

AALTO-YLIOPISTO
Perustieteiden korkeakoulu
Teknillisen fysiikan ja matematiikan tutkinto-ohjelma

Yrjänä Hynninen

Asiantuntija-arvioinnit asejärjestelmien kustannustehokkuuden arvioinnissa

Mat-2.4108 Sovelletun matematiikan erikoistyöt

Espoo, 28.10.2011

Työn saa tallentaa ja julkistaa Aalto-yliopiston avoimilla verkkosivuilla. Muilta osin kaikki oikeudet pidätetään.

Sisältö

1 Johdanto	1
2 Asiantuntija-arvioinnit tietolähteenä	2
2.1 Asiantuntijoiden valinta	2
2.2 Arviointitietojen määrittäminen	2
2.3 Arviointitietojen hyödyntäminen	3
2.4 Asiantuntija-arvioinnin haasteet	4
3 Asejärjestelmien kustannustehokkuuden arviointi	7
3.1 Lähtökohdat ja tavoitteet	7
3.2 Arviointiviitekehys	7
3.3 Asiantuntija-arvioinnit	8
3.4 Asejärjestelmäyhdistelmien kustannustehokkuus	10
3.5 Kustannustehokkuusanalyysin tulokset	11
4 Pohdinta	14
5 Yhteenveto	15
6 Viitteet	16
7 Liitteet	17
7.1 Mallin matemaattinen formulointi	17

1 Johdanto

Asiantuntija-arvioiden käyttö päätösanalyysissä on yleistynyt viime vuosikymmeninä ja muuttunut entistä hyväksytyimmäksi tiedon hankintatavaksi. Asiantuntijoihin tukeudutaan tilanteissa, joissa on paljon epävarmuutta ja joita ei voi arvioida matemaattisessa muodossa olevan informaation - esimerkiksi historiallisten aikasarjatietojen - avulla. Tällaisissa tilanteissa asiantuntija-arviot antavat lähtötietoja, joiden pohjalta voidaan tuottaa perusteltuja päätössuosituksia.

Asiantuntija-arvioinnissa on kaksi eri haaraa (Morris, 1974): (i) Miten kerätä informaatiota asiantuntijoilta mahdollisimman tarkoituksenmukaisesti ja (ii) miten yhdistää ja käyttää kerättyä informaatiota hyödyllisesti? Molempien kysymysten kohdalla alalla on erilaisia suuntauksia, mutta laajasti vakiintuneita menetelmiä ei ole vielä kehittynyt.

Keskeisimpinä haasteina asiantuntijoihin tukeutumisessa ovat mielipiteiden epävarmuudet, ihmisten heuristiikat ja erilaiset ajatusmallit. Asiantuntijat eivät ole poikkeuksia näiden inhimillisten piirteiden omistajina (Lichtenstein ja Fischhoff, 1977). Lisäksi jokainen tilanne on ainutlaatuinen eikä mikään menetelmä ole paras kaikissa tilanteissa (Clemen ja Winkler, 1999). Päätökset haluttavien asioiden kyselymenetelmistä datan mallinnukseen ja järkevään yhdistämiseen liittyvät aina tarkasteltavaan kokonaisuuteen.

Alan kirjallisuudessa kyselymenetelmät vaihtelevat asiantuntijoiden itsenäisestä työskentelystä ryhmätyöhön sekä näiden kahden ääripään välillä oleviin muunnelmiin, kuten Delfoi-menetelmään (Dalkey ja Helmer, 1963). Asiantuntija-arvioiden yhdistämisessä on suosittu lähinnä painotettuja keskiarvoja ja bayesiläistä lähestymistapaa (Sandri ym., 1995). Perusideana on ollut yhdistää useat arviot yhdeksi arvioksi, jonka perusteella suosituksia ja päätöksiä tehdään.

Tässä työssä kehitetään asejärjestelmien kustannustehokkuusanalyysiin liittyvän esimerkin kautta uusi lähestymistapa asiantuntija-arvioiden hyödyntämiseen. Lähestymistapa perustuu arvioiden yhdistämisen sijasta erillisten arvioiden hyödyntämiseen ja robustien ratkaisujen löytämiseen. Menetelmä antaa päätöksentekijälle perusteltuja suosituksia ja ohjaa huomiota olennaisiin näkökohtiin.

Työn rakenne on seuraavanlainen. Seuraavassa luvussa tehdään kirjallisuuskatsaus asiantuntija-arvioinnin menetelmiin ja niihin liittyviin haasteisiin. Kolmannessa luvussa esitellään case-esimerkin kautta uuden lähestymistavan perusteet ja tarkastellaan sen tuottamia tuloksia. Neljännessä luvussa pohditaan menetelmän etuja ja haasteita, ja viidennessä luvussa esitetään yhteenveto.

2 Asiantuntija-arvioinnit tietolähteenä

Kirjallisuudessa asiantuntijoiden käyttö jaetaan usein kahteen osaprosessiin (Morris, 1974): Ensimmäisessä vaiheessa arviot haetaan asiantuntijoilta erilaisin kysely- ja haastattelumenetelmin ja toisessa vaiheessa arvioita tarkastellaan analysointimenetelmin ja annetaan päätösuosituksia. Molempia vaiheita edeltää asiantuntijuuden määrittäminen ja asiantuntijoiden valinta. Seuraavaksi käsitellään näitä kolmea osaprosessia, esitellään yleisimpiä menetelmiä ja tarkastellaan keskeisimpiä haasteita.

2.1 Asiantuntijoiden valinta

Asiantuntijuuden määrittäminen ja asiantuntijoiden valinta ovat lähtökohtia asiantuntijoiden hyödyntämisessä. Alan kirjallisuus korostaa, että asiantuntijat tulisi valita ammatillisen maineen, kokemuksen ja - jos mahdollista - asiantuntemusalueella olevan julkaisuhistorian perusteella (Cooke ja Goossens, 2004). Asiantuntijapaneelilla tulee olla tehtävään oikeanlaista osaamista ja sen tulee kattaa tehtävän asettamat vaatimukset.

Asiantuntijapaneelin koko on merkittävä tekijä. Hoffmann ym. (2007a) ovat tutkineet suurien (yli 20 asiantuntijaa) ja heterogeenisten paneelien käyttöä. He pitävät niiden etuna mahdollisuutta parempaan tilastolliseen analyysiin. Toisaalta haasteena on löytää tarpeeksi suuri määrä ammattitaitoisia ja tehtävään haluavia asiantuntijoita. Paneelin koon kasvun ei tule johtaa asiantuntijoiden osaamistason alenemiseen.

Usein tavoitteena on saada arviot koko pätevien asiantuntijoiden muodostamalta populaatiolta tai sitä edustavalta otokselta. Molemmissa tapauksissa haasteena on selvittää asiantuntijapopulaatio, joka ei ole yleisesti tiedossa. Tämän tyyppisten populaatioiden selvittämiseen on kehitetty menetelmiä, joista eräs tunnettu on niin kutsuttu lumipal-lomenetelmä (Kalton ja Anderson, 1986). Tässä menetelmässä muutamaa tunnettua ja pätevää asiantuntijaa pyydetään tunnistamaan kyseiseen tehtävään soveltuvat asiantuntijat ammatillisesta ympäristöstään. Tämän jälkeen kaikkia ehdotettuja asiantuntijoita pyydetään tekemään sama tunnistamistehtävä. Tätä toistetaan muutaman kerran, kunnes asiantuntijat ehdottavat lähinnä jo aiemmin mainittuja asiantuntijoita. Voidaan olettaa, että populaatio olisi täten riittävän kattavasti tunnistettu.

2.2 Arviointitietojen määrittäminen

Asiantuntija-arvioiden hakumenetelmät vaihtelevat asiantuntijoiden täysin itsenäisestä työskentelystä ryhmän yhteisen konsensuksen hakemiseen. Näiden ääripäiden välimaastossa on menetelmiä kuten Delfoi, joka sisältää yhtäältä itsenäisen ja anonyymin

arvioimisen, ja toisaalta arvioiden yhteisen vertailemisen välisiä kierroksia (Dalkey ja Helmer, 1963).

