakkaalta.
4. **Kirjallisuuteen tutustuminen:** Projektia varten hyödynnetään tieteellisiä julkaisuja ja oppikirjoja liittyen mallien valintaan ja validointiin. Tilastollisten mallien osalta tutustutaan tarkemmin aikasarja-analyysiin (ARMA-mallit) ja klassifiointimenetelmiin (logistinen regressio ym.), joita tullaan käyttämään junien peruuntumistodennäköisyyden estimoinnissa.
5. **Junien peruuntumisen vaikutus työvuoroon:** Työkalun kehittämisessä on kolme päätehtävää, joista tämä on ensimmäinen. Työkalua varten mallinnetaan yksittäisen junan peruuntuminen vakavuutta. Tehtävän toteuttaminen vaatii paljon keskustelua työvuorosuunnittelijoiden kanssa, jotta mallissa otetaan huomioon mahdollisimman hyvin junan peruuntumisen vaikutukset työvuoron toteuttamiseen. Monen junan yhtäaikainen peruuntuminen jätetään projektin ulkopuolelle, sillä se vaikeuttaisi analyysia huomattavasti.

6. **Historiadataan hyödyntäminen:** Tämä on työkalun kehittämisen toinen päätehtävä. Liikennevirastolla on avoin tietokanta junien kulkutiedoista. Tämän historiadataan perusteella luodaan tilastollinen aikasarjamalli junien peruuntumisen ennustamiseksi. Tehtävään kuuluu datan lukeminen ja käsittely sekä mallin luominen ja testaaminen. Projektin muut vaiheet ovat vahvasti riippuvaisia tämän tehtävän onnistumisesta. Sen vuoksi historiadataan hyödyntäminen aloitetaan hyvin yksinkertaisen mallin luomisella ja mallia tarkennetaan työn edetessä.
7. **Työkalun käyttöliittymän luominen:** Projektin viimeinen päätehtävä on peruuntumisennusteen yhdistäminen tietoon peruuntumisen vaikutuksista, ja näiden tietojen muokkaaminen päätöksentekoa helpottavaan muotoon. Työvuorosuunnittelijoiden kanssa keskustellaan siitä, miten työkalu toimisi toivotulla tavalla. Työkalusta luodaan demoversioita, joita kokeillaan yhdessä asiakkaan kanssa.
8. **Väliraportti:** Väliraportin kirjoittaminen on osa kurssisuoritusta. Raportin palautuspäivämäärä on 21. huhtikuuta. Väliraportissa tarkistetaan, miten projekti on edennyt, mitä on tehty ja mitä on vielä tekemättä. Koska väliraportin palautus on lähellä kurssin loppua, tulisi suurin osa työstä olla valmis väliraporttiin mennessä.
9. **Dokumentointi:** Työkalu ja sen sisältämät mallit dokumentoidaan tarkasti, jotta asiakas voi halutessaan kehittää niitä tai integroida ne omaan järjestelmäänsä. Dokumentointi koskee erityisesti ohjelmaa, jolla luodaan tilastollinen malli junien peruuntumistodennäköisyyksistä, sillä todennäköisyyksiä tulee tulevaisuudessa pystyä päivittämään uusien kulkutietojen avulla.
10. **Verifiointi ja validointi:** Työkalun verifiointia ja validointia suoritetaan sekä tiimin kesken että asiakkaan kanssa. Tapaamisissa varmistetaan työkalun toimivuus ja tarkoituksenmukaisuus. Junien peruuntumistodennäköisyyksiä mallinnettaessa mallin oletukset käydään läpi huolellisesti ja mallin vahvuudet sekä heikkoudet raportoidaan.
11. **Loppuraportti:** Loppuraportin kirjoittaminen on osa kurssisuoritusta. Raportin palautuspäivämäärä on 19. toukokuuta. Loppuraportissa kuvataan laajasti projektin tulokset. Loppuraporttia tehtäessä päätöksentekoa tukeva työkalu on viimeistelty ja toimitettu asiakkaalle.

## 4 Aikataulu

Projektiaikataulu on hahmoteltu alla näkyvään taulukkoon viikon tarkkuudella. Tärkeimmät päivämäärät koskevat raportteja, joiden tulee olla palautettu aina merkityn ajanjakson loppuun mennessä.

Tehtävä \ Viikkonumero	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Projektin suunnittelu																		
2. Dataan tutustuminen																		
3. Projektisuunnitelma																		
4. Kirjallisuuteen tutustuminen																		
5. Junien peruuntumisen vaikutus																		
6. Historiadatan hyödyntäminen																		
7. Käyttöliittymän luominen																		
8. Väliraportti																		
9. Dokumentointi																		
10. Verifiointi ja validointi																		
11. Loppuraportti																		

## 5 Resurssit

Projektitiimin koko on kolme henkilöä: projektipäällikkö Tapio Hautamäki, Juhani Mutikainen ja Johanna Piipponen. Kurssin mitoitus on viisi opintopistettä eli 133 tuntia. Tämä tarkoittaa 400 henkilötyötuntia. VR:n puolelta asiakkaita ovat junaliikennöintidivisioonan Antti Punkka, Joonas Ollila (alkuvaiheessa), Teppo Koskinen ja Kalle Rantanen. Tapaamiset VR:n kanssa hoidetaan toimistotyöaikaisten puitteissa ja niitä on enintään kahdesti kuukaudessa.

