



**TURUN
YLIOPISTO**

HIUKASSUODIN- JA HIUKASSILOITINALGORITMIT SEKÄ NIIDEN
SOVELTAMINEN AOA-MENETELMÄÄN PERUSTUVASSA
BLUETOOTH-SISÄTILAPAUKANNUKSESSA

Lasse Rintakumpu

Pro gradu -tutkielma
Maaliskuu 2024

Tarkastajat:

Ohjaajan titteli (Prof./Dos./FT) ja nimi

Toisen tarkastajan titteli (Prof./Dos./FT) ja nimi

MATEMATIIKAN JA TILASTOTIETEEN LAITOS

Turun yliopiston laatujaarjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck-järjestelmällä

TURUN YLIOPISTO
Matematiikan ja tilastotieteen laitos

LASSE RINTAKUMPU: Hiukassuodin- ja hiukassiloitinalgoritmit sekä niiden soveltaminen AoA-menetelmään perustuvassa Bluetooth-sisätilapaikannuksessa
Pro gradu -tutkielma, X s.
Tilastotiede
Maaliskuu 2024

Tutkielmassa esitetään hiukassuodin- ja hiukassiloitinalgoritmien teoria Bayesilaisessa tilastotieteellisessä viitekehyksessä. Lisäksi tutkielmassa käsitellään hiukassuotimien varianssin estimointia.

Empiirisenä esimerkkinä tutkielmassa tarkastellaan hiukassuodin- ja hiukassiloitinalgoritmien käyttöä AoA-teknologiaan perustuvassa Bluetooth-sisätilapaikannusratkaisussa.

Asiasanat: SMC-menetelmät, Monte Carlo -menetelmät, sekventiaalinen Monte Carlo, suodinongelma, hiukassuodin, hiukassiloitin, SIR-algoritmi, sisätilapaikannus, BLE, AoA, triangulaatio, Bayesilainen päättely

Sisällys

1	Johdanto	3
1.1	Notaatioista	3
1.2	Suodinongelma	4
1.3	Suodin- ja siloitteluongelmien historiaa	5
1.4	Monte Carlo -menetelmistä	7
1.4.1	Monte Carlo -approksimaatio	8
1.4.2	Tärkeytysotanta	8
1.5	Bayesilainen suodin	9
1.6	Kalman-suotimen ja hiukassuotimen yhteydestä ja eroista	10
2	Hiukassuotimet	11
2.1	Saapasremmisuodin	11
2.2	SIS-algoritmi	11
2.3	SIR-algoritmi	11
2.3.1	Parametrien valinta	13
2.3.2	Konvergenssituloksia	16
2.3.3	Marginaalijakauma	16
2.3.4	Aikakompleksisuus	17
2.4	Varianssin estimointi	17
3	Hiukassilottimet	19
3.1	Offline-algoritmit	19
3.1.1	SIR-siloitin	19
3.2	Online-algoritmit	19
4	Hiukassuodin ja -siloitin sisätilapaikannuksessa	21
4.1	Koeasetelma	21
4.2	Datan kuvaus	23
4.3	Ongelman kuvaus	25
4.4	Algoritmin toteutuksesta	27
4.5	Parametrien valinta	27
4.6	Teknologian kuvaus	28
4.6.1	Kalibraatioalgoritmi	28
4.7	AoA-menetelmistä	28
4.7.1	MUSIC-algoritmi	28
4.8	Datan kuvaus	29

4.9	Ongelman ja mallien kuvaus	29
4.9.1	Datan valinta	29
4.9.2	Uskottavuusmallit	29
4.9.3	Dynaaminen malli	29
4.10	Algoritmien toteutuksesta	29
4.11	Karttasovitusalgoritmi	29
4.12	Reitinhakualgoritmi	29
4.13	Parametrien valinnasta	29
4.14	Tulokset	29
5	Lopuksi	31
6	(Liite) Other stuff {-}(followed by# A chapter‘)	33

Luku 1

Johdanto

Hiukassuotimet ovat joukko Monte Carlo -algoritmeja, joiden avulla voidaan ratkaista ns. suodinongelma, kun ongelma on epälineaarinen ja/tai ongelmaan liittyvä kohina ei noudata normaali jakaumaa. Hiukassuotimille on lukuisia sovellutuksia esimerkiksi Bayesilaisessa tilastotieteessä, fysiikassa ja robotiikassa.

Tämän tutkielman tavoitteena on esittää hiukassuotimien teoria sekä joitakin menetelmäperheeseen kuuluvia algoritmeja. Tutkielman ensimmäisessä luvussa kuvataan yleisellä tasolla sekä suodinongelma että sen ratkaisujen historiaa ja esitetään joitakin Monte Carlo -menetelmiin liittyviä yleisiä tuloksia sekä Bayesilainen viitekehys suodinongelmalle. Toisessa luvussa kuvataan kaksi hiukassuodinalgoritmia, saapasremmisuodin sekä SIR-algoritmi ja perehdytään hiukassuotimen varianssin estimointiin. Kolmannessa luvussa tarkastellaan suodinongelmaan läheisesti liittyvää siloitteluongelmaa ja esitetään hiukassiloitinalgoritmeja tämän ongelman ratkaisemiseksi. Neljäs luku keskittyy hiukassuotimen käyttöön empiirisessä AoA/Bluetooth-teknologiaan perustuvassa sisätilapaikannussovelluksessa. Tässä luvussa esitetään myös hiukassuodinalgoritmit radiosignaalin tulokulman arviointiin sekä radiovastaanottimen kalibrointiin. Lisäksi käsitellään lyhyesti sisätilapaikannuksessa hyödynnettävää karttasovitusalgoritmia.

Hiukassuodin- sekä hiukassiloitinalgoritmien osalta tutkielman esitykset seuraavat erityisesti Simo Särkän kirjaa *Bayesian Filtering and Smoothing* (2013) [8], Fredrik Gustafssonin artikkelia “Particle Filter Theory and Practice with Positioning Applications” (2010) [2] sekä Olivier Cappén, Simon J. Godsillin ja Eric Moulines’n artikkelia “An overview of existing methods and recent advances in sequential Monte Carlo” (2007) [7]. Hiukassuotimien varianssin estimointi seuraa artikkeleita TODO.

1.1 Notatioista

Tässä tutkielmassa käytetään seuraavia notaatioita. Vektoreita merkitään pienellä kirjaimella, esimerkiksi z . Hiukassuotimen hiukkaset sisältäviä vektoreita merkitään x_k^i , missä alaindeksi viittaa ajanhetkeen $k, k = \{1, \dots, T\}$ ja yläindeksi partikkeliin i , missä $i = \{1, \dots, N\}$. Ajanhetkien $k, k = \{1, \dots, T\}$ havainnot sisältäviä vektoreita merkitään $\{y_1, \dots, y_k\}$. Lähtökohtaisesti kaikki tutkielmassa esitetyt muuttujat ovat ylä- ja alaindeksejä lukuunottamattavektoreita. Skalaareihin pyritään viittaamaan

Taulukko 1.1: Lyhenteet ja symbolit

Lyhenne tai symboli	Selitys
SMC	<i>Sequential Monte Carlo</i> , sekventiaallinen Monte Carlo -menetelmä, synonyymi hiukassuotimelle

isoilla kirjaimilla, esimerkiksi Z . Milloin tämä ei ole mahdollista, selviää muuttujan skalaariarvoisuus asiayhtedestä. Prosesseihin viitataan alaindeksoidulla isolla kirjaimella, esimerkiksi X_k . Matriiseja merkitään isolla lihavoidulla kirjaimella, esimerkiksi \mathbf{X} ja funktiota TODO. Taulukossa 1.1 esitetään tutkielman keskeisimmät lyhenteet ja symbolit.

1.2 Suodinongelma

Stokastisten prosessien teoriassa suodinongelmaksi kutsutaan tilannetta, jossa halutaan muodostaa keskineliövirheen mielessä paras mahdollinen estimaatti jonkin järjestelmän tilan arvoille, kun ainoastaan osa tiloista voidaan havaita ja/tai havaintoihin liittyy kohinaa. Tavoitteena on toisin sanoen laskea jonkin prosessin posteriorijakauma kyseisten havaintojen perusteella. Ongelmaa havainnollistaa kaavio (1.1).

$$\begin{array}{ccccccc}
 x_1 & \longrightarrow & x_2 & \longrightarrow & x_3 & \longrightarrow & \dots & \text{piilossa olevat tilat} \\
 \downarrow & & \downarrow & & \downarrow & & & \\
 y_1 & & y_2 & & y_3 & & \dots & \text{havainnot}
 \end{array} \tag{1.1}$$

Tässä tutkielmassa keskitytään erityisesti epälineaarisen, ns. Markovin piilomallin posteriorijakauman Bayesilaiseen ratkaisuun. Ongelmassa tiedetään, miten havaitut muuttujat y_k kytkeytyvät ”piilossa oleviin” tilamuuttujiin x_k sekä osataan sanoa jotain tilamuuttujien todennäköisyyksistä. Oletetaan myös, että piilossa oleville tiloille X_k pätee Markov-ominaisuus, jolloin kutakin hetkeä seuraava tila x_{k+1} riippuu menneistä tiloista $x_{1:k}$ ainoastaan tilan x_k välityksellä. Lisäksi havaittu tila y_k riippuu tiloista x_k ainoastaan jonkin x_k :n funktion kautta. Kun aika-avaruus on diskreetti ja ajanhetkellä $k = \{1, \dots, t\}$ piilossa olevan prosessin tilaa merkitään x_k ja havaittua prosessia y_k , saadaan mallit

$$x_{k+1} = f(x_k, \nu_k), \tag{1.2}$$

$$y_k = h(x_k) + e_k. \tag{1.3}$$

Lisäksi tiedetään prosessin alkuhetken jakauma $x_0 \sim p_{x_0}$, tähän liittyvän kohinaprosessin jakauma $\nu_k \sim p_{\nu_k}$ sekä malliin y_k liittyvä kohina $e_k \sim p_{e_k}$. Koska hiukassuodinalgoritmit pyrkivät ratkaisemaan juurikin epälineaarisen, ei-Gaussisen

suodinongelman, voivat funktiot $f(\cdot)$ ja $h(\cdot)$ olla epälineaarisia eikä kohinan tarvitse olla normaalijakautunutta.