Asiantuntija-arvioinnissa on oleellista suunnitella kartoitusmenetelmät silmälläpitäen sitä, miten saatuja vastauksia analysoidaan. Clemen ja Winkler (1999) toteavat artikkelissaan, että mikään yksittäinen menetelmä ei ole paras kaikkiin tilanteisiin ja että prosessi voi sisältää ominaisuuksia useista menetelmäluokista. Heidän mielestään asiantuntijoiden haastatteluprosessin laatijoihin ja toteuttajiin on tarpeen kuulua sekä substanssialan tuntijoita että analysoinnin osajia. Cooke ja Goossens (2008) esittävät lisäksi, että kyselyä on testattava vähintään yhdellä asiantuntijalla kysymysmuotojen ja ymmärrettävyyden varmistamiseksi.

Monien arviointeihin liittyvien heuristiikkojen ja harhojen minimoimiseksi asiantuntijat on valmennettava huolellisesti. Ensinnäkin on tärkeitä, että asiantuntijat ymmärtävät arviointitehtävän merkityksen ja rakenteen sekä sen, miten heidän arvioitaan hyödynnetään. Toisekseen formaali muoto, jossa tietoa pyydetään asiantuntijoilta - esimerkiksi todennäköisyyspohjaisesti tai luottamusväleillä - voi oudoksuttaa. Tämän takia on tärkeitä saattaa asiantuntijat mahdollisimman tasa-arvoiseen asemaan kyselymenetelmiin nähden. Valmennus myös lisää asiantuntijoiden motivaatiota.

Kyselytilanteen tavoitteena on tallentaa asiantuntijan tietämyksen tila tarkasteltavasta asiasta, mikä voidaan tehdä todennäköisyyksien avulla. Todennäköisyydet tarjoavat tähän matemaattisen keinon (Keeney ja von Winterfeldt, 1989). Pyytämällä asiantuntijaa arvioimaan vaihtoehtoisten tapahtumien todennäköisyyksiä, toisin sanoen määrittämään todennäköisyysjakauma, pystytään arvioimaan asiantuntijan varmuus ja tietämys tarkasteltavasta asiasta. Arvioiden varmuuden selvittäminen on tärkeitä, sillä yleensä asiantuntijoilta haetaan näkemystietoa tilanteissa, joihin liittyy merkittävästi epävarmuutta.

Eräs yleinen tapa selvittää arvioihin liittyvää varmuutta on pyytää asiantuntijoita määrittämään luottamusväli todennäköisimmäksi arvioimansa tapahtuman ympärille. Yleensä käytetään 90 % intervallia eli asiantuntijaa pyydetään määrittämään väli, johon hän uskoo todellisen arvon sijoittuvan 90 % todennäköisyydellä. Tällaisella menetelmällä selvitettyä varmuutta voidaan käyttää hyväksi arvioiden yhdistämisessä, painottamisessa ja analysoinnissa.

2.3 Arviointitietojen hyödyntäminen

Arviointitietojen analysointimenetelmät jaetaan usein kahteen luokkaan: matemaattisiin menetelmiin ja käyttäytymistieteellisiin menetelmiin (Clemen ja Winkler, 1999). Empiiristen kokeiden perusteella molemmat menetelmät ovat tarkkuudeltaan yhtä hy-

viä, kuitenkin siten, että matemaattisilla menetelmillä on pieni etu. Tässä työssä keskitytään matemaattisiin menetelmiin.

Alan tutkimuksessa ja kirjallisuudessa hallitsevana tapana on hyödyntää arviointitietoja yhdistämällä ne yhdeksi jakaumaksi, jonka oletetaan kuvaavan todellisuutta parhaiten ja jonka avulla päätökset tehtäisiin suoraviivaisesti. Eräs luonnollinen tilanne, jossa arviot yhdistetään, on lähtötietojen vieminen suurempaan malliin. Yleisimpinä yhdistämismenetelminä käytetään painotettuun keskiarvoon perustuvia menetelmiä sekä bayesiläiseen lähestymistapaan perustuvia menetelmiä (Sandri ym., 1995).

Painokertoimille perustuvissa menetelmissä arviointitiedot yhdistetään painottamalla niitä tasaisesti tai tietyin perustein annetuina kertoimin. Painotuksen tavoitteena on antaa luotettavalle arviolle suurempi painoarvo kuin vähemmän luotettavalle. Haasteena on luotettavuuden määrittäminen ja asiantuntijoiden kyvykkyyden mittaaminen. Eräitä perusteita painotukselle ovat asiantuntijan historiatiedot ja itsearviointit (Helmer, 1963).

Cooke (1991, ss. 187-198) esittelee kirjassaan painotettuun summaan perustuvan niin sanotun klassisen mallin, jossa yksittäisen asiantuntijan kyvykkyys eli painoarvo määritetään testikysymysten avulla. Testikysymykset ovat alaan liittyviä kysymyksiä, joiden oikeat tarkat vastaukset tiedetään. Asiantuntijat vastaavat testikysymyksiin täysin samalla menetelmällä kuin todellisiin tutkimuskysymyksiin. Vertaamalla asiantuntijoiden vastauksia ja todellisia arvoja määritetään jokaisen asiantuntijan niin sanottu kalibraatio (kuinka lähellä oikeata arvoa asiantuntijan arvio on) ja informatiivisuus (kuinka varma asiantuntija on oman arvionsa paikkansapitävyydestä). Asiantuntijoiden arviota todellisesta tutkittavasta asiasta painotetaan yhteisjakaumassa näiden kahden tekijän perusteella.

Bayesiläiseen lähestymistapaan perustuvissa menetelmissä käytetään arviotietoja havaintoina todellisesta arvosta. Arviotietojen perusteella päivitetään päätöksentekijän oletus- eli priorijakaumaa Bayesin teoreemaa käyttäen. Kirjallisuudessa on esitelty useita erilaisia bayesiläisiä malleja, joista yleiskatsauksen saa muun muassa Cooken kirjasta (Cooke, 1991, ss. 176-184).

2.4 Asiantuntija-arvioinnin haasteet

Arviointiin liittyy erilaisia heuristiikkoja ja harhoja eivätkä asiantuntijat ole niiden suhteen poikkeuksia. Tutkimuksen laatijan ja analysoijan täytyykin tuntea ja ottaa huomioon keskeisimmät haasteet saadakseen luotettavia tuloksia. Erilaisia heuristiikkoja ja harhoja (yleisimpinä ankkurointi, saatavuusheuristiikka, edustavuusheuristiikka) on käsitelty paljon alan kirjallisuudessa (esim. Cooke, 1991, ss. 63-79), joten niihin ei tässä

työssä keskitytä. Kappaleen loppuosassa käsitellään hieman todennäköisyyspohjaiseen arviointiin keskeisesti liittyvää liiallisen itseluottamuksen harhaa.

Seuraavassa esitellään arvioiden haku- ja hyödyntämismenetelmiin liittyviä haasteita.

Asiantuntijoiden arvioihin perustuviin tutkimuksiin on varsinkin ennen suhtauduttu varsin varauksellisesti. Yleisön epäluulon poistamiseen auttaa huomattavasti tutkimuksessa käytettyjen asiantuntijoiden nimien ja taustan paljastaminen (Hoffmann ym., 2007b). Arvioiden luotettavuus perustuu yksittäisten asiantuntijoiden ammattitaitoon (Hoffmann ym., 2007a). Tämä on eräs syy, jonka takia vastaajien nimilistan tulisi olla julkinen. Toistettavuuden ja tutkimuksen perusteltavuuden takia myös nimet ja vastaukset on pystyttävä yhdistämään, mutta ei tarvittaessa julkisesti.