Työmateriaaleja on saatavissa useita. VR tarjoaa projektin käyttöön aidon työvuoroaineistoa ajalta 23.1.–12.2.2017. Tiedostot jaetaan VR:n Sharepoint-työtilassa, jonka tallennustilaa voidaan tarvittaessa käyttää muuhunkin. Liikennevirasto julkaisee junien täsmällisyystietoja avoimen rajapinnan kautta Digitraffic-palvelussa (<https://rata.digitraffic.fi/api/v1/doc/index.html>). Rajapintaa käyttää myös Julia-palvelu, jonka visuaalisesta esityksestä on hyötyä varsinkin projektin alkuvaiheessa. (<https://julia.dy.fi/>)

Tiimi käyttää työskentelyyn Slack-keskustelualustaa ja Google Drive -pilvipalvelua. Slack on ryhmille ja organisaatioille suunnattu viestintätyökalu, joka mahdollistaa asiakkaiden tiiviimmän osallistumisen projektiin. Google Drive toimii pöytäkirjojen ja palautettavien raporttien työvälineenä.

Projektissa tuotettava työkalu toteutetaan R-kielellä. R on avoimeen lähdekoodiin perustuva tilastollisen laskennan ohjelmointikieli, jota myös VR:n suunnittelijat käyttävät.

## 6 Riskit

Projektiin liittyvät riskit on jaoteltu dokumentin lopussa olevaan taulukkoon niiden todennäköisyyden ja vaikutuksen mukaan (pieni, keskisuuri, suuri). Suurin osa riskeistä voidaan välttää hyvällä suunnittelulla ja avoimella kommunikaatiolla tiimin ja asiakkaan välillä.

Tiimin sisäiseen dynamiikkaan liittyvät uhkia ovat tiimin jäsenten poissaolot ja projektipäällikön epäonnistuminen. Molempien riski on pieni, mutta vaikutukset ovat suuria, mikäli kaksi jäsentä joutuu tekemään kolmen jäsenen työt, tai keskisuuria, mikäli tiimin jäsenet joutuvat ohjaamaan toinen toistaan ilman yhtä yhteistä päällikköä. Molempiin uhkiin pätevät suunnilleen samat keinot: tiimin sisäinen keskustelu, palautteenanto ja avoin vuoropuhelu.

Aikatauluongelmien todennäköisyys on keskisuuri, sillä projektia tehdään kaikkien muiden kurssi- ja työtehtävien ohella. Aikatauluongelmien tapahtuessa stressi kasvaa ja työn laatu voi kärsiä. Uhkaa voidaan lieventää tapaamalla useammin tiimin kanssa, työskentelemällä yhdessä asetettujen määräaikojen puitteissa ja tapaamalla useammin asiakasta, mikä on rohkaisevaa varsinkin projektin alkupuolella.

Itse työn tekemisessä ongelmia voi syntyä historiadatan käsittelyn kanssa. Hankalaksi tämän uhan tekee se, että historiadata ei tule VR:ltä vaan Liikennevirastolta, joten ongelmien kanssa tulisi kääntyä heidän puoleensa. Tiimin tulee yhdessä keskittyä historiadatan purkamiseen, jotta koko vastuu ei ole vain yhdellä jäsenellä. Jos historiadatan käsittely onnistuu, ongelmia voi syntyä mallinvalinnan kanssa. Epäonnistuminen mallinvalinnassa on suuri isku koko projektille ja aiheuttaa tappioita asiakkaalle. Uhan toteutumista voi estää aikasarja-analyysin kertaamisella tiimin voimin ja erilaisten mallien kokeilun dokumentoinnilla.

Projektin tavoitteena on luoda työkalu, jota suunnittelijat voivat käyttää päättäessään työvuorojen sijoittelusta. Mikäli työkalu on suunnittelijoille epämieluisen, sitä ei käytetä. Toisaalta työkalun taustatiedot kuten peruuntumistodennäköisyydet olisivat edelleen olemassa ja niitä voitaisiin hyödyntää muilla tavoin. Uhkaa estetään tiiviillä työskentelyllä suunnittelijoiden kanssa ja beetatestaamalla työkaluja aikaisessa vaiheessa, jotta niitä voidaan kehittää palautteen perusteella.

Uhka	Todennäköisyys	Vaikutus	Torjuminen
<b>Tiimin jäsenet eivät ole sitoutuneita.</b>	Pieni	Suuri: Työmäärä on mitoitettu kolmelle jäsenelle	Tiimin sisäinen keskustelu ennen projektin aloitusta
<b>Projektipäällikkö ei suoriudu tehtävästään.</b>	Pieni	Keskisuuri: Suunniteltu aikataulu ei pidä ja työn laatu kärsii.	Palautteenantoa tiimin sisällä, tiimin jäsenten välinen vuoropuhelu
<b>Aikataulu ei pidä.</b>	Keskisuuri	Keskisuuri: Stressi kasvaa ja työn laatu kärsii.	Tiimitapaamisia useammin, tiimin sisäisiä määräaikoja, tapaamiset asiakkaan kanssa
<b>Ongelmat historiadatan kanssa</b>	Pieni	Keskisuuri: Aikasarjamallin luominen vaikeutuu ja hidastuu.	Dataan keskitytään tiimin voimin, lisätietoja voi kysyä myös Liikennevirastolta.
<b>Mallin valinta epäonnistuu.</b>	Keskisuuri	Suuri: Tärkein osa koko projektia ja aiheuttaa tappioita asiakkaalle.	Aikasarja-analyysin kertaaminen, oletuksien läpikäynti VR:n kanssa, dokumentointi
<b>Suunnittelutyökalu ei ole suunnittelijoiden mieleen.</b>	Pieni	Keskisuuri: Työkalua ei käytettäisi, mutta taustatiedot olisivat olemassa	Tiivis työskentely suunnittelijoiden kanssa, työkalujen beetatestaus