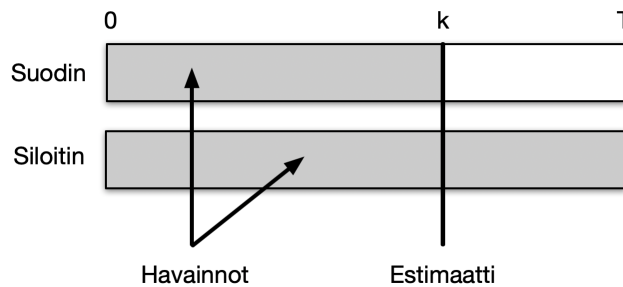
Mallit voidaan esittää myös yleisemmässä jakaumamuodossa

$$x_{k+1} \sim p(x_{k+1}|x_k), \quad (1.4)$$

$$y_k \sim p(y_k|x_k). \quad (1.5)$$

Tutkielman teoriaosassa käytetään ensisijaisesti yhtälöiden (1.4) ja (1.5) muotoilua. Empiirisessä osassa palataan yhtälöiden (1.2) ja (1.3) muotoiluun.

Suodinongelmaa lähellä on myös ns. siloitteluongelma (*smoothing problem*), jossa ollaan kiinnostuneita prosessin x_k posteriorijakaumasta $p(x_k|y_k)$ jokaisena ajanhetkenä $\{1, \dots, k\}$ ei ainoastaan haluttuna ajanhetkenä k . Hiukassuodinalgoritmit näyttävät ratkaisevan siloitteluongelman ilmaiseksi, mutta tähän liittyy kuitenkin joidenkin mallien kohdalla mahdollista epätarkkuutta, joten tarvittaessa tasoitusongelma pitää ratkaista erikseen. Tähän ongelmaan palataan tutkielman luvussa 3. Kuva 1.1 selittää suodin- ja siloitteluongelmien eron.



Kuva 1.1: Suodin- ja siloitteluongelma

1.3 Suodin- ja siloitteluongelmien historiaa

Tämä alaluku esittää pääpiirteittään suodinongelmalle esitettyjen ratkaisujen historian. Lineaarisen suodinongelman osalta alaluku noudattaa Dan Crisanin artikkelia “The stochastic filtering problem: a brief historical account” (2014) [1] sekä Mohinder S. Grewalin ja Angus P. Andrewsin artikkelia “Applications of Kalman Filtering in Aerospace 1960 to the Present” (2010) [4]. Hiukassuotimien osalta lähteenä toimii Cappé & al (2007) [7].

Suodinongelma nousi esille insinööritieteiden sekä sotateollisuuden käytännön ongelmista 2. maailmansodan aikana, vaikkakin suodinongelman diskreetin ajan ratkaisut juontavat jo Andrei N. Kolmogorovin 30-luvun artikkeleihin. Jatkuvan ajan tilanteessa ensimmäisen optimaalisen, kohinan sallivan suotimen esitti matemaatikko, kybernetiikan kehittäjä Norbert Wiener. Wiener-suotimenä tunnettua ratkaisuaan varten Wiener muotoili seuraavat kolme ominaisuutta, jotka prosessin X estimaatin \hat{X}_t pitää toteuttaa.

1. *Kausaliteetti*: X_t tulee estimoida käyttäen arvoja Y_s , missä $s \leq t$.
2. *Optimaalisuus*: X_t :n estimaatin \hat{X}_t tulee minimoida keskineliövirhe $\mathbb{E}[(X - \hat{X}_t)^2]$.
3. *On-line -estimointi*: Estimaatin \hat{X}_t tulee olla saatavissa minä hyvänsä ajanhetkenä t .

Wiener sovelsi ratkaisussaan stationaaristen prosessien spektriteoriaa. Tulokset julkaistiin salaisina Yhdysvaltojen asevoimien tutkimuksesta vastanneen National Defense Research Committeeen (NDRC) raportissa vuonna 1942. Tutkimus tunnettiin sodan aikana lempinimellä ”Keltainen vaara” sekä painopaperinsa värin että vaikeaselkoisuutensa vuoksi. Myöhemmin Wiener esitti tuloksensa julkisesti kirjassaan *Extrapolation, Interpolation and Smoothing of Stationary Time Series* (1949). Wienerin alkuperäiset kolme perusperiaatetta pätevät edelleen kaikille suodinongelman ratkaisuille, myös hiukassuotimille.

Kenties tärkein ja varmasti tunnetuin lineaariseen suodinongelman ratkaisu on Kalman-suodin. Suotimen kehittivät R.E. Kalman ja R.S. Bucy 1950- ja 60-lukujen taitteessa Yhdysvaltain kylmän sodan kilpavarustelutarpeisiin perustetussa Research Institute for Advanced Studies -tutkimuslaitoksessa (RIAS). Kalman-suodin on suodinongelman diskreetin ajan ratkaisu, kun taas Kalman-Bucy-suodin on jatkuvan ajan ratkaisu. Kohinan ollessa normaalijakautunutta on Kalman-suodin Wiener-suotimen tavoin lineaarisen suodinongelman optimaalinen ratkaisu. Wiener-suotimella ja Kalman-suotimella on kuitenkin erilaiset oletukset, minkä vuoksi erityisesti säätö- ja paikannussovelluksissa Kalman-suotimen käyttö on luontevampaa. Suotimien oletuksia ja oletusten välisiä eroja ei käsitellä tässä tutkielmassa, mutta alaluvussa 1.6. TODO:LINKKI käsitellään Kalman-suotimen formaalia yhteyttä hiukassuotimiin.

Kalman-suodinta voidaan soveltaa myös epälineaarisisessa tapauksessa, kunhan suodinongelman funktiot $f(\cdot)$ ja $h(\cdot)$ ovat derivoituvia ja niihin liittyvä kohina oletetaan normaalijakautuneeksi. Tätä ratkaisua kutsutaan laajennetuksi Kalman-suotimeksi (extended Kalman filter, EKF). Suodin kehitettiin 60-luvulla NASA:n Apollo-ohjelman tarpeisiin, vaikkakin itse avaruusalusten laitteistot hyödynsivät lentoratojen laskennassa Kalman-suotimen perusversiota. Laajennetun Kalman-suotimen toimintaperiaate perustuu epälineaaristen funktioiden linearisointiin Taylorin kehitelmän avulla kulloisenkin estimaatin ympärillä. Laajennettu Kalman-suodin on erityisesti paikannussovellusten *de facto* -suodinstandardi, mutta suodin ei kuitenkaan ole epälineaarisen ongelman optimaalinen estimaattori.

Kalman-suotimesta on lisäksi olemassa lukuisia muita epälineaarisiiin ongelmiin soveltuvia laajennuksia, muun muassa paikkaratkaisun Kalman-suodin (*position Kalman filter*, PKF), hajustamaton Kalman-suodin (*unscented Kalman filter*, UKF) sekä tilastollisesti linearisoitu Kalman-suodin (*statistically linearized Kalman filter*, SLF). Kuitenkin jos prosessin X mallia ei tunneta tarkasti tai kohinaa ei voida olettaa normaalijakautuneeksi, ovat hiukassuotimet eli sekventiaaliset Monte Carlo -menetelmät Kalman-suotimen johdannaisia parempia ratkaisuja. Vaikka tila-avaruuden dimensioiden kasvaessa kasvaa myös hiukassuotimien vaatima laskentateho, ovat hiukassuotimet aina sitä parempia mitä epälineaarisempia mallit ovat ja mitä kauempana normaalijakaumasta kohina on. Viimeisten vuosikymmenten

aikana myös laskennan teho on kasvanut merkittävästi samalla kun laskennan hinta on vastaavasti romahtanut, mikä puoltaa Monte Carlo -menetelmien käyttöä entistä useammissa ongelmissa.

Joitakin suodinongelman rekursiivisia Monte Carlo -ratkaisuja löytyy jo 1950–70-luvuilta, erityisesti säätöteoriaan piiristä. Olennainen nykyalgoritmeihin periytynyt oivallus varhaisissa suodinalgoritmeissa oli tärkeytysotannan käyttö halutun jakaumasta p ja teemme sen sijaan havaintoja p ja teemme sen sijaan havaintoja q , jota painotamme niin, että tuloksena saadaan jakauman p harhaston estimaatti. Algoritmi on kuvattu tarkemmin tutkielman alaluvussa 2.