Kyselytilanteen ohjaajan kokemuksella on myös merkitystä. Kyselytilanne on edelleen eräänlainen "taiteenmuoto", jossa kokeneet kyselyn ohjaajat saavat parempia tuloksia aikaisiksi kuin kokemattomat (Lin ja Bier, 2008). Itse kyselytilanne ei saisi ylittää kestoltaan puolta tuntia (Cooke ja Goossens, 2008).

Keskiarvoistamiseen perustuvia menetelmiä on kritisoitu muun muassa niiden luotettavuuden takia (Sandri ym., 1995). Esimerkkinä on tilanne, jossa kahdella luotettavalla asiantuntijalla on toisistaan selkeästi poikkeavat mielipiteet tietystä arvosta. Tällöin keskiarvoistamalla saadaan kaksikon yhteismielipidettä kuvaamaan arvo, joka on arvioitujen arvojen välimaastossa eli ei kummankaan asiantuntijan mielestä lähellä todellista arvoa. Painotetun keskiarvon metodi saattaa myös vaikuttaa varianssiin siten, että tuloksen varianssista tulee pienempi kuin yksittäisten arvioiden varianssista (Sandri ym., 1995). Tämä ilmiö on hyväksyttävää, jos asiantuntijat ovat toisistaan riippumattomia. Asiantuntijoilla on kuitenkin usein suuri määrä samaa - esimerkiksi teknistä - informaatiota käytössään ja riippumattomuusoletus on tällöin kyseenalainen. Lisäksi keskiarvon käyttäminen kadottaa suuren osan arvioiden informaatiosta. Esimerkiksi yksittäisen asiantuntijan antamien vastausten väliset suhteet katoavat keskiarvoa käytettäessä.

Bayesiläiseen lähestymistapaan perustuvien menetelmien eräs suurimmista heikkouksista on se, että tutkittavasta ilmiöstä tarvitaan a priori -jakauma (Sandri ym., 1995). Toisin sanoen asiantuntijoiden apua tarvitsevan analysoijan pitäisi itsekkin olla asiantuntija. Monissa tapauksissa analysoijalla ei kuitenkaan ole käytettävää pohjatietoa arvosta.

Yksi yleisimmistä harhoista on liiallinen itseluottamus. Myös asiantuntijat luottavat omiin arvioihinsa enemmän kuin mihin arviointitarkkuus antaa edellytyksiä. Lichtenstein ja Fischhoff (1977) toteavatkin, että "liiallinen itseluottamus korostuu korkean vaikeustason tehtävissä". Esimerkiksi Russon ja Schoemakerin (1992) tutkimuksessa yritysjohtajien arvioimista 90% varmuusväleistä vain 40%-60% sisälsi todellisen arvon.

Liiallinen itseluottamus liittyy ihmisen perusluonteeseen ja on yleistä alasta tai koulutuksesta riippumatta. Yksilöiden välisiä eroja kuitenkin on (Lin ja Bier, 2008), mikä tekee ilmiöstä tutkimuksen kannalta osin ongelmallisen. Vaikutukset voivat olla suuret yhdistettäessä asiantuntija-arvioita yhdeksi jakaumaksi. Käytettäessä subjektiivisia todennäköisyysarvioita riskianalyysin puolella on erityisen tärkeää ottaa mahdollisesti liiallinen itseluottamus huomioon, etenkin ohjattaessa asiantuntijoita vastaustilanteeseen.

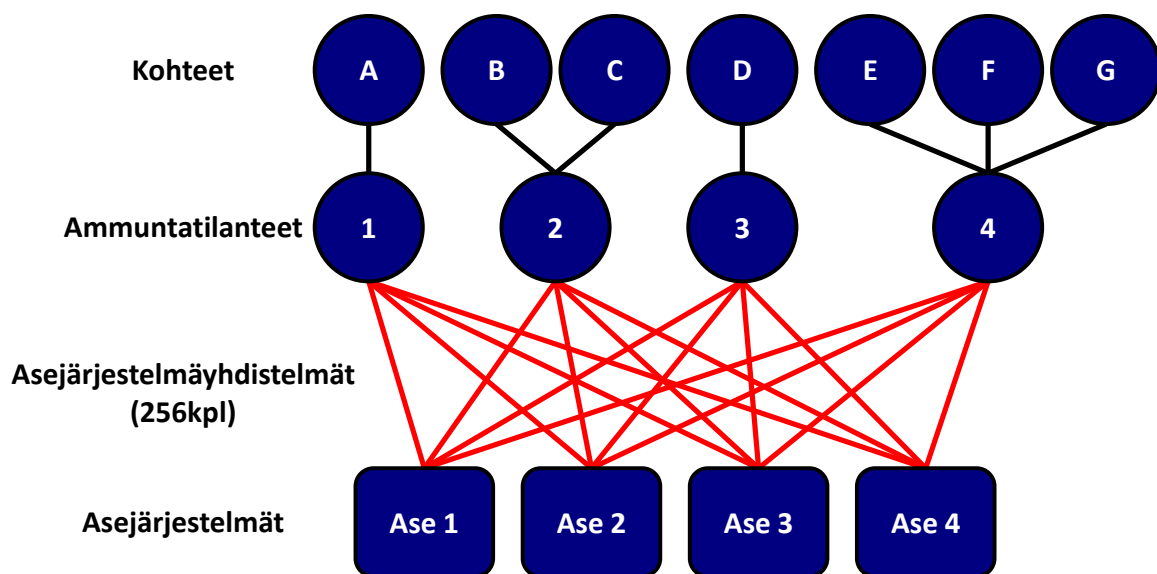
3 Asejärjestelmien kustannustehokkuuden arviointi

3.1 Lähtökohdat ja tavoitteet

Työn loppuosassa esitellään esimerkin kautta eräs tapa hyödyntää useita asiantuntija-arvioita. Huomattavaa käytetyssä arviointien hyödyntämistavassa on se, että yksittäisten asiantuntijoiden arvioita ei yhdistetä yhdeksi arvioksi esimerkiksi keskiarvoistamalla, kuten kirjallisuudessa usein tehdään, vaan niitä käytetään hyväksi erikseen. Menetelmän toiminta esitellään sovelluksen valossa. Matemaattinen formulointi tehdään liitteessä 1.

Esimerkki perustuu projektiin, jonka tarkoituksena oli tukea asiakkaan asejärjestelmien hankintapäätöstä. Tavoitteena oli selvittää neljän asejärjestelmän joukosta asiakkaan tarpeiden kannalta kustannustehokkaat valinnat. Asejärjestelmien rinnakkaiskäyttö oli myös mahdollista, joten yksittäisistä järjestelmistä koostuvien järjestelmäyhdistelmien eli -portfolioiden kustannustehokkuutta selvitettiin myös. Kustannustehokkuudella tarkoitetaan tässä yhteydessä sitä, että haluttu vaikutus saavutetaan mahdollisimman pienillä resursseilla.

3.2 Arviointiviitekehys



Kuva 1: Arvioinneissa käytetty viitekehys. Neljällä asejärjestelmällä vaikutetaan neljään ammuntatilanteeseen, jotka koostuvat yhdestä tai useammasta kohteesta. Erilaisia asejärjestelmäyhdistelmiä on olemassa 256 kpl.

Asejärjestelmien ja niiden yhdistelmien vaikuttavuutta tarkasteltiin neljän ammuntatilanteen avulla. Ammuntatilanteet koostuivat yhdestä tai useammasta kohteesta kuvan

1 mukaisesti. Asejärjestelmien vaikuttavuutta mitattiin kohteisiin aiheutettujen vaurioiden suhteen. Ammuntatilanteisiin liittyvät yksityiskohdat (esimerkiksi kohteiden etäisyydet, niiden suorituskyky ja varustus) oli määritetty asiakkaan toimesta.

Asejärjestelmäyhdistelmällä tarkoitetaan asejärjestelmistä ja niiden käyttötavoista muodostettua kokonaisuutta. Asejärjestelmien käyttötapoja rajoitettiin siten, että yhdessä ammuntatilanteessa on mahdollista käyttää ainoastaan yhtä asejärjestelmää. Tällöin vaihtoehtoisten yhdistelmien lukumäärä on 256.

3.3 Asiantuntija-arvioinnit

Asiakas määrittä asejärjestelmiin liittyvät suhteelliset yksikkökustannukset taulukon 1 mukaisesti. Kustannukset sisälsivät yhdistettynä hankinta- ja elinkaarikustannukset. Kussakin ammuntatilanteessa käytettiin eri asejärjestelmiä yhtä monta yksikköä, jolloin jokaisen asejärjestelmäyhdistelmän kokonaiskustannus pystyttiin laskemaan yksikkökustannusten avulla.

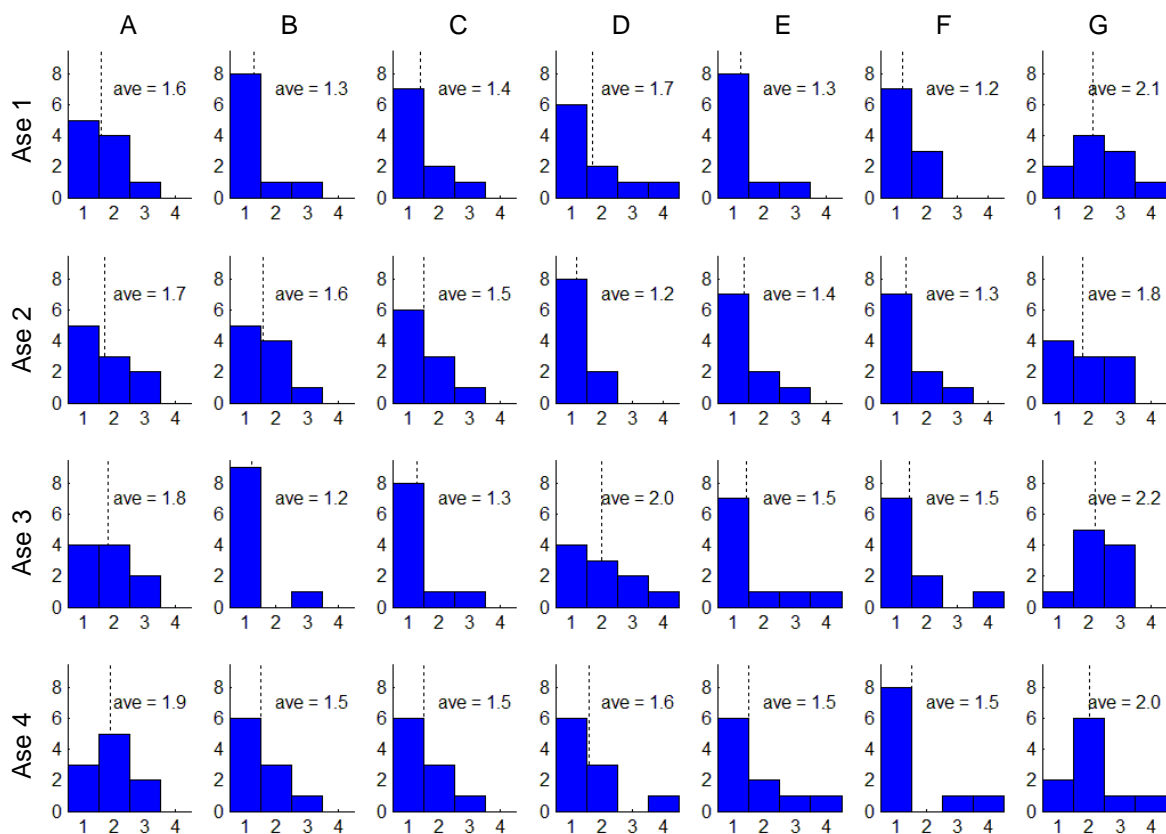
Taulukko 1: Asejärjestelmien suhteelliset yksikkökustannukset

Asejärjestelmä	Suhteellinen kustannus
1	1,04
2	1,12
3	1,00
4	1,29

Arviointeja varten järjestettiin kaksi työpajaa. Ensimmäisessä työpajassa selvitettiin kymmenen asiantuntijan avulla asejärjestelmien vaikuttavuuksia eri kohteiden suhteen. Asiantuntijat arvioivat jokaiselle kohteelle vauriotason, jonka eri asejärjestelmät niille todennäköisimmin aiheuttaisivat. Vauriotasojen arvioinnissa käytettiin neliportaista asteikkoa 1 – 4 siten, että arvo 1 tarkoittaa kohteen täydellistä tuhoutumista ja arvo 4 kohteen pysymistä vaurioitumattomana. Analysointia varten arviot muutettiin vaikuttavuusarvoiksi skaalaamalla niitä lineaarisesti välille 0 – 1 siten, että arvo 1 kuvaa kohteen täydellistä tuhoutumista.

Jälkimmäisessä työpajassa selvitettiin 19 asiantuntijan avulla kohteiden suhteellisia tärkeyksiä. Asiantuntijat arvioivat, miten tärkeitä tarkasteltavaan kohteeseen vaikuttaminen on suhteessa muihin kohteisiin. Arviointiperusteena toimi kohteiden potentiaalinen tuho vaikutus. Arviot tehtiin asteikolla 0 – 100 siten, että arvo 100 kuvaa tärkeintä kohdetta. Arviot skaalattiin lineaarisesti välille 0 – 1 siten, että jokaisen asiantuntijan arviot summautuvat yhteen ja että suurempi arvo kuvaa suurempaa tärkeyttä.

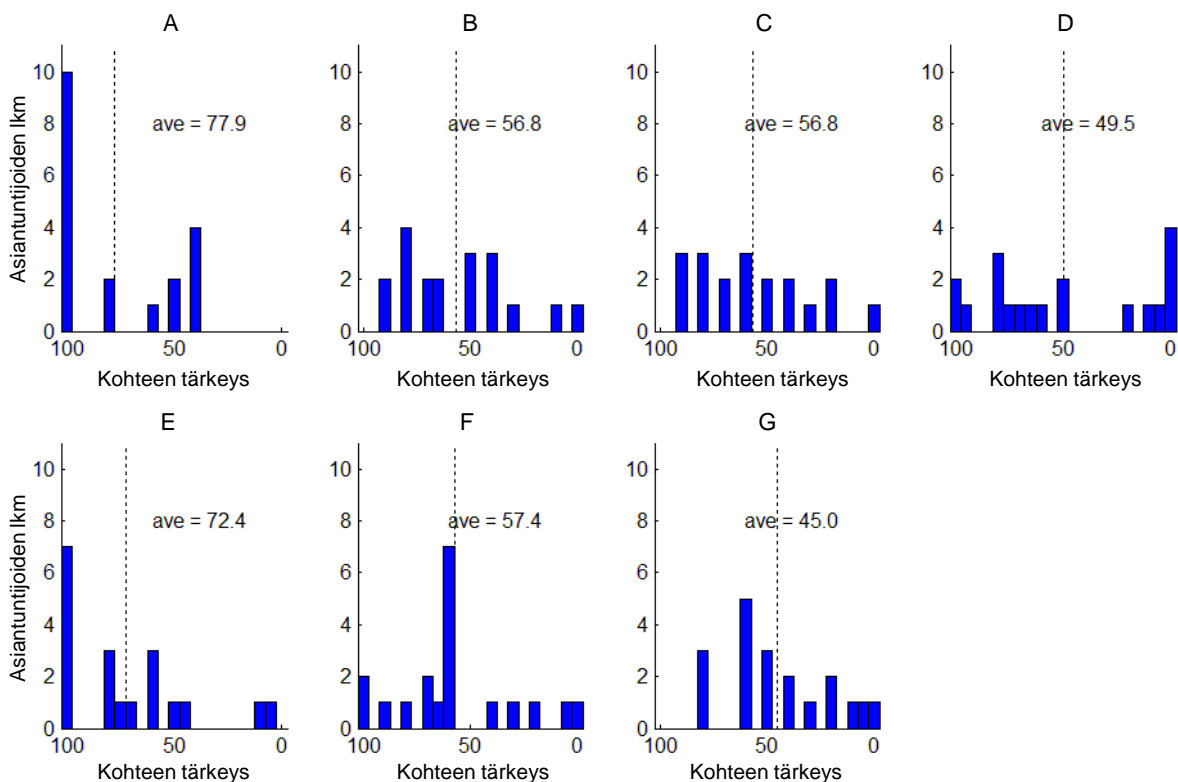
Molemmissa työpajoissa asiantuntijat edustivat ammattitaidoiltaan alansa parhaimmistoa. Asiantuntijat oli myös valittu vaihtelevilta aloilta, jolloin paneeleissa oli erilaista