Tärkeytysotantaa käyttävä suodinongelman ratkaiseva SIS-algoritmi (sequential importance sampling) ei kuitenkaan vielä 70-luvulla löytänyt suurta käytännön suosiota. Osin tämä johtui puutteellisesta laskentatehosta, mutta algoritmi kärsi myös otosten ehtymisenä (sample impoverishment) tunnetusta ongelmasta. Monissa ongelmissa SIS-algoritmia käytettäessä suuri osa painoista päättyy vain tietyille partikkeleille, jolloin vastaavasti suuri osa partikkeleista ei enää estimoi haluttua jakaumaa. Tähän ongelmaan palataan myöhemmin.

Merkittävän ratkaisun ehtymisongelmaan esittivät Gordon, Salmond ja Smith artikkelissaan “Novel approach to nonlinear/non-Gaussian Bayesian state estimation” (1993). [6] Artikkelin ratkaisu kulki nimellä *bootstrap filter*, saapasremmisuodin. Saapasremmisuodin vältti ehtymisen uudelleenotannalla, jossa matalapainoiset partikkelit korvattiin otoksilla korkeapainoisemmista partikkeleista. Ratkaisussa painot eivät myöskään riippuneet partikkelien aiemmista poluista vaan ainoastaan havaintojen uskottavuusfunktioista. Vastaavaa ratkaisua käytetään tämän tutkielman uudemmas-
sa SIR-algoritmissa (*sampling importance resampling*), jossa myös uudelleenotantaan sovelletaan tärkeytysotantaa.

Sekventiaalisissa Monte Carlo -menetelmissä stokastisen prosessin posteriorijakauman esittämiseen käytettyjä otoksia kutsutaan myös partikkeleiksi ja menetelmiä siten hiukassuotimiksi. Erityisesti myöhemmin esitettävää SIR-algoritmia kutsutaan usein hiukassuotimeksi. Termiä hiukassuodin käytti ensimmäisen kerran Del Moral artikkelissa “Nonlinear Filtering: Interacting Particle Resolution” (1996) [5], SMC-menetelmät termiä Liu ja Chen artikkelissa “Sequential Monte Carlo Methods for Dynamic Systems” (1998) [3]. Tässä tutkielmassa käytetään yleisemmin käytettyä termiä hiukassuotimet.

1.4 Monte Carlo -menetelmistä

Tässä alaluvussa kuvataan lyhyesti hiukassuotimissa käytettävien Monte Carlo -menetelmien peruseriaate todennäköisyysjakauman estimoinnissa. Lisäksi esitetään tärkeytysotanta-algoritmi (*importance sampling*), jonka tarkoituksena on estimoida harhattomasti jakaumaa $p(x|y_{1:k})$, josta emme voi suoraan tehdä otoksia, mutta jota voimme approksimoida toisella jakaumalla q . Esitykset noudattavat Särkkää (2013) [8].

1.4.1 Monte Carlo -approksimaatio

Bayesilaisessa päättelyssä ollaan yleisesti kiinnostuttu laskemaan johonkin posteriori-
tiheysjakaumaan p liittyvää odotusarvoa

$$\mathbb{E}[g(x)|y_{1:k}] = \int g(x)p(x|y_{1:k})dx, \quad (1.6)$$

missä g on tila-avaruuden mielivaltainen funktio ja $p(x|y_{1:t})$ on havaintoihin $\{y_1, \dots, y_k\}$ liittyvä x :n posterioritiheysjakauma. Odotusarvo on laskettavissa suljetussa muodossa vain harvoissa tapauksissa, suodinongelman kohdalla silloin, kun kyseessä on lineaarinen ja Gaussinen malli. Odotusarvoa voidaan kuitenkin approksimoida niin sanoituilla Monte Carlo -menetelmillä. Menetelmien peruseriaa-
te on tehdä riippumattomia otoksia estimoitavasta jakaumasta ja laskea haluttu odotusarvo otosten avulla. Jos tehdään N otosta jakaumasta $x^i \sim p(x|y_{1:t})$, missä $i = \{1, \dots, N\}$ saadaan näiden otosten avulla laskettua odotusarvon estimaatti

$$\mathbb{E}[g(x)|y_{1:k}] \simeq \frac{1}{N} \sum_{i=1}^N g(x^i). \quad (1.7)$$

Monte Carlo -estimaatti konvergoi keskeisen raja-arvolauseen nojalla ja sen estimointivirheen voidaan osoittaa olevan luokkaa $O(\frac{1}{\sqrt{N}})$ riippumatta tilamuuttujan x dimensiosta. Hiukassuotimet hyödyntävät Monte Carlo -estimointia sekventiaalisesti, jolloin estimaatti lasketaan rekursiivisesti kullekin ajanhetkelle $k = \{1, \dots, t\}$. Tähän palataan luvuissa 3 ja 4.

1.4.2 Tärkeytysotanta

Tilanteessa, jossa Monte Carlo -otoksia ei voida tehdä suoraan jakaumasta p , voidaan hyödyntää jakaumaa p approksimoivaa tärkeytys- tai ehdotusjakaumaa $q(x|y_{1:k})$ sekä ns. tärkeytysotantaa. Oletetaan, että tunnetaan priorijakauma $p(x)$ ja on olemassa havaintomalli $p(y_{1:k}|x)$ sekä valittu ehdotusjakauma $q(x|y_{1:k})$, josta voidaan tehdä otoksia. Ehdotusjakaumalta edellytetään lisäksi, että sen kantaja on suurempi tai yhtä suuri kuin jakauman $p(x|y_{1:k})$ ja että se saa nollasta poikkeavia arvoja kaikkialla missä $p(x|y_{1:k})$ saa nollasta poikkeavia arvoja. Kirjoitetaan halutun posteriorijakauman odotusarvo integraalina

$$\int g(x)p(x|y_{1:k})dx = \int g(x)\frac{p(x|y_{1:k})}{q(x|y_{1:k})}q(x|y_{1:k})dx, \quad (1.8)$$

jolle voidaan muodostaa Monte Carlo -approksimaatio tekemällä N otosta jakaumasta $x^i \sim q(x|y_{1:k})$.

Muodostetaan näin odotusarvo

$$\mathbb{E}[g(x)|y_{1:k}] \simeq \frac{1}{N} \sum_{i=1}^N \frac{p(x^i|y_{1:k})}{q(x^i|y_{1:k})} g(x^i) = \sum_{i=1}^N w^i g(x^i), \quad (1.9)$$

missä $g(x)$ on jokin estimoinnissa hyödyllinen, mielivaltainen funktio. Tutkielmassa käytetty notaatio x_k^i viittaa ajanhetken k partikkeliin i , missä $i = \{1, \dots, N\}$. Tärkeytysotantaa kuvaa nyt algoritmi (1). Kun posteriorijakauman estimaatti muodostetaan kyseisellä algoritmilla voidaan tulos kirjoittaa

$$\hat{p}(x|y_{1:k}) = \sum_{i=1}^N w^i \delta(x - x^i), \quad (1.10)$$

missä $\delta(x)$ on Diracin deltafunktio.

Algoritmi 1: Tärkeytysotanta

```

begin
  for  $i = 1, 2, \dots, N$  do
    begin
      Otetaan  $N$  otosta ehdotusjakaumasta  $x^i \sim q(x|y_{1:k})$ .
    begin
      Lasketaan normalisoimattomat painot  $w_*^i = p(y_{1:k}|x^i)p(x^i)/q(x^i|y_{1:k})$ .
      ja normalisoidut painot  $w^i = w_*^i / \sum_{j=1}^N w_*^j$ .
    begin
      Estimoidaan  $p$  laskemalla tiheydelle approksimaatio
       $\mathbb{E}[g(x)|y_{1:k}] \simeq \sum_{i=1}^N w^i g(x^i)$ .
    end
  end
end

```

1.5 Bayesilainen suodin

Suodinongelmassa ollaan kiinnostuttu tilavektorin posteriorijakauman $p(x_k|y_{1:k})$ estimoinnista. Tässä alaluvussa käydään läpi yleinen rekursiivinen, Bayesilainen posteriorijakauman laskenta. Tällaista suodinongelman ratkaisua kutsutaan myös Bayesilaiseksi suotimeksi. Koska epälineaarisessa, ei-normaalijakautuneessa tilanteessa rekursiota ei voida laskea analyyttisesti, pitää estimoinnissa käyttää numeerisia menetelmiä. SMC-menetelmissä tämä tarkoittaa jakauman sekventiaalista Monte Carlo -approksimointia, jonka toteutus esitetään alaluvun 4 algoritmissa. Molemmat esitykset noudattavat Gustafssonia (2010).

Bayesilainen ratkaisu tilavektorin posteriorijakauman estimaatille $\hat{p}(x_k|y_{1:k})$ saadaan seuraavalla rekursiolla (käydään läpi jokaiselle ajanhetkelle $k = \{1, \dots, t\}$). Lasketaan ensin

$$p(x_k|y_{1:k}) = \frac{p(y_k|x_k)p(x_k|y_{1:k-1})}{p(y_k|y_{1:k-1})}, \quad (1.11)$$

joka saadaan suoraan Bayesin kaavasta $P(A|B) = P(B|A)P(A)/P(B)$. Normalisointivakio lasketaan integraalina

$$p(y_k|y_{1:k-1}) = \int_{\mathbb{R}^{n_x}} p(y_k|x_k)p(x_k|y_{1:k-1}) dx_k, \quad (1.12)$$

joka saadaan kokonaistodennäköisyyskaavasta $P(A) = \mathbb{E}[P(A|X)] = \int_{-\infty}^{\infty} P(A|X=x) f_X(x) dx$. Merkintä R^{n_x} vastaa tässä piilossa olevan tilavektorin dimensiota n .