Kuva 2: Vauriotasoarviointin tulokset. Sinisen palkin korkeus vastaa kyseisen arvion antaneiden asiantuntijoiden lukumäärää (max 10). Vaaka-akselilla vauriotasot (1 = kohde tuhoutuu, 4 = kohde ei vaurioidu). Katkoviivalla merkitty arvioiden keskiarvo (ave).

ammattitaitoa. Asiantuntijat tekivät arvionsa itsenäisesti ja heillä oli käytössään samat yksityiskohtaiset tiedot ammuttilanteista. Asiantuntijoita pyydettiin olemaan vastaamatta, jos he tunsivat itsensä kykenemättömiksi vastaamaan luotettavasti kysyttyihin asioihin. Muutama asiantuntija jätti vastaamatta.

Kuvissa 2 ja 3 on esitetty työpajojen tulokset histogrammien avulla. Molemmat kuvat osoittavat, että asiantuntijoiden arviot olivat varsin vaihtelevia. Arvioissa oli myös havaittavissa yhdenmukaisuusvajetta, sillä arviot menivät usein ristiin asiantuntijoiden välillä: toinen arvioi ase 1 eniten vaikuttavaksi ja ase 4 vähiten vaikuttavaksi ja joku päinvastoin.



Kuva 3: Tulokset kohteiden suhteellisten tärkeyksien arvioinnista. Sininen palkki kuvaa kyseisen arvion antaneiden asiantuntijoiden lukumäärää (max 19). Vaaka-akselilla tärkeysarvot (100 = tärkein, 0 = vähiten tärkeä). Katkoviivalla merkitty arvioiden keskiarvo (ave).

3.4 Asejärjestelmäyhdistelmien kustannustehokkuus

Asejärjestelmäyhdistelmien vertailussa käytetään kokonaisvaikuttavuutta, joka lasketaan kohdekohtaisten vaikuttavuuksien painotettuna summana. Painokertoimena käytetään kohteille arvioituja suhteellisia tärkeyksiä. Koska vaikuttavuusarviot ovat välillä 0 – 1 ja kohteiden suhteelliset tärkeydet summautuvat yhteen, kokonaisvaikuttavuuden vaihteluväliksi saadaan 0 – 1 siten, että 1 vastaa suurinta mahdollista kokonaisvaikuttavuutta.

Taulukossa 2 on esitetty esimerkki kokonaisvaikuttavuuden laskemisesta. Asiantuntija on arvioinut sekä kohteiden kokemat vauriotasot kyseisen asejärjestelmäyhdistelmän tapauksessa että kohteiden suhteelliset tärkeydet. Arviot on skaalattu edellä mainittuun tapaan. Vaikuttavuusarviot kerrotaan kohteittain skaalatuilla tärkeysarvioilla ja summaamalla näin saadut arvot yhteen saadaan kokonaisvaikuttavuus, joka esimerkin tapauksessa on 0,89.

Käyttämällä molempien työpajojen jokaista arviota saadaan mille tahansa yhdistelmälle laskettua $10 \times 19 = 190$ kokonaisvaikuttavuusarvoa. Tällä tavalla saadut kokonaisvaikuttavuusarvot voivat olla hyvinkin vaihtelevia. Oleellista on, että jos jokainen

Taulukko 2: Esimerkki kokonaisvaikuttavuuden laskemisesta

Kohde	A	B	C	D	E	F	G	Σ
Vauriotaso, (1-4)	2	1	1	1	1	1	3	
Vaikuttavuus, v (0-1)	0,67	1	1	1	1	1	0,33	
Tärkeys, (0-100)	80	60	50	100	40	30	20	380
Skaalattu tärkeys, w (0-1)	$\frac{80}{380} = 0,21$	0,16	0,13	0,26	0,11	0,08	0,05	1,00
$v \times w$	0,14	0,16	0,13	0,26	0,11	0,08	0,02	0,89
Kokonaisvaikuttavuus	0,14 + 0,16 + 0,13 + 0,26 + 0,11 + 0,08 + 0,02 = 0,89							

yksittäinen asiantuntija oletetaan päteväksi arviointitehtävässään, jokainen kokonaisvaikuttavuuden arvio on pätevä.

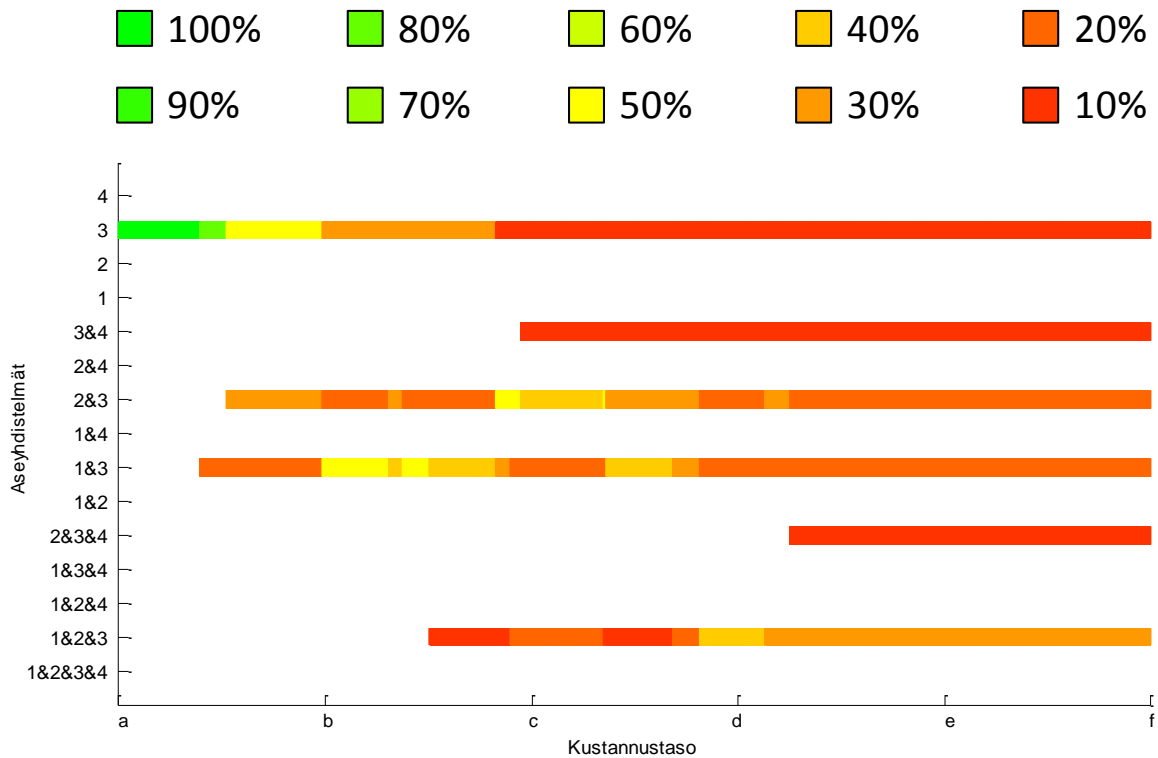
Tarkasteltavan esimerkin tapauksessa kaikkia kokonaisvaikuttavuusarvioita käytetään erillisinä hyväksi kustannustehokkuuden määrittämisessä. Tavoitteena on löytää sellaiset asejärjestelmäyhdistelmät, joille ei *minkään* arvioyhdistelmän mukaan olisi olemassa kokonaisvaikuttavuudeltaan suurempaa ja kustannuksiltaan korkeintaan yhtä kallista vaihtoehtoa. Tällainen yhdistelmä on määritelmän mukaan kustannustehokas. Toisin sanoen karsitaan vaihtoehtojen joukosta pois ne, joille on kaikkien mahdollisten arvioyhdistelmien mukaan olemassa kokonaisvaikuttavuudeltaan suurempi ja korkeintaan yhtä kallis vaihtoehto.

Tällaisella analyysillä löydetty kustannustehokkaat asejärjestelmäyhdistelmät ovat päätöksentekijän kannalta rationaalisesti järkeviä vaihtoehtoja. Päätöksentekijän ei kannata valita kustannustehotonta yhdistelmää, koska sille on olemassa vaihtoehto, joka (*i*) tuottaa saman suorituskyvyn halvemmalla tai (*ii*) tuottaa suuremman suorituskyvyn samalla kustannuksella tai (*iii*) tuottaa suuremman suorituskyvyn halvemmalla.

3.5 Kustannustehokkuusanalyysin tulokset

Arvioiden perusteella saatiin jokaiselle 256:lle asejärjestelmäyhdistelmälle määritettyä $10 \times 19 = 190$ kokonaisvaikuttavuusarvoa. Yhdistelmien joukosta selvitettiin kustannustehokkaat vaihtoehdot, joita oli 19 kappaletta eli noin 7,4% kaikista yhdistelmistä. Kun mahdollisia yhdistelmiä oli verraten vähän, kustannustehokkaat vaihtoehdot pystyttiin selvittämään pareittaisella dominanssivertailulla. Jos yhdistelmiä olisi huomattavasti suurempi määrä, menetelmää pitäisi kehittää laskennan nopeuttamiseksi.

Jäljelle jääneistä kustannustehokkaista yhdistelmistä pystytään tuottamaan päätöksentekoa tukevaa informaatiota esimerkiksi Robust Portfolio Modeling -menetelmässä käytetyn niin kutsutun ydinluvun muodossa (Liesiö ym., 2007). Asejärjestelmäyhdistelmälle laskettu ydinluku kuvaa sitä arviointitietojen osuutta, jonka perusteella yhdistelmä on tehokas tarkasteltavalla kustannustasolla. Ydinluku lasketaan yhdistelmille kaikilla

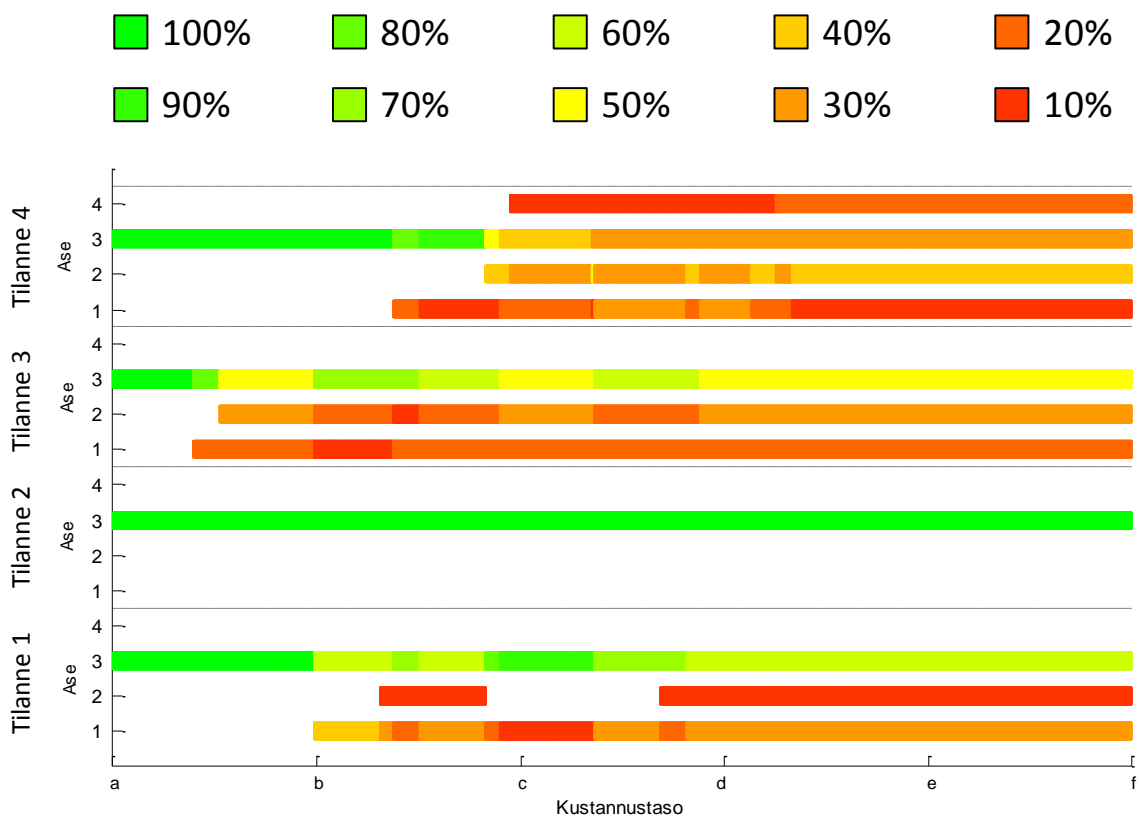


Kuva 4: Asejärjestelmäyhdistelmien tehokkuus kustannustasoittain. Väriskaala kuvaa sitä arviointitietojen osuutta, jolla yhdistelmä on tehokas kyseisellä kustannustasolla.

kustannustasoilla erikseen. Jos ydinluku on 100 %, yhdistelmä on tehokas kaikkien arvioiden valossa. Vastaavasti jos yhdistelmän ydinluku on 0 %, se ei ole tehokas minkään arviointitiedon perusteella.

Ydinlukua voidaan käyttää hyväksi esimerkiksi kuvan 4 mukaisesti. Kuvassa on havainnollistettu asejärjestelmien käyttöä jakamalla kustannustehokkaat yhdistelmät ryhmiin yhdistelmissä käytettyjen asejärjestelmien mukaan. Syntyy 15 ryhmää, joissa ei oteta kantaa, mihin kohteeseen vaikutetaan milläkin asejärjestelmällä. Jokaiselle ryhmälle voidaan jokaisella kustannustasolla laskea ydinluku, jota kuvataan kuvassa eri värein. Kustannusväli (a-f) on esitetty siten, että taso a on pienin kustannustaso, jolla voidaan hankkia halvin mahdollinen asejärjestelmäyhdistelmä ja tasolla f on jo mahdollista hankkia jokainen mahdollinen yhdistelmä. Kustannusväli (a-f) on jaettu yhtä suuriin osiin.

Vihreä väri tarkoittaa, että kyseinen asejärjestelmäyhdistelmä on kaikkien arvioiden mukaan tehokas kyseisellä kustannustasolla. Mitä punaisempi väri on, sitä harvempien arvioiden mukaan yhdistelmä on tehokas. Yhdistelmä on tehoton kaikkien arvioiden mukaan, jos sen kohdalla ei ole värillistä palkkia. Yhdistelmän tehokkuuden edellytyksenä on, että se on mahdollista hankkia kyseisellä kustannustasolla. Tämän takia kuvan pienimmillä kustannustasoilla (a) halvin asejärjestelmä 3 on ainoa



Kuva 5: Yksittäisten asejärjestelmien tehokkuus kustannustasoittain eri tilanteissa.

tehokas yhdistelmä, sillä mitään muita yhdistelmiä ei pystytä hankkimaan.