Lopuksi lasketaan päivitysaskel ajalle, joka saadaan edelleen kokonaistodennäköisyydellä

$$p(x_{k+1}|y_{1:k}) = \int_{\mathbb{R}^{n_x}} p(x_{k+1}|x_k) p(x_k|y_{1:k}) dx_k . \quad (1.13)$$

Rekursioin avulla voimme laskea jakauman $p(x_k|y_{1:k})$ estimaatti käymällä rekursion läpi k kertaa.

1.6 Kalman-suotimen ja hiukassuotimen yhteydestä ja eroista

TODO

Luku 2

Hiukassuotimet

2.1 Saapasremmisuodin

2.2 SIS-algoritmi

2.3 SIR-algoritmi

Tässä alaluvussa esitetään SMC-menetelmiin kuuluva SIR-algoritmi, epälineaarisen suodinongelman ratkaisemiseksi. Algoritmi on numeerinen toteutus alaluvussa 3 kuvatussa Bayesilaisesta suotimesta. Esitetty algoritmi perustuu Gustafssoniin (2010).

Algoritmi alustetaan jakaumasta $x_1^i \sim p_{x_0}$ generoiduilla N -kappaleella partikkeleita. Jokaiselle partikkelille annetaan alustuksessa sama paino $w_{1|0}^i = 1/N$. Algoritmi suoritetaan jokaiselle partikkelille $i = \{1, 2, \dots, N\}$ jokaisella ajanhetkellä $k = \{1, 2, \dots, t\}$.

Seuraava toistetaan jokaiselle ajanhetkelle $k = \{1, 2, \dots, t\}$. Algoritmin ensimmäisessä vaiheessa päivitetään painot yhtälön (2.1) mukaan.

$$w_{k|k}^i = \frac{1}{c_k} w_{k|k-1}^i p(y_k | x_k^i). \quad (2.1)$$

Tämä vastaa yllä esitetyn Bayes-suotimen päivitysvaihetta (1.11). Normalisointipaino c_k lasketaan puolestaan yhtälöstä (2.2), mikä vastaa Bayes-suotimen normalisointivaiheen laskemista (1.12) ja asettaa painojen summaksi $\sum_{i=1}^N w_{k|k}^i = 1$.

$$c_k = \sum_{i=1}^N w_{k|k-1}^i p(y_k | x_k^i). \quad (2.2)$$

Seuraavassa vaiheessa estimoidaan p laskemalla tiheyden $p(x_{1:k} | y_{1:k})$ Monte Carlo -estimaatti yhtälön (2.3) perusteella

$$\hat{p}(x_{1:k} | y_{1:k}) = \sum_{i=1}^N w_{k|k}^i \delta(x_{1:k} - x_{1:k}^i). \quad (2.3)$$

Tämän jälkeen suoritetaan valinnainen uudelleenotanta. Uudelleenotanta voidaan tehdä jokaisella askeleella tai efektiivisen otoskoon perusteella alla kuvatun kynnyksarvoehdon $\hat{N}_{eff} < N_{th}$ täyttyessä, jolloin uudelleenotantaa kutsutaan adaptiiviseksi uudelleenotannaksi. Tällaista uudelleenotantaa hyödynnetään esitetyssä algoritmissa (2). Uudelleenotantaa tarkastellaan lähemmin alaluvussa 4.1.2. Lopuksi päivitetään aika (jos $k < t$), luodaan uudet ennusteet partikkeleille ehdotusjakaumasta (2.4)

$$x_{k+1}^i \sim q(x_{k+1} | x_k^i, y_{k+1}) \quad (2.4)$$

ja päivitetään partikkelien painot tärkeytysotannalla (2.5), sen mukaan kuinka todennäköisiä partikkelien ennusteet ovat

$$w_{k+1|k}^i = w_{k|k}^i \frac{p(x_{k+1}^i | x_k^i)}{q(x_{k+1}^i | x_k^i, y_{k+1})}. \quad (2.5)$$

Vaiheet 2.4 ja 2.5 vastaavat Bayes-suotimen aikapäivitystä (1.13).

Alla käsitellään algoritmiin liittyvän uudelleenotantamenetelmän, partikkelien määrän ja ehdotusjakauman valinta. Lopuksi esitetään algoritmin konvergenssia, marginaalijakaumaa sekä aikakompleksisuutta koskevia tuloksia.

Algoritmi 2: SIR

Result: Posteriorijakauman $p(x_{1:k}|y_{1:k})$ estimaatti.

Data: Havainnot y_k . Generoitu $x_1^i \sim p_{x_0}$ missä $i = \{1, \dots, N\}$ ja jokainen partikkeli saa saman painon $w_{1|0}^i = 1/N$.

```
begin
  for  $k = \{1, 2, \dots, t\}$  do
    for  $i = \{1, 2, \dots, N\}$  do
      begin
        Päivitetään painot  $w_{k|k}$ .
      begin
        Estimoidaan  $p$  laskemalla tiheydelle approksimaatio
         $\hat{p}(x_{1:k}|y_{1:k}) = \sum_{i=1}^N w_{k|k}^i \delta(x_{1:k} - x_{1:k}^i)$ .
      begin
        Lasketaan efektiivinen otoskoko  $\hat{N}_{eff}$ .
      if  $\hat{N}_{eff} < N_{th}$  then
        begin
          Otetaan uudet  $N$  otosta palauttaen joukosta  $\{x_{1:k}^i\}_{i=1}^N$ , missä
          otoksen  $i$  todennäköisyys on  $w_{k|k}^i$ .
        begin
          Asetetaan painot  $w_{k|k}^i = 1/N$ .
        if  $k < t$  then
          begin
            Aikapäivitys.
            Luodaan ennusteet partikkeleille ehdotusjakaumasta
             $x_{k+1}^i \sim q(x_{k+1}|x_k^i, y_{k+1})$ ,
            päivitetään partikkelien painot tärkeytysotannalla.
```

2.3.1 Parametrien valinta

Ennen algoritmin suorittamista valitaan ehdotusjakauma $q(x_{k+1}|x_{1:k}, y_{k+1})$, uudelleenotantamenetelmä sekä partikkelien määrä N . Ehdotusjakauman ja uudelleenotantamenetelmän valinnassa tärkeimpänä päämääränä on välttää otosten ehtymistä, kun taas partikkelien määrä säätelee kompromissia algoritmin suorituskyvyn ja tarkkuuden välillä.

2.3.1.1 Otoksoon N valinta

Yleispätevää sääntöä otoksoon/partikkelien lukumäärän N valinnalle on vaikeaa antaa, sillä vaadittava estimointitarkkuus riippuu usein käsillä olevasta ongelmasta. Gordon & al. (1993) esittävät kuitenkin kolme tekijää, jotka vaikuttavat partikkelien lukumäärän valintaan

- a. tila-avaruuden ulottuvuuksien lukumäärä n_x ,

- b. tyypillinen päällekkäisyys priorin ja uskottavuuden välillä
- c. sekä tarvittava aika-askelten lukumäärä.

Ensimmäisen tekijän vaikutus on selvä. Mitä useammassa ulottuvuudessa otantaa tarvitsee tehdä, sen korkeammaksi on N asetettava, jotta jokainen ulottuvuus pystytään kattamaan. Tekijät (b) ja (c) puolestaan seuraavat uudelleenotannasta. Jos se osa tila-avaruutta, jossa uskottavuus $p(y_k|x_k)$ saa merkittäviä arvoja on pieni verrattuna siihen osaan, jossa priorijakauma $p(x_k|y_{1:k-1})$ saa merkittäviä arvoja, suuri osa partikkeleista saa pieniä painoja eikä näin valikoidu uudelleenotantaan.

Yleisesti ottaen N kannattaa asettaa sellaiseksi, että se paitsi tuottaa riittävän tarkan estimaatin, on se käytettävissä olevan laskentatehon sekä vaadittavan laskentanopeuden kannalta järkevää. Tähän palataan tutkielman lopuksi empiirisessä paikannusesimerkissä.

2.3.1.2 Uudelleenotantamenetelmän valinta

Ilman uudelleenotantaa on mahdollista, että algoritmi alkaa kärsiä SIS-algoritmillemme ominaisesta otosten ehtymisestä. Toisin sanoen kaikki painot alkavat keskittyä vain muutamalle partikkelille eikä algoritmi enää approksimoi tehokkaasti haluttua jakaumaa. Uudelleenotanta tarjoaa osittaisen ratkaisun tähän ongelmaan, mutta hävittää samalla informaatiota ja siten lisää satunnaisotantaan liittyvää epävarmuutta. Yleisesti ottaen uudelleenotanta kannattaa aloittaa vasta siinä vaiheessa algoritmin suorittamista, kun siitä on otosten ehtymisen kannalta hyötyä, esimerkiksi efektiivisen otoskoon pudottua jonkin kynnysarvon alapuolelle (adaptiivinen uudelleenotanta). Efektiivinen otoskoko saadaan laskettua variaatiokertoimesta c_ν kaavalla

$$N_{eff} = \frac{N}{1 + c_\nu^2(w_{k|k}^i)} = \frac{N}{1 + \frac{\text{Var}(w_{k|k}^i)}{(\mathbb{E}[w_{k|k}^i])^2}} = \frac{N}{1 + N^2 \text{Var}(w_{k|k}^i)}. \quad (2.6)$$

Näin laskettu efektiivinen otoskoko maksimoituu ($N_{eff} = N$), kun kaikille painoille pätee $w_{k|k}^i = 1/N$ ja minimoituu ($N_{eff} = 1$), kun $w_{k|k}^i = 1$ todennäköisyydellä $1/N$ ja $w_{k|k}^i = 0$ todennäköisyydellä $(N-1)/N$. Normalisoitujen painojen avulla saadaan efektiiviselle otoskoolle ajanhetkellä k laskennallinen approksimaatio

$$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^N (w_{k|k}^i)^2}. \quad (2.7)$$

Sekä määritelmälle (2.6) että (2.7) pätee $1 \leq \hat{N}_{eff} \leq N$. Yläraja saavutetaan, kun jokaisen partikkelin paino on sama. Alarajalle päädytään, kun kaikki paino keskittyy yksittäiselle partikkelille. Tästä saadaan määriteltyä algoritmillemme SIR-uudelleenotantaehto $\hat{N}_{eff} < N_{th}$. Gustafsson (2010) esittää uudelleenotannan kynnysarvoksi esimerkiksi $\hat{N}_{th} = 2N/3$.

Uudelleenotanta ei muuta approksimoitavan jakauma p odotusarvoa, mutta se lisää jakauman Monte Carlo -varianssia. On kuitenkin olemassa esimerkiksi osittamiseen perustuvia uudelleenotantamenetelmiä, jotka pyrkivät minimoimaan varianssin lisäyksen. Varianssin pienennysmenetelmät jätetään tämän tutkielman ulkopuolelle.

2.3.1.3 Ehdotusjakauman valinta

Yksinkertaisin muoto ehdotusjakaumalle on $q(x_{1:k}|y_{1:k})$ eli jokaisella algoritmin suorituskerralla käydään läpi koko aikapolku $1 : k$. Tämä ei kuitenkaan ole tarkoituksenmukaista, erityisesti jos kyseessä on reaaliaikainen sovellutus. Kirjoitetaan ehdotusjakauma muodossa

$$q(x_{1:k}|y_{1:k}) = q(x_k|x_{1:k-1}, y_{1:k})q(x_{1:k-1}|y_{1:k}). \quad (2.8)$$

Jos yhtälöstä (2.8) poimitaan ehdotusjakaumaksi ainoastaan termi $q(x_k|x_{1:k-1}, y_{1:k})$ voidaan tämä kirjoittaa edelleen Markov-ominaisuuden nojalla muotoon $q(x_k|x_{k-1}, y_k)$. Tämä on suodinongelman kannalta riittävää, koska olemme kiinnostuneita posteriorijakaumasta ja arvosta x ainoastaan ajanhetkellä k (tasoitusongelmassa tarvitsisimme koko polun $x_{1:k}$). Alla tarkastellaan edelleen Gustafssonia (2010) seuraten kahta ehdotusjakauman valintatapaa, prioriotantaa (prior sampling) sekä uskottavuusotantaa (likelihood sampling).

Ennen ehdotusjakauman tarkastelua määritellään mallille signaali-kohinasuhde uskottavuuden maksimin ja priorin maksimin välisenä suhteena

$$\text{SNR} \propto \frac{\max_{x_k} p(y_k|x_k)}{\max_{x_k} p(x_k|x_{k-1})}. \quad (2.9)$$

Yhdistetään lisäksi ehdotusjakaumia varten yhtälöt (2.1) ja (2.2), jolloin saadaan painojen päivitys muotoon

$$w_{k|k}^i \propto w_{k-1|k-1}^i \frac{p(y_k|x_k^i)p(x_k|x_{k-1}^{k-1})}{q(x_k|x_{k-1}^i, y_k)}. \quad (2.10)$$

Kun suhde (2.9) on matala, on prioriotanta luonnollinen valinta. Tässä käytetään ehdotusjakaumana tilavektorin ehdollista prioria eli

$$q(x_k|x_{1:k-1}, y_k) = p(x_k|x_{k-1}^i). \quad (2.11)$$

Yhtälön (2.11) perusteella saadaan edelleen prioriotannan painoiksi

$$w_{k|k}^i = w_{k-1|k-1}^i p(y_k|x_k^i) = w_{k-1|k-1}^i p(y_k|x_{k-1}^i). \quad (2.12)$$

Kun signaali-kohinasuhde on kohtalainen tai korkea, on parempi käyttää ehdotusjakaumana skaalattua uskottavuusfunktioita (2.14). Tarkastellaan ensin tekijöihin jakoa

$$p(x_k|x_{k-1}^i, y_k) = p(y_k|x_k) \frac{p(x_k|x_{k-1}^i)}{p(y_k|x_{k-1}^i)}. \quad (2.13)$$

Kun SNR on korkea ja uskottavuusfunktio on integroitava pätee $p(x_k|x_{k-1}^i, y_k) \propto p(y_k|x_k)$, jolloin voidaan asettaa (2.14)

$$q(x_k|x_{k-1}^i, y_k) \propto p(y_k|x_k). \quad (2.14)$$

Yhtälön (2.14) perusteella saadaan edelleen uskottavuusotannan painoiksi (2.15).

$$w_{k|k}^i = w_{k-1|k-1}^i p(x_k^i|x_{k-1}^i). \quad (2.15)$$

2.3.2 Konvergenssituloksia

Alla esitetään kaksi SIR-algoritmiin liittyvää konvergenssitulosta. Se, kuinka hyvin esitetyllä algoritmilla arvioitu posterioritiheys $\hat{p}(x_{1:k}|y_{1:k})$ approksimoi todellista tiheysfunktioita $p(x_{1:k}|y_{1:k})$ sekä mikä on approksimaation keskineliövirhe. Tulokset noudattavat Crisanin ja Doucet'n artikkeleita "Convergence of Sequential Monte Carlo Methods" (2000) ja "A Survey of Convergence Results on Particle Filtering Methods for Practitioners" (2002).

Konvergenssitulos 1: Kun $N \rightarrow \infty$ algoritmille pätee $\forall k$ tulos (2.16).

$$\hat{p}(x_{1:k}|y_{1:k}) \xrightarrow{a.s.} p(x_{1:k}|y_{1:k}). \quad (2.16)$$

Konvergenssitulos 2: Keskineliövirheelle pätee asymptoottinen konvergenssi (2.17).

$$\mathbb{E}(\hat{g}(x_k) - \mathbb{E}(g(x_k)))^2 \leq \frac{p_k \|g(x_k)\|}{N}, \quad (2.17)$$

missä g on mikä hyvänsä piilossa olevan tila-avaruuden rajoitettu Borel-mitallinen funktio ($g \in \mathcal{B}(\mathbb{R}^{n_x})$), $\|g(\cdot)\|$ kyseisen funktion supremum-normi ja p_k jokin äärellinen vakio, jolle pätee ajanhetkestä k riippumatta $p_k = p < \infty$. Konvergenssituloksia ei tämän tutkielman puitteissa todisteta.

2.3.3 Marginaalijakauma

Edellä kuvattu algoritmi 1 tuottaa approksimaation koko prosessin posteriorijakaumalle $p(x_{1:k}|y_{1:k})$. Jos halutaan tietää ainoastaan posteriorijakauman $p(x_k|y_{1:k})$ estimaatti, voidaan käyttää yksinkertaisesti viimeisestä tilasta x_k laskettua estimaattia

$$\hat{p}(x_k|y_{1:k}) = \sum_{i=1}^N w_{k|k}^i \delta(x_k - x_k^i). \quad (2.18)$$

Toinen, tarkempi vaihtoehto on käyttää laskennassa tärkeytyspainoa

$$w_{k+1|k}^i = \frac{\sum_{j=1}^N w_{k|k}^j p(x_{k+1}^i | x_k^j)}{q(x_{k+1}^i | x_k^i, y_{k+1})} \quad (2.19)$$

painon (2.5) sijaan. Tällöin jokaisella aikapäivitysaskeleella lasketaan painot kaikkien mahdollisten tila-aika-avaruuspolkujen yli. Samoin kuin uudelleenotanta tämä pienentää painojen varianssia.

2.3.4 Aikakompleksisuus

Algoritmin perusmuodon aikakompleksisuus on $\mathcal{O}(N)$. Uudelleenotantamenetelmän tai ehdotusjakauman valinta ei suoraan vaikuta aikakompleksisuuteen. Sen sijaan marginalisointi tärkeytyspainolla (2.19) lisää algoritmin aikakompleksisuutta $\mathcal{O}(N) \rightarrow \mathcal{O}(N^2)$, koska jokaisen partikkelin kohdalla painot lasketaan jokaisen tila-aika-avaruuspolun yli. On selvää, että erityisesti isoilla otoskoon N arvoilla ei yllä esitetty marginalisointi enää ole mielekäästä.

Tällaisia tilanteita varten algoritmista on olemassa $\mathcal{O}(N \log(N))$ -versioita, jotka perustuvat esimerkiksi N :n kappaleen oppimiseen (N-body learning). Näiden algoritmien käsittely jää tämän tutkielman ulkopuolelle, mutta katsauksen algoritmeista ovat esittäneet esimerkiksi Klaas & al. artikkelissa “Toward Practical N^2 Monte Carlo: the Marginal Particle Filter” (2012).