Kuvasta havaitaan, että asejärjestelmä 3 - halvin asejärjestelmä - on mukana kaikissa kustannustehokkaissa yhdistelmissä. Täten sen hankkiminen on hyvin perusteltua. Vastaavasti asejärjestelmiä 1, 2 ja 4 ei ole kustannustehokasta käyttää, jos käytettävissä ei ole muita asejärjestelmiä. Asejärjestelmä 4, joka on vaihtoehtoista kallein, on tehokas vain hyvin harvojen arviointilähtötietojen perusteella.

Edellisen tarkastelun perusteella ei vielä voida ottaa kantaa, missä ammuntatilanteissa mitään asejärjestelmiä tulisi käyttää. Kuvassa 5 on esitetty yksittäisten asejärjestelmien ydinluvut eri ammuntatilanteissa. Kuvasta havaitaan muun muassa, että ammuntatilanteessa 2 ainoa tehokas vaihtoehto on asejärjestelmä 3 kustannustasosta riippumatta. Kyseinen asejärjestelmä on myös suhteellisen kustannustehokas kaikissa muissa ammuntatilanteissa. Vastaavasti asejärjestelmä 4 on kustannustehoton kaikkien muiden paitsi kohteen 4 kohdalla. Asejärjestelmien 1 ja 2 kustannustehokkuuksissa ei ole selkeitä eroja.

4 Pohdinta

Esimerkin perusteella menetelmä näyttää soveltuvan hyvin myös tilanteisiin, joissa asiantuntijoiden arviot ovat varsin erilaisia. Esimerkin tapauksessa 256 vaihtoehdosta pystyttiin karsimaan pois 237 (noin 92,6%) tehottomina vaihtoehtoina, vaikka asiantuntija-arviot poikkesivat toisistaan. Arvioiden ollessa yhdenmukaisempia menetelmä karsisi kustannustehokkaiden yhdistelmien määrän esimerkkitapaustakin pienemmäksi.

Arvioiden poikkeavuuksien syitä voi olla useita. Molemmat työpajat koostuivat heterogeenisestä asiantuntijajoukosta, mikä tarkoittaa, että asiantuntijoiden arviointiperusteet saattoivat olla hyvinkin erilaisia. Tiedot asejärjestelmistä, ammunnan kohteista ja niiden merkityksestä voivat olla peräisin eri lähteistä.

Eräs merkittävimmistä suuren ja heterogeenisen asiantuntijapaneelin eduista on juuri se, että asiantuntemusta saadaan laajalta kentältä. Onnistunut heterogeenisen paneelin hyödyntäminen edellyttää kuitenkin sitä, että arviointikysymykset ja -menetelmät ovat tasapuolisia kaikille asiantuntijoille. Asiantuntijoilla tulee olla yhtäläinen käsitys arviointitilanteen tavoitteista ja kontekstista. Tähän voidaan kiinnittää huomiota panostamalla asiantuntijoiden perehdyttämiseen. Paneelin totetuttajalla on tässä suuri vastuu.

Menetelmän merkittävimpiä etuja on se, että se säilyttää yksittäiset asiantuntija-arviot kokonaisuuksina, jolloin asiantuntijakohtainen informaatio säilyy tuloksiin asti. Ei tapahdu informaatiohäviötä, kuten esimerkiksi keskiarvoistamista käyttämällä. Tulosten esittäminen ja perustelu yksinkertaistuu, kun voidaan todeta, että ”jokaisen yksittäisen asiantuntijan mielestä nämä vaihtoehdot ovat tehottomia ja siten pitäisi jättää harkinnan ulkopuolelle”.

Menetelmän haasteet liittyvät epävarmuuden mallintamiseen. Esimerkissä ei käsitelty yksittäisten asiantuntijoiden luotettavuutta tai heidän arvioidensa epätarkkuutta. Epätarkkuuden mallintaminen on oleellinen piirre asiantuntijoiden järkevässä hyödyntämisessä ja menetelmään voisikin tulevaisuudessa lisätä epävarmuutta kuvaavan mittarin, esimerkiksi Cooken (1991) klassisen mallin tapaan.

Tällaisenaan menetelmä ohjaa merkittävästi päätöksentekoprosessia jättäen kuitenkin lopullisen vastuun päätöksentekijälle. Menetelmä soveltuu varsinkin tilanteisiin, joissa vaihtoehtoisten yhdistelmien suurta määrää on karsittava paljon. Jäljelle jääneistä kustannustehokkaista yhdistelmistä saadaan arvokasta lisätietoa ydinluvun avulla. Hyödyllistä on myös, että tehokkuustarkastelu tehdään kaikilla järkevillä kustannustasoilla. Tällöin on mahdollisuus saavuttaa merkittäviä säästöjä, jos havaitaan, että kustannusten lisääminen ei muuta päätössuosittelusta ja siten saavutetaan pienemmällä panoksella sama vaikuttavuus kuin suuremmalla.

5 Yhteenveto

Työn ensimmäisen osan kirjallisuuskatsauksessa tarkasteltiin keskeisimpiä menetelmiä asiantuntijapaneelin muodostamiseen, asiantuntija-arvioiden määrittämiseen ja niiden analysointiin liittyen. Todettiin, että yleisin arvioiden hyödyntämistapa on niiden yhdistäminen yhdeksi jakaumaksi joko painotetun summan tai Bayesin kaavalla. Katsauksessa käsiteltiin molemmille menetelmille esitettyä kritiikkiä ja esiteltiin esimerkinomaisesti hyvin käyttökelpoiselta vaikuttava muunnelma painotetun summan menetelmästä, Cooken (1991) klassinen malli.

Työn jälkimmäisessä osassa esiteltiin eräs kirjallisuudesta poikkeava arviointitietojen hyödyntämismenetelmä esimerkin kautta. Esimerkki perustuu todelliseen tilanteeseen, jossa tuettiin päätöksentekoa asejärjestelmäyhdistelmän hankintaan liittyen. Tässä esitellyt menetelmä eroaa yleisemmistä menetelmistä siten, että siinä ei pyritä yhdistämään asiantuntija-arvioita yhdeksi jakaumaksi esimerkiksi keskiarvon avulla, vaan käytetään kaikkia yksittäisiä arvioita erillisinä. Menetelmässä selvitetään sellaiset vaihtoehdot, jotka ovat kaikkien arviointitietojen perusteella järkeviä. Näistä järkevistä vaihtoehdoista pystytään lisäksi tuottamaan hyödyllistä lisäinformaatiota esimerkiksi työssä esitellyn ydinluvun muodossa.

Menetelmän haasteet liittyvät keskeisesti asiantuntija-arvioiden vaihtelevuuteen ja tämän seurauksiin. Menetelmää voisi tulevaisuudessa kehittää esimerkiksi arvioimalla ja ottamalla huomioon yksittäisten asiantuntijoiden luotettavuus ja heidän arvioidensa epävarmuus. Tämän tiedon perusteella voisi luoda useita analyysejä arvioiden luotettavuuden suhteen.