§

2.4 Varianssin estimointi

Varianssin estimoinnissa ollaan kiinnostuneita XXX. Yksinkertaisin tapa estimoida hiukassuodinalgoritmin varianssia on yksinkertaisesti *** Monte Carlo -varianssi *** todistus. Monte Carlo -varianssin laskeminen on kuitenkin laskennallisesti tehotonta, sillä se edellyttää algoritmin ajamista useita, todennäköisesti kymmeniä kertoja. Monissa käytännön sovelluksissa jo yhden hiukassuodinalgoritmin ajaminen vaatii runsaasti laskentatehoa, jolloin Monte Carlo -varianssin laskeminen ei ole mahdollista. Varianssia ei luonnollisesti voi myöskään laskea analyttisesti. Varianssin estimointiin on kuitenkin viime vuosina kehitetty muutamia algoritmeja, jotka mahdollistavat varianssin estimoinnin aikakompleksisuudella TODO.

Luku 3

Hiukassilottimet

Hiukassiloittimissa blahblah. Tässä luvussa käsiteltävät algoritmit jaetaan kahteen pääkategoriaan, offline-algoritmeihin, joita sovelletaan XXX sekä online-algoritmeihin, jotka ajetaan yhdessä hiukassuodinalgoritmin kanssa. Käytännössä online-algoritmit vaativat XXX.

3.1 Offline-algoritmit

Lorem ipsum

3.1.1 SIR-siloitin

3.2 Online-algoritmit

Lorem ipsum

Luku 4

Hiukassuodin ja -siloitin sisätilapaikannuksessa

Sisätilapaikannus tarkoittaa nimensä mukaisesti ihmisten tai esineiden automaattista paikantamista sisätiloissa. Koska GPS toimii sisätiloissa huonosti tai ei lainkaan, tarvitaan näihin ympäristöihin muita paikannusratkaisuja. Yleinen valinta ovat erilaiset Bluetooth-standardiin tai muuhun radioteknologiaan perustuvat lähetin-vastaanotinratkaisut. Turkulainen teknologia- ja analytiikkayritys Walkbase käyttää Bluetooth-sisätilapaikannusta erityisesti ruokakaupoissa sekä tavarataloissa asiakkaiden käyttäytymistä koskevan datan keräämiseen. Tyypillisessä asennusskenaariossa lähettimet (tagit) kiinnitetään ostoskärryihin sekä -koreihin ja paikantimet kiinnitetään liiketilan kattoripustuksiin. Markkinoilla on lukuisia sisätilapaikannusratkaisuja, mutta kustannussyistä Walkbase on kehittänyt oman laitteistoratkaisunsa, jonka tavoitteena on tarjota kaikissa ympäristöissä 95% varmuudella alle metrin paikannustarkkuus.

Esimerkissä käytetään SMC-algoritmia Bluetooth-paikannussovelluksessa lähettimen sijainnin laskemiseen. Paikannukseen käytettävä data kerättiin toimistoympäristössä Bluetooth Low Energy (BLE) -lähettimen sekä kattoon sijoitettujen vastaanottimien avulla. Havainnot koostuvat vastaanottimien lähettimien signaalien perusteella laskemista, BLE5.1-standardin mukaisista signaalin tulokulmista eli AoA-havainnoista (angle of arrival). Lopuksi esimerkissä analysoidaan ja vertaillaan algoritmin eri versioiden suorituskykyä sekä suorituskyvyn että paikannustarkkuuden näkökulmasta. Vertailuarvona käytetään perinteistä triangulaatio-algoritmia.

4.1 Koeasetelma

Paikannusesimerkissä lähettimenä toimi 25 Bluetooth-paikannustagista koostuva Walkbase Foculator -testilaitte (kuva 4.1), vastaanottimena toimistoympäristöön asennetut neljä Walkbase XR-2 -vastaanotinta (kuva 4.2). Jokainen vastaanotin sisältää kuusitoista antennia, joiden vastaanottamien lähetinsignaalien perusteella vastaanottimet laskevat signaalin tulokulman suhteessa vastaanottimeen. Tarkka tulokulmien laskemiseen käytetty algoritmi on paikantimen antennit toimittaneen Silicon Laboratories, Inc. -yrityksen liikesalaisuus, mutta perusperiaate on arvioida

tulokulma mittaamalla eri antennien välistä vaihe-eroa ns. IQ-signaalin avulla.

Esteettömässä ympäristössä koeasetelmassa käytetyn järjestelmän kulmavirhe on hyvin pieni ja paikannusongelma voidaan ratkaista riittävällä tarkkuudella suoraan triangulaatio-algoritmillä. Tässäkin tilanteessa voi paikannusta parantaa suodattimen käytöllä, mutta jo triangulaatio-algoritmin perusversio tuottaa halutun tarkkuuden. Toimistoympäristö on kuitenkin haastava, sillä erityisesti näyttöruudut sekä heijastavat että estävät radiosignaaleja. Silicon Labs lupaa omalle AoA-järjestelmälleen vastaavassa toimisympäristössä (seitsemällä paikantimella) kulmavirheen välillä $3.7^\circ - 5.7^\circ$. Tämä ei kuitenkaan riitä johdonmukaisesti haluttuun alle metrin paikannustarkkuuteen, joten AoA-paikannus toimistoympäristössä tarjoaa hyvän motivaation SMC-menetelmien käytölle.

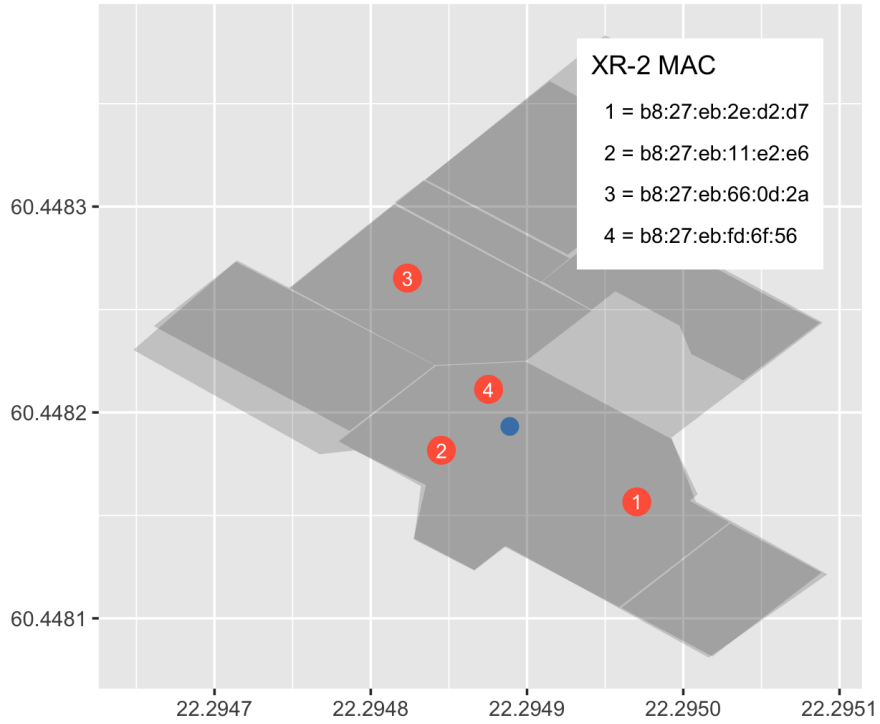


Kuva 4.1: Walkbase Foculator



Kuva 4.2: Walkbase XR-2

Data on kuvattu tarkemmin alaluvussa 5.2. Koeympäristön pohjapiirustus on esitetty kuvassa 4.3. Piirustuksessa XR-2-paikantimet on kuvattu punaisilla numeroituilla ympyröillä. Foculator-testilaitteen sijainti on kuvattu sinisellä ympyrällä. Toimistoympäristön pohjapiirustus on kuvattu harmaalla niin, että piirustuksesta erottuvat toimiston väliseinät.



Kuva 4.3: Koeasetelman pohjapiirustus

Toimisto valittiin testiympäristöksi, koska siellä testaaminen on helpompaa ja halvempaa kuin aidossa kauppaympäristössä. Lisäksi toimistossa hyvin toimiva järjestelmä todennäköisesti toimii sellaisenaan vähemmän esteisessä kauppaympäristössä.

Dataa kerättiin sekä liikkuvalla että paikallaan olevalla testilaitteella. Paikallaan oleva testilaitte asetettiin pöydälle 1.5 metrin korkeudelle. Liikkuvassa tapauksessa testilaitte kiinnitettiin robottiin, joka oli ohjelmoitu seuraamaan haluttua liikerataa vastaavia lattiamerkintöjä ennalta määrätyllä vakionopeudella. Yksinkertaisuuden vuoksi tässä tutkielmassa tarkastellaan ainoastaan ongelmaa, jossa testilaitte on paikoillaan. Esitetyt mallit soveltuvat kuitenkin myös liikkuvalla laitteella. Data kerättiin yöaikaan, jolloin toimiston käyttöaste oli alhainen. Tällä minimoitiin radiosignaalien tielle osuvien ihmisten vaikutus signaaleihin. Maan pinnan kaarevuus on otettu huomioon koeasetelmassa pisteiden välisiä etäisyyksiä laskettaessa.

4.2 Datan kuvaus

Riippuen päällä olevien paikannustagien lukumäärästä sekä kulmaestimaattien luomiseen käytetystä laskentamenetelmästä tuottaa testilaitte noin 60 havaintoa sekunnissa jokaista vastaanotinta kohden. Koetta varten käytettiin yhden minuutin aikana kertyneitä havaintoja. Havaintoja on datassa yhteensä $N_{obs} = 15018$ kappaletta. Havaintojen aikaleimat on tallennettu sekunnin tuhannesosan tarkkuudella. Jokainen havainto koostuu taulukossa 4.1 kuvatuista muuttujista.

Muuttuja	Kuvaus	Esimerkkiarvo
ts	havainnon aikaleima	21:38:20.998+00
locator_mac	vastaanottimen MAC-osoite	b8:27:eb:66:0d:2a
azimuth_angle	atsimuuttikulma θ (astetta)	102.7
converted_azimuth	napapohjoisesta laskettu kulma Φ (astetta)	34
snr_ss	antennikohtaisten signaali-kohinasuhteiden neliösumma	6996.473
rss	signaalin vahvuus (dBm)	-81
distance	arvioitu etäisyys lähettimeen (m)	13.5
opposite_angle	vastakkainen kulma (lähetin-vastaanotin) Φ' (radiaania)	3.735
lat	vastaanottimen sijainti (leveyspiiri)	60.448265
lon	vastaanottimen sijainti (pituuspiirit)	22.294823
bearing	suuntimakulma η (astetta)	34
height	vastaanottimen korkeus (m)	2.22

Taulukko 4.1: Havaintomuuttujat

Etäisyys on estimoitu signaalin vahvuudesta käyttäen valmistajan omaa propaatiomallia. Etäisyyttä tai signaalin vahvuutta ei käytetä paikantamiseen, joten tämän mallin käsittely jätetään tutkielman ulkopuolelle. Munoz (2009) luku 2 sisältää yleiskatsauksen propaatiomalleista.

Atsimuuttikulma ϕ lasketaan aina vastaanottimen tietyltä sivulta, joten se vastaa napapohjoista ainoastaan siinä tapauksessa, että vastaanottimen kyseinen sivu on asetettu kohtisuoraan napapohjoiseen nähden. Käytännön syistä tämä ei ole aina mahdollista eikä vaihe-erojen mittaamisen kannalta edes suotavaa. Tämän vuoksi jokaiselle vastaanottimelle on tietokantaan tallennettu oma suuntimakulma η . Kokeessa käytetään napapohjoisesta laskettuja kulmia Φ , jotka lasketaan jokaiselle havainnolle havainnon vastaanottimen suuntimakulman avulla

$$\Phi = (\theta + \eta + 360^\circ) \mod 360^\circ. \quad (4.1)$$

Suuntimakulma Φ kertoo vastaanottimen ja lähettimen välisen kulman. Koska triangulaatiossa tarvitaan päinvastaista kulmaa, lasketaan lisäksi yhtälöllä (4.2) suuntimakulman perusteella lähettimen ja vastaanottimen välinen kulma Φ' . Samalla asteissa oleva kulma muunnetaan radiaaneiksi triangulaatiota varten.

$$\Phi' = \frac{\pi}{180^\circ}(\Phi + 180^\circ \mod 360^\circ). \quad (4.2)$$

Lisäksi havainnot sisältävät antennikohtaisista SNR-arvoista lasketun neliösumman, jota käytetään erityisesti triangulaatiossa vahvimpien havaintojen valitsemiseen.

Havaintomuuttujien ohella koetilanteessa tallennettiin Foculator-testilaitteen todellinen sijainti sekunnin tarkkuudella. Tallennettu sijainti perustuu koeympäristön lattiaan pohjapiirrustusten sekä laser-mittausten avulla tehtyihin merkintöihin. Näin saadut testimuuttujat on kuvattu taulukossa (4.2). Kun testilaitte on paikoillaan, ovat nämä havainnot luonnollisesti samat jokaiselle ajanhetkelle.

Muuttuja	Kuvaus	Esimerkkiarvo
foculator_ts	aikaleima	21:38:20+00
foculator_lat	testilaitteen sijainti (leveyspiiri)	60.44819
foculator_lon	testilaitteen sijainti (pituuspiiri)	22.29493

Taulukko 4.2: Testimuuttujat

Testimuuttujia käytetään paikannusalgoritmien paikannusvirheen laskemisessa.

4.3 Ongelman kuvaus

Tarkoituksena on estimoida testilaitteen/robotin sijainti sekunnin tarkkuudella. Merkitään estimoitavaa tilasarjaa $x_{1:k} = \{x_1, \dots, x_k\}$. Lisäksi merkitään x_0 testilaitteen lähtösijaintia. Jokainen tilasarjan havainto koostuu suuntimakulmasta sekä pituudesta leveyskoordinaateista (x_k^x, x_k^y) . Määritellään tilalle testilaitteen/robotin kulkua kuvaava vektorisuunnistukseen (dead reckoning) perustuva malli (4.3)

$$x_{k+1} = f(x_k, \nu_k) = x_k + D_k \begin{bmatrix} \cos \psi_k \\ \sin \psi_k \end{bmatrix} + \nu_k, \quad (4.3)$$

missä D_k on testilaitteen/robotin ajanhetkenä k kulkema matka ja ψ_k testilaitteen/robotin atsimuutti/suuntimakulma kyseisenä ajanhetkenä. ν_k on kohinaa, joka syntyy mittausvirheestä ja jolle voidaan olettaa $\sim \mathcal{N}(\mu_x, \sigma_x^2)$. Kun laite on paikallaan, yksinkertaistuu malli muotoon $x_{k+1} = f(x_k) = \text{id}(x_k) = x_k$, missä $\text{id}(\cdot)$ on identiteettifunktio.

Vastaavasti $y_{1:k} = \{y_1, \dots, y_k\}$ kuvaa paikantimien ja lähettimien välillä laskettuja kulmahavaintoja. Näin ollen jokainen havainto koostuu (maksimissaan) paikantimien määrää vastaavasta neljästä kulmasta. Havainnot lasketaan sekunnin tarkkuudella, mutta todellinen havaintotarkkuus on tiheämpi. Tästä syystä jokaisen sekunnin sensorihavainnot $\{s_k^1, \dots, s_k^4\}$ muodostetaan keskiarvona kaikista kyseisen sekunnin aikana tapahtuvista paikannin-lähetinparien kulmahavainnoista

$$y_k = \begin{bmatrix} \frac{1}{n} \sum_{j=1}^n s_{k,j}^1 \\ \vdots \\ \frac{1}{n} \sum_{j=1}^n s_{k,j}^4 \end{bmatrix}. \quad (4.4)$$

Lisäksi tunnetaan sensoreihin $\{s^1, \dots, s^4\}$ liittyvät pituus- ja leveyskoordinaatit (λ, ϕ)

$$u = \begin{bmatrix} \lambda^1 & \phi^1 \\ \vdots & \vdots \\ \lambda^4 & \phi^4 \end{bmatrix}. \quad (4.5)$$

Määritellään havainnoille malli

$$y_k = h(x_k, u) + e_k = \text{atan2}\left(\begin{bmatrix} \phi^1 - x_k^y \\ \vdots \\ \phi^4 - x_k^y \end{bmatrix}, \begin{bmatrix} \lambda^1 - x_k^x \\ \vdots \\ \lambda^4 - x_k^x \end{bmatrix}\right) + e_k, \quad (4.6)$$

missä

$$\text{atan2}(y, x) = \begin{cases} \arctan(\frac{y}{x}) & \text{jos } > 0, \\ \arctan(\frac{y}{x}) + \pi & \text{jos } < 0 \text{ ja } y \geq 0, \\ \arctan(\frac{y}{x}) - \pi & \text{jos } > 0 \text{ ja } < 0, \\ +\frac{\pi}{2} & \text{jos } x = 0 \text{ ja } > 0, \\ -\frac{\pi}{2} & \text{jos } x = 0 \text{ ja } < 0, \\ \text{ei määritelty} & \text{jos } x = 0 \text{ ja } y = 0 \end{cases} \quad (4.7)$$

ja kohina noudattaa moniulotteista normaali jakaumaa $e_k \sim \mathcal{N}(0, \Sigma)$.

Kovarianssimatriisin estimaattina käytetään kunakin ajanhetkenä k antennikohtaisista havainnoista estimoituja otosvariansseja $\text{diag}(\hat{\sigma}_k^1, \dots, \hat{\sigma}_k^4)^2 = \text{diag}(\frac{1}{n-1} \sum_{i=1}^n (s_i^1 - \bar{s})^2, \dots, \sum_{i=1}^n (s_i^4 - \bar{s})^2)$. Määrittelemätön atan2-tapaus, jossa $x = 0$ ja $y = 0$ on käytetyllä mittaustarkkuudella käytännössä mahdoton. Jos tapaus halutaan välttää, voidaan nolla-arvot tarpeen vaatiessa korvata joillakin hyvin lähellä nollaa olevalla arvolla. Saadaan uskottavuusfunktioiksi

$$p(y_k | x_k) \propto \prod_{j=1}^4 \exp \left\{ -\frac{\|h(x_k^j, u) - y_k^j\|^2}{2(\hat{\sigma}_i^j)^2} \right\}, \quad (4.8)$$

missä $j = \{1, 2, 3, 4\}$ vastaa nyt kutakin XR-2-vastaanotinta.