6 Viitteet

- Clemen, R., Winkler, R., 1999. Combining probability distributions from experts in risk analysis. *Risk Analysis*, 19(2), 187-203.
- Cooke, R., 1991. *Experts in Uncertainty: Opinion and Subjective Probability in Science*, Oxford University Press.
- Cooke, R., Goossens, H., 2004. Expert judgement elicitation for risk assessment of critical infrastructures. *Journal of Risk Research*, 7(6), 643-656.
- Cooke, R., Goossens, H., 2008. TU Delft expert judgment data base. *Reliability Engineering and System Safety*, 93(5), 657-674.
- Dalkey, N., Helmer, O., 1963. An experimental application of the Delphi method to the use of experts. *Management Science*, 9(3), 458-467.
- Hoffmann, S., Fischbeck, P., Krupnick, A., McWilliams, M., 2007a. Elicitation from large, heterogeneous expert panels: Using multiple uncertainty measures to characterize information quality for decision analysis, *Decision Analysis*, 4(2), 91-109.
- Hoffmann, S., Fischbeck, P., Krupnick, A., McWilliams, M., 2007b. Using expert elicitation to link foodborne illnesses in the United States to foods. *Journal of Food Protection*, 70(5), 1220-1229.
- Kalton, G., Anderson, D., 1986. Sampling rare populations. *Journal of the Royal Statistical Society*, 149(1), 65-82.
- Keeney, R., von Winterfeldt, D., 1989. On the uses of expert judgment on complex technical problems. *IEEE Transactions on Engineering Management*, 36(2), 83-86.
- Lichtenstein, S., Fischhoff, B., 1977. Do those who know more also know more about how much they know? *Organizational Behavior Human Performance*, 20(2), 159-183.
- Lin, S.-W., Bier, V., 2008. A study of expert overconfidence. *Reliability Engineering and System Safety*, 93(5), 711-721.
- Liesiö, J., Mild, P., Salo, A., 2007. Preference programming for robust portfolio modeling and project selection. *European Journal of Operational Research*, 181(3), 1488-1505.
- Morris, P., 1974. Decision analysis expert use. *Management Science*, 20(9), 1233-1241.
- Russo, J., Schoemaker, P., 1992. Managing overconfidence. *Sloan Management Review*, 33(2), 7-17.
- Sandri, S., Dubois, D., Kalfsbeek, H., 1995. Elicitation, assessment, and pooling of expert judgments using possibility theory. *IEEE Transactions on fuzzy systems*, 3(3), 313-335.

7 Liitteet

7.1 Mallin matemaattinen formulointi

Olkoon tarkasteltavien järjestelmien lukumäärä I ja kohteiden lukumäärä J .

Olkoon $p_{i,j} = 1$, jos järjestelmää i käytetään kohteeseen j ja muuten $p_{i,j} = 0$. Tällöin järjestelmäyhdistelmiin voidaan viitata matriisiin \mathbf{P} avulla seuraavasti:

$$\mathbf{P} = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,J} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,J} \\ \vdots & \vdots & \ddots & \vdots \\ p_{I,1} & p_{I,2} & \cdots & p_{I,J} \end{bmatrix} \in \{0, 1\}^{I \times J} \quad (1)$$

Asetetaan rajoitus, että yksittäiseen kohteeseen voidaan käyttää ainoastaan yhdenlaista järjestelmää kerrallaan:

$$\sum_i p_{i,j} = 1 \quad \forall j = 1, \dots, M. \quad (2)$$

Olkoon $c_{i,j}$ järjestelmän i kustannukset kohteeseen j . Tällöin järjestelmäyhdistelmän kokonaiskustannukset ovat

$$C(\mathbf{P}) = \sum_{i=1}^I \sum_{j=1}^J c_{i,j} p_{i,j}. \quad (3)$$

Määritetään käypien yhdistelmien joukko \mathcal{P}_F kustannustasolla B seuraavasti:

$$\mathcal{P}_F = \{\mathbf{P} \in \{0, 1\}^{I \times J} \mid C(\mathbf{P}) \leq B, \sum_{i=1}^I p_{i,j} = 1 \quad \forall j = 1, \dots, J\}. \quad (4)$$

Olkoon järjestelmien aiheuttamia vauriotasoja arvioivien asiantuntijoiden lukumäärä K . Olkoon ${}^0v_{i,j}^k \in \{1, 2, 3, 4\}$ asiantuntijan k antama arvio järjestelmän i aiheuttamasta vauriotasosta kohteeseen j siten, että arvo 1 kuvaa suurinta vauriotasoa. Skaalataan nämä arviot vaikuttavuusarvioiksi välille $[0, 1]$ seuraavasti:

$$v_{i,j}^k = \frac{1}{3}(4 - {}^0v_{i,j}^k) \in [0, 1] \quad \forall k = 1, \dots, K; \quad i = 1, \dots, I; \quad j = 1, \dots, J. \quad (5)$$

Merkitään matriisillä \mathbf{v}^k asiantuntijan k antamia arvioita kaikkien järjestelmien ja kohteiden suhteen:

$$\mathbf{v}^k = \begin{bmatrix} v_{1,1}^k & v_{1,2}^k & \dots & v_{1,J}^k \\ v_{2,1}^k & v_{2,2}^k & \dots & v_{2,J}^k \\ \vdots & \vdots & \ddots & \vdots \\ v_{I,1}^k & v_{I,2}^k & \dots & v_{I,J}^k \end{bmatrix} \in [0, 1]^{I \times J}. \quad (6)$$

Merkitään kaikkien vaikuttavuusarvioiden joukkoa $S_v = \{\mathbf{v}^1, \dots, \mathbf{v}^K\}$.

Olkoon kohteiden suhteellisia tärkeyksiä arvioivia asiantuntijoita L kappaletta. Olkoon ${}^0w_j^l \in [0, 100]$ asiantuntijan l antama arvio kohteen j suhteellisesta tärkeydestä. Skaalataan arviot välille $[0, 1]$ siten, että niiden summa on 1 seuraavasti:

$$w_j^l = \frac{{}^0w_j^l}{\sum_{j^*=1}^J {}^0w_{j^*}^l} \quad \forall l = 1, \dots, L; \quad j = 1, \dots, J. \quad (7)$$

Merkitään asiantuntijan l antamia skaalattuja tärkeysarvioita kaikkien kohteiden suhteen vektorilla $\mathbf{w}^l = [w_1^l \dots w_J^l] \in [0, 1]^{1 \times J}$ ja kaikkien tärkeysarvioiden joukkoa $S_w = \{\mathbf{w}^1, \dots, \mathbf{w}^L\}$.

Järjestelmäyhdistelmälle \mathbf{P} määritetään kokonaisvaikuttavuus V kaikkien asiantuntijaarviointien suhteen seuraavasti:

$$V^{k,l}(\mathbf{P}) = V(\mathbf{P}, \mathbf{v}^k, \mathbf{w}^l) = \sum_{j=1}^J w_j^l \sum_{i=1}^I p_{i,j} v_{i,j}^k \quad \forall k = 1, \dots, K; \quad l = 1, \dots, L. \quad (8)$$

Määritetään kaikkien mahdollisten arvioiyhdistelmien joukko $S = S_v \times S_w$. Kustannustehokkaiden yhdistelmien joukko on tällöin

$$\mathcal{P}_E(S) = \{\mathbf{P} \in \mathcal{P}_F \mid \nexists \mathbf{P}' \in \mathcal{P}_F \text{ jolla } \left\{ \begin{array}{l} V^{k,l}(\mathbf{P}') \geq V^{k,l}(\mathbf{P}) \quad \forall (\mathbf{v}^k, \mathbf{w}^l) \in S \\ C(\mathbf{P}') \leq C(\mathbf{P}) \end{array} \right\}\}, \quad (9)$$

missä vähintään yksi vahva epäyhtälö pätee jollekin $(\mathbf{v}^k, \mathbf{w}^l) \in S$.

Kaikille yhdistelmille määritetään eri kustannustasoilla ydinluku CI , joka kertoo, kuinka suuren arviotietojen osan perusteella yhdistelmä $\mathbf{P} \in \mathcal{P}_F$ on tehokas tarkasteltavalla kustannustasolla B :

$$CI(\mathbf{P}, B) = \frac{|\{(\mathbf{v}^k, \mathbf{w}^l) \mid \mathbf{P} \in \mathcal{P}_E(S)\}|}{|S_v \times S_w|} \cdot 100\%, \quad (10)$$

missä merkintä $|\{\cdot\}|$ tarkoittaa joukon sisältämien alkioiden lukumäärää.