Kumpikaan funktiosta $h(\cdot)$ ja $f(\cdot)$ ei ole lineaarinen, joten SIR-algoritmi on sopiva valinta ongelman ratkaisemiseksi. Koetuloksia arvioidaan ensisijaisesti paikannusvirheen avulla. Paikannusvirhe e_k lasketaan jokaisen ajanhetken k posteriorijakaumaestimaatista \hat{p}_k painotettuna keskiarvona

$$\epsilon_k = \sum_{i=1}^N w_i^k d(x_k^i, y_k), \quad (4.9)$$

missä w_i^k on ajanketken k partikkelien normalisoitu paino ja $d(x_k^i, y_k)$ partikkelien ja testilaitteen todellisen sijainnin välisen etäisyyden laskeva funktio.

4.4 Algoritmin toteutuksesta

Koeasetelmassa käytetty SMC-algoritmi on toteutettu R-kielellä. Algoritmin toteutus on pääosin vektorisoitu ja tehokas. For-silmukkaa on käytetty ainoastaan ajanhetkien läpikäyntiin. Koska tämän silmukan muuntaminen vektorisoituun muotoon ei ole mahdollista, voidaan toteutusta pitää näiltä osin hyvin optimoituina. Algoritmin rungon datan käsittely on toteutettu suorituskyyvyltään hyvällä `data.table`-kirjastolla. Uskottavuusfunktiossa on käytetty hitaampia `dpylr`-kirjaston `tibble`-taulukkoja. Toteutustapa on valittu koodin ymmärrettävyyden parantamiseksi ja on edelleen riittävän nopea koeasetelman tarpeisiin. Koodin profilointi paljastaa, että suurin yksittäinen osa algoritmin ajasta kuluu `tibble`-taulukkojen käsittelyssä, joten suorituskyykyä on tarpeen vaatiessa mahdollista myös parantaa.

Koska algoritmin uskottavuudet sekä painot ovat hyvin pieniä, on laskenta-tarkkuusongelmien välttämiseksi toteutuksessa käytetty logaritmoituja painoja ja uskottavuusfunktioita. Tämä ei vaikuta lainkaan itse algoritmin toimintaan, mutta estää numeeristen ongelmien syntymisen. Tilanteessa, jossa tietyn paikantimen ja kaikkien partikkelien välinen uskottavuus on nolla, päätyvät kaikki painot nolliksi eikä algoritmi enää toimi. Tällaisessa tilanteessa uskottavuusfunktion R-toteutus kutsuu itseään rekursiivisesti uudelleen niin, että kyseinen paikannin on tiputettu havainnoista.

Paikannusvirheen laskemisessa on etäisyysfunktiona $d(\cdot)$ käytetty `raster`-kirjaston `pointDistance()`-funktioita. Koodissa on korostettu tiiviyn sijaan luettavuutta ja koodi on kommentoitu kattavasti. SMC-algoritmfunktion sekä uskottavuusfunktion R-koodit löytyvät osoitteesta <https://github.com/rintakumpu/luk>, mistä löytyy myös asetelmassa käytetty data csv-muodossa sekä kokeen muun koodin sisältävä R Markdown -notebook.

4.5 Parametrien valinta

Priorijakaumana p_{x_0} käytettiin kahta toisistaan riippumatonta otosta tasajakaumista, joista toinen vastasi leveys- ja toinen pituusasteita. Jakaumien alkupisteet valittiin niin, että ne vastasivat pienimpiä paikantimien leveys- ja pituusasteista. Vastaavasti päätepisteet valittiin niin, että ne vastasivat suurimpia paikantimien leveys- ja pituusasteita.

$$p_{x_{0_{\text{lon}}}} \sim \mathcal{U}(\min \lambda, \max \lambda), \quad (4.10)$$

$$p_{x_{0_{\text{lat}}}} \sim \mathcal{U}(\min \phi, \max \phi). \quad (4.11)$$

Koska järjestelmän on tarkoitus toimia ainoastaan paikantimien muodostaman suorakaiteen sisäpuolella, ovat valitut jakaumien päätepisteet riittävät. Kummastakin jakaumasta otettiin \sqrt{N} otosta, jolloin N partikkelia x_0^i saatiin näiden otosten permutaatioina.

Paikannus suoritettiin yhteensä 17 kertaa. Ensimmäisessä vaiheessa tarkasteltiin partikkelien määrän N vaikutusta paikannuskeskivirheeseen ilman uudelleenotantaa sekä priori- että uskottavuusotannalla. Seuraavassa vaiheessa valittiin ehdotusjakautuma edellisen vaiheen tulosten perusteella ja tarkasteltiin tilannetta sekä adaptiivisella että jokaisella iteraatiolla suoritettavalla uudelleenotannalla.

Kaikkien tulosten vertailukohtana käytettiin Pierlot & al. artikkelissa “A New Three Object Triangulation Algorithm Based on the Power Center of Three Circles” (2011) esittämää ToTal-triangulaatioalgoritmia. Triangulaatio-algoritmia ei käsitellä tässä tarkemmin, mutta se on esitetty algoritmissa REFERENSSI TOTALIIN. Algoritmia varten valittiin kunakin ajanhetkenä k ne kolme paikanninta ja kulmahavaintoa, joiden SNR-arvo oli korkein.

Algoritmit ajettiin RStudio versiossa 1.4.1106 R-ohjelmointikielen versiolla 4.0.4. Tietokoneena käytettiin vuoden 2017 mallia olevaa MacBook Pro -kannettavaa, jossa oli 3.1 Ghz Quad-Core i7 -prosessori sekä 16 gigatavua 2133 Mhz LPDDR3 -muistia. Suoritusnopeuden mittaamiseen käytettiin `microbenchmark`-kirjastoa. Jokainen algoritmi toistettiin kymmenen kertaa ja suoritusnopeus laskettiin näiden toistojen aritmeettisena keskiarvona.

4.6 Teknologian kuvaus

Here is an equation.

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad (4.12)$$

You may refer to using `\@ref(eq:binom)`, like see Equation (4.12).

4.6.1 Kalibraatioalgoritmi

4.7 AoA-menetelmistä

4.7.1 MUSIC-algoritmi

Footnotes are put inside the square brackets after a caret `^[]`. Like this one ¹.

Citations. Reference items in your bibliography file(s) using `@key`.

¹This is a footnote.

For example, we are using the **bookdown** package (check out the last code chunk in index.Rmd to see how this citation key was added) in this sample book, which was built on top of R Markdown and Note that the `.bib` files need to be listed in the index.Rmd with the YAML `bibliography` key.

The RStudio Visual Markdown Editor can also make it easier to insert citations: <https://rstudio.github.io/visual-markdown-editing/#/citations>

4.8 Datan kuvaus

4.9 Ongelman ja mallien kuvaus

4.9.1 Datan valinta

Convex hull goes here.

4.9.2 Uskottavuusmallit

4.9.3 Dynaaminen malli

4.10 Algoritmien toteutuksesta

4.11 Karttasovitusalgoritmi

4.12 Reitinhakualgoritmi

4.13 Parametrien valinnasta

4.14 Tulokset

Luku 5

Lopuksi

Tässä tutkielmassa on esitetty pääpiirteittäin SMC-menetelmien teoria Bayesilaisessa tilastotieteellisessä viitekehyksessä. Lisäksi tutkielmassa on käyty läpi uudelleenotantaa efektiivisen otoskoon perusteella hyödyntävä SIR-suodinalgoritmi. Lopuksi tutkielmassa on tarkasteltu SIR-algoritmin parametrien valintaan, suoritussykyyn sekä konvergenssiin liittyviä tuloksia.

Tutkielmassa on lisäksi tarkasteltu miten eri valinnat vaikuttavat algoritmin suoritussykyyn yksinkertaisen mutta kattavan ja todelliseen ongelmaan sekä dataan perustuvan paikannusesimerkin avulla.

Luku 6

(Liite) Other stuff {-} (followed
by# A chapter‘)

Lähteet

- [1] Dan Crisan. The stochastic filtering problem: A brief historical account. *Journal of Applied Probability*, 51A:13–22, 2014.
- [2] Fredrik Gustafsson. Particle filter theory and practice with positioning applications. *IEEE Aerospace and Electronic Systems Magazine*, 25(7):53–82, 2010.
- [3] R. Chen J. Liu. Sequential monte carlo methods for dynamic systems. *Journal of the American Statistical Association*, 93(443):1032–1044, 1998.
- [4] Angus P. Andrews Mohinder S. Grewal. Applications of kalman filtering in aerospace 1960 to the present. *IEEE Control Systems Magazine*, 30(3):69–78, 2010.
- [5] Pierre Del Moral. Nonlinear filtering: Interacting particle resolution. *Markov Processes and Related Fields*, 2(4):555–580.
- [6] A.F.M Smith N.J. Gordon, D.H. Salmond. Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEE Proceedings F (Radar Signal Process)*, 140(2):107–113, 1993.
- [7] Eric Moulines Olivier Cappé, Simon J. Godsill. An overview of existing methods and recent advances in sequential monte carlo. *Proceedings of the IEEE*, 95(5): 899–924, 2007.
- [8] Simo Särkkä. *Bayesian Filtering and Smoothing*. Cambridge University Press, 2013. URL https://users.aalto.fi/~ssarkka/pub/cup_book_online_20131111.pdf.